



PhD-FSTM-2022-025
The Faculty of Sciences, Technology and Medicine

DISSERTATION

Defense held on 30/03/2022 in Luxembourg
to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG
EN PHYSIQUE

by

Emanuele Penocchio
Born on 10/02/1992 in Manerbio (Brescia), Italy

THERMODYNAMICS OF CHEMICAL ENGINES:
A CHEMICAL REACTION NETWORK APPROACH

Dissertation defense committee:

Dr. Massimiliano Esposito (supervisor)
Professor, Université du Luxembourg

Dr. Étienne Fodor
Professor, Université du Luxembourg

Dr. Thomas Hermans
*Professor, Université de Strasbourg, Institut de Science et d'Ingénierie
Supramoléculaires (ISIS)*

Dr. David Sivak
Professor, Simon Fraser University

Dr. Alexander Skupin
*Professor, Université du Luxembourg, Luxembourg Centre for Sys-
tems Biomedicine (LCSB)*

Thermodynamics of Chemical Engines: A Chemical Reaction Network Approach



Emanuele Penocchio

ABSTRACT

Chemical processes in closed systems inevitably relax to equilibrium. Energy can be employed to counteract such tendency and drive reactions against their spontaneous direction. This nonequilibrium driving is implemented in open systems, which living organisms provide the most spectacular examples of. In recent years, experiments in supramolecular chemistry, photochemistry and electrochemistry demonstrated that, by opening synthetic systems to matter and/or energy exchanges with the environment, artificial systems with life-like behaviours can be realized and used to convert energy inputs of different nature into work at both the nanoscopic and the macroscopic level. However, one tool that is still lacking is a firm grasp of the thermodynamics of these *chemical engines*. In this thesis, we provide it by leveraging the most recent developments of the thermodynamic description of deterministic chemical reaction networks. As main theoretical results, we extend the current theory to encompass nonideal and light-driven systems, thus providing the fundamental tools to treat electrochemical and photochemical systems in addition to the chemically driven ones. We also expand the scope of information thermodynamics to bipartite chemical reaction networks characterized by macroscopic non-normalized concentration distributions evolving in time with nonlinear dynamics. This framework potentially applies to almost every synthetic chemical engine realized until now, and to many models of biological systems too. Here, we undertake the thermodynamic analysis of some of the epitomes in the field of artificial chemical engines: a model of chemically driven self-assembly, an experimental chemically driven molecular motor, and an experimental photochemical bimolecular pump. The thesis provides a thermodynamic level of understanding of chemical engines that is general, complements previous analyses based on kinetics and stochastic thermodynamics, and has practical implications for designing and improving synthetic systems, regardless of the particular type of powering or chemical structure.

CONTENTS

List of publications	vii
Acknowledgements	ix
1 INTRODUCTION	1
1.1 Steam engines and the birth of thermodynamics	2
1.1.1 Aeolipile	2
1.1.2 Pumps	3
1.1.3 Rotary and linear work	4
1.1.4 Classical Thermodynamics	5
1.1.5 Final considerations	10
1.2 Chemical Engines	12
1.2.1 Molecular aeolipiles	13
1.2.2 Molecular pumps	15
1.2.3 Rotary and linear macroscopic work done by chemical engines	20
1.2.4 Systems Chemistry	20
1.2.5 Final considerations	22
1.3 Thermodynamics of chemical systems	28
1.3.1 Mass action kinetics	28
1.3.2 Chemical potentials	30
1.3.3 Gibbs free energy	32
1.3.4 Local equilibrium	33
1.3.5 Local Detailed Balance	35
1.3.6 Open systems	37
1.3.7 Chemical reaction networks	39
1.3.8 Stochastic thermodynamics	39
1.3.9 Final considerations	41
1.4 Contributions of this thesis	47
2 THE CHEMICAL REACTION NETWORK APPROACH	49
Article: [<i>Journal of Chemical Physics</i> 154 , 094114 (2021)]	50
Article: [<i>Nature Communications</i> 10 , 3865 (2019)]	64
3 PHOTOCHEMICAL ENGINES	81
Article: [<i>Journal of Chemical Physics</i> 155 , 114101 (2021)]	82
Article: [<i>ChemRxiv</i> (2021)]	97
4 TOWARDS AN INFORMATION THERMODYNAMIC PERSPECTIVE	129
Article: [<i>Manuscript</i> (2022)]	130
Article: [<i>ChemRxiv</i> (2021)]	149
5 CONCLUSIONS	181

LIST OF PUBLICATIONS

- 1) G. FALASCO, T. COSSETTO, E. PENOCCHIO and M. ESPOSITO, *New J. Phys.* **21**. (2019), 073005.
- 2) E. PENOCCHIO, R. RAO and M. ESPOSITO, *Nat. Commun.* **10**. (2019), 3865.
- 3) A. SABATINO, E. PENOCCHIO, G. RAGAZZON, A. CREDI and D. FREZZATO, *Angew. Chem. Int. Ed.* **58**. (2019), 14341.
- 4) F. AVANZINI, E. PENOCCHIO, G. FALASCO and M. ESPOSITO, *J. Chem. Phys.* **154**. (2021), 094114.
- 5) S. AMANO, M. ESPOSITO, E. KREIDT, D. A. LEIGH, E. PENOCCHIO and B. M. W. ROBERTS, *Nat. Chem.* **14**. (2022), 530–537.
- 6) E. PENOCCHIO, R. RAO and M. ESPOSITO, *J. Chem. Phys.* **155**. (2021), 114101.
- 7) S. CORRA, M. T. BAKIĆ, J. GROPPi, M. BARONCINI, S. SILVI, E. PENOCCHIO, M. ESPOSITO and A. CREDI, *Nature Nanotechnology* (2022).
- 8) E. PENOCCHIO, F. AVANZINI and M. ESPOSITO, *arXiv* **2204.02815**. (2022), Accepted in *The Journal of Chemical Physics*.

I contributed to all the aspects of research and publication in 2,4,5,6, and 8, and developed the theoretical modelling in 7. I contributed to devise the research in 1 and 3, and performed numerical simulations for 1. Publications 1 and 3 are not included as chapters of this thesis.

All my work was funded by the European Research Council project NanoThermo (ERC-2015-CoG Agreement No. 681456).

ACKNOWLEDGEMENTS

*There was an overused expression at that time: “a good comrade”.
And when is someone a good comrade?
You say that, but you are never fully aware of what you are saying.
I have come to my own definition, here it is:
if a good comrade were to design the world, I would live well in it.*

— Luca Rastello, *Piove all’insù*
(Bollati Boringhieri, 2006. Free translation from Italian)

I have been questioning myself on how to become a *good scientist* for a long time. Even more after meeting excellent examples of skilled people who are far from representing an idea of science I am willing to embrace. Over the last four and a half years, what I have been most grateful for has been sharing my everyday life with outstanding colleagues: if they were to design the scientific community, I would live well in it. You have been the good scientists I needed to meet. Without you, I would have lacked the necessary motivation to put so many efforts into this thesis – you now know why I was so excited to organize our group meetings!

Moving to the contents of my doctoral dissertation, I shall start by thanking Massimiliano, who has been encouraging and guiding my ideas since we first met. Such ideas would have never come to me without the massive number of discussions with Giulio, an invaluable source of inspiration throughout all my academic career. Such discussions would have never been translated into formulas without all the chalk sacrificed by Riccardo to teach me thermodynamics. I also explicitly thank Francesco, the person with whom I worked best in these years, together with Gianmaria and Tommaso, with whom I shared the best physics – and physical – activities.

The present thesis has also benefited from two precious experimental collaborations. One with Ben, Dave, Elisabeth, and Shuntaro, which will probably remain the most incredible birthday present I have ever received. Undoubtedly, understanding each other properly was never easy, but we managed to chisel our Rosetta Stone in the end. The other one with Alberto, Jessica, Marina, Massimo and Stefano, which has been a pleasant blast from my university years in Bologna. Particular thanks go to Stefano for our Friday not-too-early morning meetings and for the rich exchanges we had.

Among the good scientists I got to know over the years, I cannot help but mention Diego, who has unconditionally supported me even when it was something cumbersome to do. I also hope not to have made Yolande’s life too difficult with my bureaucratic incompetence, and Daniele’s and Massimo’s ones with my tutoring inexperience.

It is truly impossible to fully thank and mention all the people who have indirectly contributed in so many different ways to make my doctoral studies such an amazing period of personal growth. I will just list the few ones to whom I think I owe more. My family and Danilo, who has been a colleague, a friend, a flatmate, and finally even beat my mum in the fastest birthday wishes competition. Oliviero and Tommaso for our meta-scientific discussions, and all the eXtemporanea community for always trying to go *against method* together. Claudio and Lorenzo because we never literally walk alone, except when chased by starved Neapolitan mastiffs in the middle of Sicily.

A final word to Giulia, my best comrade in the last two years, to whom this thesis is explicitly NOT dedicated, as I have broken up with all the partners I have ever mentioned in dedications. Good scientists are probably not superstitious, but I have good reasons to be so.

The cover image of this thesis represents Arnath, the demon of irreversibility, and it was drawn by the cartoonist Zerocalcare. I also thank him for keeping me company with his books and for always reminding us that life is almost never just about tearing along the dotted line. Sometimes, as the vicissitudes of my previous doctoral studies testify, dropping out of the dotted line may even lead to better results.

Luxembourg, 18th February 2022

Emanuele Penocchio

1

INTRODUCTION

The only pure myth is the idea of a science devoid of all myth.

— Michel Serres in *We have never been Modern* by Bruno Latour
(Harvard University Press, 1993. Original version in French: 1974)

The goal of this thesis is to expand the scope of thermodynamics towards – and by – the study of those chemical systems which exploit energy to undergo non-spontaneous processes. Such systems will be generically called *chemical engines*, in analogy to steam engines or electrical engines. Chemical engines are ubiquitous in living organisms, where they enable phenomena requiring energy and information processing.

Understanding the detailed functioning of life in its distinct aspects such as energetics is one of the greatest challenges of contemporary science. However, while of general scope in its theoretical part, this thesis is exclusively concerned with the thermodynamics of artificial chemical engines in its applied part. Examples of this kind of systems are found among synthetic molecular motors and machines, which earned the 2016 Nobel Prize in Chemistry to some of the key players in the field.

The choice of focusing on the field of artificial chemical engines has three main motivations. The first – personal – motivation is the same as the one that George Mallory (1886-1924) gave when he was asked to justify his efforts for first ascending Mount Everest: “because it’s there”. With this, spontaneous fascination for a research topic is claimed as a legitimate reason to undertake doctoral studies. The second – methodological – motivation, is that many explanations have been proposed to disclose the working principles of biological chemical engines, but the complexity of living systems makes it hard to test them and to tell the more insightful ones apart. This is not the case for synthetic systems, which have the great advantage that their functioning can be fully described by relatively simple models. Furthermore, these models can be accurately tested and used to improve the design of experimentally synthesized systems. They thus provide an ideal intermediate platform to go from the current understanding of the thermodynamics of chemical systems to its fruitful application in the life sciences. In particular, pursuing a comprehensive description of artificial chemical engines grounded on thermodynamics is a way to probe the status of the theory and to pinpoint what is missing. In the light of this, we see the choice to study artificial chemical engines as a necessary – or, at least, a convenient – step for advancing thermodynamics. The third – technological – motivation, is inspired by the one that the Royal Swedish Academy of Sciences gave when awarding the 2016 Nobel Prize in Chemistry: “in terms of development, the molecular motor is at the same stage as the electric motor was in the 1830s, when scientists displayed various spinning cranks and wheels, unaware that they would lead to washing machines, fans and food processors. Molecular machines will most likely be used in the development of things such as new materials, sensors and energy storage systems.” Here, we extend the comparison in both scope and time by arguing that the current situation of chemical engines in general shares some similarities also

with the one of steam engines at the beginning of the 19th century. As such, they are calling for a firm grasp of the thermodynamic principles underlying their functioning, which can significantly contribute to their rational development.

This introduction is organized in 4 sections aimed at contextualizing the thesis with respect to past and contemporary research areas, and at providing a self-contained presentation of the main ideas and results which the following chapters are based on. In Section 1.1, we discuss some of the milestones in the historical development of steam engines and show how the fundamental ideas of classical thermodynamics emerged. The conceptual framework and the so-called *laws* of thermodynamics are presented. In Section 1.2, a selection of the most relevant experiments in the field of artificial chemical engines is discussed, some of which will be taken as case studies in the following chapters. Our definition of chemical engines is further elaborated through examples and the third motivation point given above is substantiated. In Section 1.3, the most relevant concepts and tools from chemical thermodynamics are reviewed with the goal to provide a simple but complete introduction to the theory used throughout the thesis. In view of the current status of the theory, the next chapters are presented and contextualized in Section 1.4.

A word of caution is due before continuing. Notwithstanding some efforts towards giving a comprehensive introduction, this chapter reflects the *myths* underlying this thesis, namely the ideas of the author and the scientific tradition of the group left their marks on how the content is organized and presented. Recognizing that this contamination is unavoidable, we hope it will not be perceived as a drawback but rather as an asset to place the research in its rightful context.

1.1 STEAM ENGINES AND THE BIRTH OF THERMODYNAMICS

Science owes more to the steam engine than the steam engine owes to science.

— Lawrence Joseph Henderson (attributed to)

How to move objects of dimensions and at velocities way beyond human capabilities has been a central technological challenge of any society since the dawn of time. The exploitation of animal force in plowing and transports or the use of water and wind flows to actuate mills largely amplified the scope of humankind. For millennia, all sources of power available to humanity shared a common crucial feature: they are naturally in motion. In view of the latter consideration, a huge conceptual breakthrough in the history of technology has been the idea of an engine as a device to produce mechanical motion or work out of a power source which is naturally static, like burning wood or coal.

1.1.1 Aeolipile

One of the first recorded examples of such an engine is the aeolipile, also known as Hero's Engine, from Hero of Alexandria, who probably lived in the 1st century [1, 2]. As sketched in Figure 1, it basically consists of a hollow sphere with two asymmetric blowholes which starts spinning when

steam is generated inside the sphere itself. It has to be noted that the aeolipile operates autonomously, i.e., without the need of external intervention once the fire is turned on.

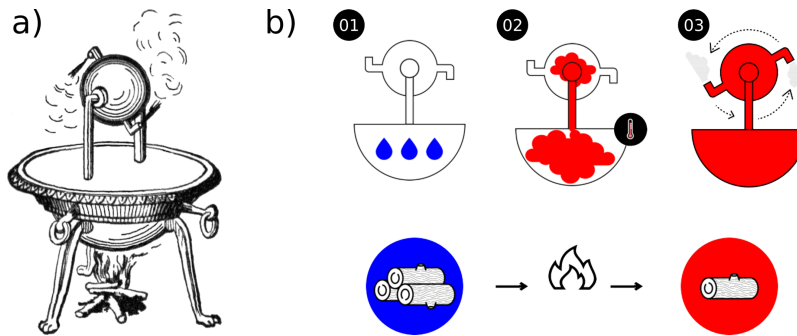


Figure 1: The aeolipile. a) An illustration of Hero's aeolipile (public domain); b) Schematic of the engine starting: water is heated up at the expense of fire to generate steam, which reaches the hollow sphere and makes it spinning by exiting through the two asymmetric blowholes. For an anachronistic but effective demonstration see <https://youtu.be/AyFMexBTpLs>.

Machines like the aeolipile were described by Hero mostly as toys, and used in public shows as entertainments. However, Hero's use of steam may also be understood as an attempt to explore new energy sources in the context of a society which already mastered hydraulic power and was actively seeking out energy sources independent of geographic location. Some authors even hypothesized that steam technology was quite widespread in Hellenistic society, also based on the argument that the level of technology developed in Hero's works is too complex to have been conceived by a single person, but no direct proofs have been found [2].

It should be stressed that modern steam engines like those mentioned below are not unrelated to the aeolipile [3]. Hero's works were extensively studied in the following centuries and the idea of using steam as motor power is for instance recovered in 1601 by Giovanni Battista Della Porta in his *"Pneumaticorum libri tres"*, directly based on Hero's *Pneumatica* [4]. Moreover, in 1615 Salomon de Caus (1576-1626) used steam to power a fountain, a kind of device which has been widely described by Hero [5].

1.1.2 Pumps

A considerable leap in the interest and scope of heat engines came with the transition, in European countries, from an economy based on mercantilism to industrial capitalism, and the subsequent incentive to constantly foster the production of goods by any means, technology included [6]. The first industrially relevant and established application of a steam engine is mainly due to Thomas Newcomen (1664-1729), who in 1712 significantly improved a pump invented by Thomas Savery (1650-1715) which was commercialized as "the Miner's friend" [7]. Newcomen's engine operates with two strokes: a piston inside a cylinder is first pushed upwards by the steam generated in a boiler, and subsequently pushed downwards by atmospheric pressure when the steam condensates due to injection of cold water inside the cylinder. Thanks to valves, the operation cycle can be repeated autonomously, contrary to Savery's machine which required constant external intervention instead. The main application of Newcomen's engine has been sucking wa-

ter from the bottom of coal mines, where the availability of coal as fuel without the need of transporting it was crucial to make the use of the pump convenient. Hundreds of samples were constructed throughout the 18th century and sold in Britain and Europe, thus quadruplicating mining activity [8].

The key improvement which made it possible to use Newcomen's pump and the subsequent steam engines in places far from coal mines such as factories has been the spatial separation of the heat source from the cooling system of the steam [9]. Indeed, in Newcomen's pump the steam is generated and heated up in the same cylinder where it is cooled down and condensed by the addition of water to complete a stroke. This mechanism has the major drawback that the whole cylinder has to be heated up during each operation cycle, thus considerably limiting the performance. If the steam, once expanded in the cylinder thanks to the heat source, is cooled down in a separate condenser, the power needed to heat up the cylinder is not lost at each stroke, and the performance of the engine is improved. This fundamental idea came to James Watt (1736-1819) in 1765, working at the time as the "Mathematical Instrument Maker" at the University of Glasgow, while trying to figure out why a scale didactic model of Newcomen's pump couldn't function properly [8]. The reason turned out to be exactly that in a smaller cylinder heat losses are so amplified that Newcomen's original mechanism is not applicable at small scales. Such an innovation allowed Watt, in collaboration with the industrialist Matthew Boulton (1728-1809), to produce pumps outperforming Newcomen's ones and to start a very successful business as an engine craftsman in Birmingham. Steam powered pumps stopped to be "just" tools to dry off the bottom of coal mines and started to play a crucial role in storing energy in the form of water tanks raised with respect to the ground, which could subsequently be used to actuate mills even in the absence of natural sources of hydraulic power.

1.1.3 Rotary and linear work

A second major contribution from Watt has been the invention in 1781 of a method to convert reciprocating motion of a piston to rotating motion. This, together with other technical improvements, opened the way to the direct use of steam for a plethora of industrial applications. The rotary engine replaced watermills, windmills and horses as the main power sources of the time, thereby freeing industries from geographical constraints and becoming one of the main agents in the so-called Industrial Revolution. Boulton and Watt were creative in marketing too: starting from the idea that the main good they were selling was power, instead of asking money for the construction and installation of a machine their contracts dictated that for a period of five years the buyers would have paid them "only" one third of the difference between the cost of the coal necessary to power the steam engine and the cost of the fodder they would have needed to feed horses delivering the same amount of work. This kind of deals obviously raised the question of how much work a horse can do, and in 1782 Watt himself combined experiments and data from a sawmill to establish the unit of horsepower (Hp) as the average amount of work a horse can deliver in one minute. At their best, Watt's engines could deliver about 20 Hp [8].

The availability of rotary motors able – contrary to the aeolipile – to deliver decent amounts of power stimulated the idea to use them to move vehicles. However, Watt's engines were too big and heavy for that scope.

The key ingredient to build smaller but more powerful steam engines consisted in increasing the working pressure of the steam, which became possible thanks to improvements in boiler technology. Richard Trevithick (1771-1833), a mining engineer, developed the first high pressure steam engine, while George Stephenson (1781-1848), a self-taught mechanic, in 1825 built the first steam locomotive successfully carrying passengers on a public rail line, the Stockton and Darlington Railway [8]. This marked the supremacy of the steam engine over water, wind and horses not only regarding the amount of delivered power, but also regarding the velocity at which objects can be moved.

1.1.4 Classical Thermodynamics

It should be clear from the above discussion that the development of the steam engine was mostly a technological rather than a scientific enterprise, at least to some extent¹. In particular, machines were invented and improved without an underlying scientific theory able to explain and predict their functioning, or quantifying their performances. With the words of Sadi Carnot (1796-1832), an engineer in the French army and alumnus of the recently established *École Polytechnique* in Paris, as well as son of the physicist and politician Lazare Carnot (1753-1823): “notwithstanding the work of all kinds done by steam-engines, notwithstanding the satisfactory condition to which they have been brought today, their theory is very little understood, and the attempts to improve them are still directed almost by chance” [10].

Actually, it can be argued that the opposite happened: namely, the tremendous technological innovation brought by the steam engine was the spark of the new scientific theory which we now call thermodynamics. We already mentioned that Watt introduced the unit of horsepower to quantify the performance of his engines, but even the current notion of power itself was introduced by John Smeaton (1724-1792) in 1759 to quantify the work per unit of time done by a pump lifting buckets of water, therefore quantifying the energy stored [11]. Also, it was the desire to push engines technology to its limits for the good of the French nation [12] that pushed Sadi Carnot to undertake his paramount theoretical “reflections on the motive power of fire” [10] in 1824.

Even though at the beginning of the 19th century heat was thought of as a real substance – the caloric – falling from hot bodies to cold ones akin to a fall of water powering hydraulic engines, there were no doubts about its enormous motive power and the fact that “to heat also are due the vast movements which take place on the earth” as “it causes the agitations of the atmosphere, the ascension of clouds, the fall of rain and of meteors, the currents of water which channel the surface of the globe, and of which man has thus far employed but a small portion” [10]. Starting from these facts, in particular from the idea that steam was just a mean allowing the caloric to flow and that could in principle be substituted by other substances, Carnot managed to abstract the functioning of heat engines by conceiving an ideal set of operations on a system – a gas or a liquid – realizing a systematic extraction of work. He realized that this extraction requires the transfer of heat between two reservoirs, at least: a heat source (the hot reservoir) and a heat sink (the cold reservoir), as they produce the necessary fall of caloric.

¹ We are loosely referring here to the common sense distinction between what is “scientific” and what is “technological”. We are aware that a neat distinction of the two domains is far from being trivially granted.

Starting from calorimetry measurements, he computed the heat contained in coal per unit of mass, and using an estimate of the amount of motive power per unit of heat, he gave an estimate of the motive power of one kilogram of coal. He used this estimate to compare the performance of known steam engines in terms of the fraction of coal motive power they were able to exploit (the best result was 5%, but it was an overestimate due to the fact that Carnot consciously underestimated the motive power of coal). By analyzing his idealized engine when driven in a reversible manner, i.e. quasi-statically between equilibrium states, he derived the protocol – the Carnot cycle – yielding the maximum efficiency that any heat engine can achieve when operated between two given temperatures, now known as the Carnot limit. Even if Carnot never wrote the explicit mathematical expression, he also understood that no work can be extracted from a single heat reservoir (or two heat reservoirs at the same temperature) and that the efficiency of power extraction is proportional to the temperature difference between the two heat sources.

Some of the basic assumptions of Carnot’s analysis have later been proved incorrect (e.g., the theory of caloric and the equations governing the adiabatic expansion of gases) and therefore prevented him to obtain solid quantitative results. However, thermodynamics stems from his contributions in so far as he introduced a model epitomizing the conceptual and practical elements forming the basis of heat engines operation, and he also conceived the ideal quasi-static protocol, which would be impossible to realize in practice but which constitutes a powerful theoretical tool to draw general conclusions. After his premature death, Carnot’s estimates were made more precise by the experiments of James Joule (1818-1889) “on the mechanical equivalent of heat” [13]. Starting from that, a number of German scientists among those Julius von Mayer (1814-1878), Hermann von Helmholtz (1821-1894), and Rudolf Clausius (1822-1888) – who could read about Carnot’s ideas mostly through the German translation of Émile Clapeyron’s (1799-1864) works, who further developed them [12] – together with William Thomson (1824-1907, also known as Lord Kelvin) developed thermodynamics basically in the form that we know today [14–16].

A blast from the past: thermodynamic analysis of the aeolipile

To recap the main results of classical thermodynamics concerning the analysis of machines, we provide here a qualitative thermodynamic analysis of the aeolipile. While simple, its functioning recapitulates the most important concepts that we need at this point. What follows is meant to be a light and intuitive presentation aimed at furnishing some background to better discuss the next Section 1.2, leaving a more rigorous attitude to Section 1.3.

For our purposes, we can model the aeolipile presented in Section 1.1.1 as a system (the metal hollow sphere containing water) in contact with two *reservoirs* (the air and the fire pit), as sketched in Figure 2a. The reservoirs are special systems whose properties do not change in time and are fully characterized by intensive quantities whose values are externally determined. We assign an absolute (positive) temperature T_c and a pressure P_c to the surrounding air, and an absolute temperature T_h to the fire pit. Instead, the system is characterized by extensive quantities which can either be conserved or varied by the dynamics. For our model aeolipile, we consider both the volume and the amount of water as conserved quantities (i.e., we neglect thermal expansion and matter losses due to steam flows). The

quantities whose time variation we will be concerned with are the internal energy U and the entropy S , which for the moment can be considered as two functions whose expressions are system and state specific.

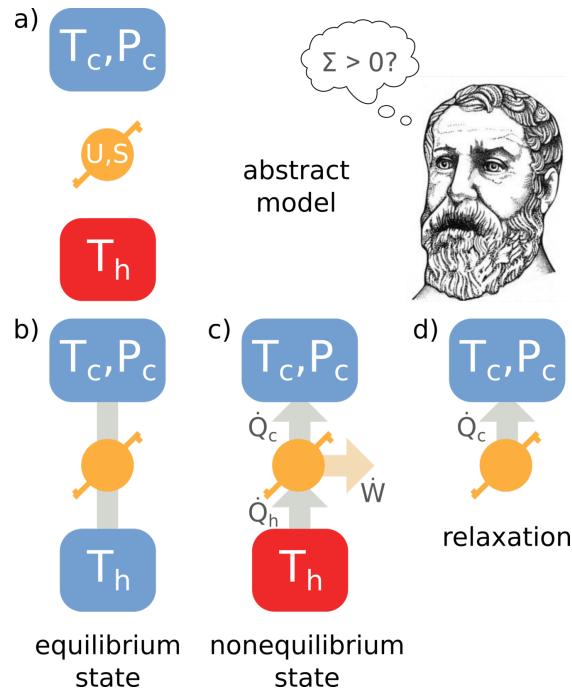


Figure 2: Thermodynamic modelling of the aeolipile. a) The aeolipile as a system in contact with two reservoirs c and h , with Hero's of Alexandria figuring out the second law. The reservoirs are characterized by temperature (T_c , T_h) and pressure (P_c). The system is characterized by internal energy (U) and entropy (S); b) Equilibrium state characterized by the absence of any net exchange between the system and the reservoirs; c) A nonequilibrium state driven by a temperature gradient ($T_h > T_c$). Heat flows from the hotter reservoir to the system (\dot{Q}_h) and from the system to the colder reservoir (\dot{Q}_c). Possibly, the system can deliver power to the environment (\dot{W}) if coupled to external forces; d) Relaxation towards the equilibrium state after the hotter reservoir is removed.

THE ZEROth LAW AND THE DEFINITION OF EQUILIBRIUM. Let's briefly consider the engine at *equilibrium*. In thermodynamics, two or more otherwise isolated systems are said to be in thermodynamic equilibrium if, upon pairwise contact, they do not exchange any quantity (e.g., heat, work, matter). This can be taken as a formulation of the so-called zeroth law of thermodynamics, and has the major implication that systems in equilibrium have identical intensive quantities. In fact, the equilibrium relation formally justifies the measurement of intensive parameters through physical instruments like thermometers. At equilibrium, we can therefore assign to our model aeolipile in Figure 2b the temperature $T = T_c$ and the pressure $P = P_c$ of the air (neither heat or matter flow between the system and the air), and deduce that there cannot be powering ongoing ($T_h = T_c$, meaning that equilibrium is possible only when the fire is off). Since there are no net exchanges between the system and the reservoirs, the values of internal energy and entropy in the equilibrium state ($U^{(eq)}$ and $S^{(eq)}$) are constant in time and can be conveniently taken as reference values. For a thermo-

dynamic system, the equilibrium state with respect to a certain reservoir is completely determined by its conserved quantities.

THE FIRST AND THE SECOND LAWS. We now go back to the general case represented in Figure 2c, where $T_h \neq T_c$ (fire on), and ask how internal energy and entropy vary in time when power in the form of heat is supplied to the system. The celebrated first and second laws of thermodynamics respectively answer both questions in the most general way. Here, the former takes the form of the following energy balance:

$$\text{First law: } d_t U = \dot{Q}_h + \dot{Q}_c + \dot{W}. \quad (1)$$

It expresses the fact that the internal energy of a system can vary in time only due to heat or work exchanges with the reservoirs. In particular, \dot{Q}_h and \dot{Q}_c denote the heat exchanged per unit of time with the hot and the cold reservoir, respectively, while \dot{W} is the mechanical power delivered (or absorbed) by the engine if, for instance, a load is attached. All the flow-like quantities denoted with dots are conventionally taken as positive when the flow goes from the reservoirs to the system (e.g., useful mechanical work can be delivered when $\dot{W} < 0$). It is worth noticing that the first law of thermodynamics is not exactly “a version of the law of conservation of energy”, as it is sometimes presented as, but rather an operational way to define and measure heat. In fact, *starting from* the assumption that energy is conserved and from the empirical observation that the energy change in a system undergoing a certain process does not always correspond to the work it exchanges, the first law prescribes that the part of the energy must have been exchanged as heat. In connection with the historical background of thermodynamics ideas traced above, the first law states the equivalence of heat and work as two interchangeable forms of energy exchange, thus condensing many results of calorimetric experiments.

Turning to the entropy balance, it takes here the following form:

$$d_t S = \frac{\dot{Q}_h}{T_h} + \frac{\dot{Q}_c}{T_c} + \dot{\Sigma}. \quad (2)$$

It expresses the fact that entropy varies in time due to heat exchanges with reservoirs plus and extra term $\dot{\Sigma}$, which is now often called entropy production rate. The second law of classical thermodynamics states that the entropy production rate can never be negative, being zero solely for those processes keeping the system in its equilibrium state:

$$\text{Second law: } \dot{\Sigma} \geq 0. \quad (3)$$

Such statement has predictive power, as it requires physical systems not to undergo transformations leading to a negative entropy production rate. However, it has no explanatory content, as it does not justify the positiveness of the entropy production rate with a mechanistic theory, nor does it prescribe how $\dot{\Sigma}$ can be directly computed.

STEADY STATES AND CARNOT EFFICIENCY Equations (1), (2), and (3) distill classical thermodynamics to its theoretical essence. They apply to virtually any physical model, allowing us to recover empirically observed bounds on systems’ behavior and to draw nontrivial connections and statements about measurable quantities. As an example, we now use them to characterize the aeolipile in its *steady state* and to show how Carnot’s limit on efficiency directly follows.

When operated between two reservoirs with a constant gradient of intensive quantities ($T_h - T_c > 0$), many systems will eventually reach a steady state, that is a nonequilibrium condition where all the state-dependent properties (e.g. internal energy, entropy) are constant in time. In the case of Hero's engine, the steady state will be characterized by stationary temperature and pressure of the steam inside the sphere, such that its outflow is also constant. Steady states are interesting for the sake of characterizing the functioning of an engine, because they can virtually be maintained for an infinite time (as long as the gradient is maintained and up to degradation of the machine) without changing the performance. By imposing the steady state condition ($d_t U = 0$ and $d_t S = 0$) in equations (1) and (2), and by using equation (3), the following inequality can be derived:

$$T_c \dot{\Sigma} = \dot{W} + \dot{Q}_h \left(1 - \frac{T_c}{T_h}\right) \geq 0, \quad (4)$$

where we also exploited the fact that we are using the absolute scale to measure temperatures (and therefore T_c is positive). Crucially, if we now evaluate the efficiency of the engine as the ratio between the (negative defined) delivered work and the heat exchanged with the hot reservoir, we find the Carnot limit

$$\frac{-\dot{W}}{\dot{Q}_h} \leq \left(1 - \frac{T_c}{T_h}\right). \quad (5)$$

It means that, when Hero's engine – or any other heat engine powered by two thermal reservoirs – is used to perform work (i.e., lifting a load) the amount of input heat that can be converted into delivered power has a universal limit that, as Carnot foresaw, depends on the temperature gradient. It should be stressed that real engines often work way below this thermodynamic limit, as the actual efficiency is governed by the extra energy dissipation occurring due to system-specific features. We also note that the resulting spinning of Hero's engine (whose work against air friction can be neglected for the purposes of this discussion) and its consequent capability to deliver some useful mechanical work thanks to the rotary motion are consequences of the asymmetric shape of the blowholes. If the latter were – say – perfectly aligned, the thermodynamic description would be the same, but extracting useful mechanical power would be unpractical. This shows that the dynamic features of systems in nonequilibrium steady states, in particular their capability to perform work as allowed by thermodynamics, strongly depends also on their specific design.

ENERGY STORAGE. It is useful to introduce one last thermodynamic concept: the energy stored in nonequilibrium states. Consider the aeolipile in its steady state, which is clearly not at equilibrium with the surrounding air ($\dot{Q}_c \neq 0$ and steam flows from the inside to the air). In particular, if $\dot{W} = 0$ (the engine is not coupled with work sources) and $T_h > T_c$ (the fire is hotter than the air), the following equation must hold

$$\dot{Q}_c \left(\frac{1}{T_h} - \frac{1}{T_c}\right) > 0, \quad (6)$$

implying $\dot{Q}_c < 0$. We can therefore conclude that the aeolipile at steady state is hotter than the surrounding air. Based on what we discussed above, if we suddenly extinguish the fire as represented in Figure 2d, the system will keep on delivering heat to the air at the expense of the internal energy

($d_t U = \dot{Q}_c$) until relaxation to equilibrium is completed. As the heat released can be in principle used to perform useful work, we say that some energy ΔU is stored in the nonequilibrium state of the system. Since we assumed that both the volume and the amount of water (the two conserved quantities uniquely determining the equilibrium state of our model aeolipile) do not significantly change during the operation, the total amount of energy stored in the system is given by the difference between the internal energy at steady state and the equilibrium one ($\Delta U = U_{\text{st.st.}} - U_{\text{eq}}$). Note that, in the absence of a detailed model allowing to compute $U_{\text{st.st.}}$ and/or U_{eq} , measuring the heat released with a calorimeter offers a practical way to determine energy differences.

1.1.5 Final considerations

After its birth, thermodynamics fulfilled its mission of guiding a rational development of heat engine technology as foreseen by Carnot. One example among many is the invention of diesel engine, which was explicitly conceived by Rudolf Diesel (1858-1913) as an attempt of creating a highly efficient engine that could work on the Carnot cycle [17]. However, it turned out that the range of applicability of thermodynamics is much larger than heat engines. With the words of Albert Einstein (1879-1955), it can be considered “the only physical theory of universal content concerning which I am convinced that, within the framework of the applicability of its basic concepts, it will never be overthrown” [18]. Moreover, after marking the abandonment of Newtonian mechanics as the sole all-embracing physical theory [6], thermodynamics also contributed to originating the debate around the problem of black body radiation [19]. Far from lacking crucial epistemological relevance, the solution of the latter given by Max Planck (1858-1947) called the beginning of quantum mechanics [20]. Also, statistical mechanics largely developed from the attempts to give a microscopic mechanistic foundation to the second law. With its most recent extensions, thermodynamics spans from quantum to planetary scales, and is routinely used to study systems way beyond engines. As we will see, beyond heat and work, thermodynamics also covers the physical effects of information.

In this thesis, thermodynamics will be applied, and some of its field-specific formulations will be slightly extended, to study chemical systems. In particular, those introduced in the next Section 1.2 are meant to be the ideal case studies for the framework developed in the following chapters. An account of the main scientific milestones which have allowed the application of thermodynamics to chemical systems will be given in Section 1.3.

REFERENCES FOR CHAPTER 1.1

- [1] HERO OF ALEXANDRIA, *The pneumatics of Hero of Alexandria*, ed. by B. WOODCROFT, Taylor, Walton and Maberly, 1851.
- [2] L. RUSSO, *The Forgotten Revolution*, Springer, 2004.
- [3] M. BOAS, “Hero’s Pneumatica: A Study of Its Transmission and Influence”, *Isis* 40.1 (1949), 38–48.
- [4] G. B. D. PORTA, *Pneumaticorum libri tres*, ed. by O. TRABUCCO, vol. 10, Edizioni Scientifiche Italiane, 2009.
- [5] S. de CAUS, *Les Raisons des Forces Mouvantes*, Jan Norton, 1615.

- [6] B. HESSEN, “The Social and Economic Roots of Newton’s Principia”, *The Social and Economic Roots of the Scientific Revolution: Texts by Boris Hessen and Henryk Grossmann*, ed. by G. FREUDENTHAL and P. McLAUGHLIN, Springer Netherlands, 2009, pp. 41–101.
- [7] J. FAREY, *A Treatise on the Steam Engine*, Longman, Rees, Orme, Brown, and Green, 1827.
- [8] J. D. BERNAL, *The Extension of Man : A History of Physics Before 1900*, Paladin, 1973.
- [9] D. CARDWELL, *From Watt to Clausius: The Rise of Thermodynamics in the Early Industrial Age*, Iowa State University Press, 1989.
- [10] S. CARNOT, *Reflections on the motive power of fire*, ed. by E. MENDOZA, Dover Publications, 1988.
- [11] J. SMEATON, “An Experimental Enquiry concerning the Natural Powers of Water and Wind to Turn Mills, and Other Machines, Depending on a Circular Motion”, *Philosophical Transactions of the Royal Society of London* 51.0 (1759), 100–174.
- [12] E. MENDOZA, “Reflections on the motive power of fire”, Dover Publications, 1988, chap. Introduction.
- [13] J. P. JOULE, *The Scientific Papers of James Prescott Joule*, Cambridge University Press, 2011.
- [14] R. CLAUSIUS, *The Mechanical Theory of Heat*, Taylor and Francis, 1867.
- [15] T. KUHN, “Critical Problems in the History of Science”, ed. by M. CLAGETT, Madison: University of Wisconsin Press, 1959, chap. Energy Conservation as an Example of Simultaneous Discovery.
- [16] Y. ELKANA, *The Discovery of the Conservation of Energy*, Harvard University Press, 1974.
- [17] R. DIESEL, *Theory and Construction of a Rational Heat Motor*, E. & F. N. Spon, 1894.
- [18] P. A. SCHILPP, ed., *Albert Einstein: Philosopher-Scientist*, MIF Books, New York, 1970.
- [19] G. CICCOTTI and E. DONINI, “L’Ape e l’Architetto”, Feltrinelli, 1976, chap. Sviluppo e crisi del meccanicismo: da Boltzmann a Planck.
- [20] M. CINI, *Un paradiso perduto*, Feltrinelli, 1994.

1.2 CHEMICAL ENGINES

We're divided into two main branches, those who rig and those who dismantle or break down, and both kinds are like blind people with sensitive fingers. I say blind because, actually, the things we handle are too small to be seen, even with the most powerful microscopes: so we've invented various intelligent gadgets to recognize them without seeing them.

— Primo Levi, *The Wrench*
(Abacus, 2013. Original version in Italian: 1978)

One of the central technological challenges of the last sixty years has been how to move and control objects of very tiny dimensions, in particular at the nanoscale where molecules reside. More broadly, the enterprise of building systems able to process and use energy to perform tasks in the realm of chemistry was embraced by scientists. We will generally refer to such kind of systems as *chemical engines*, a terminology whose usage in this sense is currently gaining momentum in the literature [1].

As it was very soon realized [2], a major complication is that such kind of molecular devices cannot be conceived as merely miniaturized versions of their macroscopic counterparts, as physics looks quite different at the nanoscale. In particular, Brownian motion and electromagnetic interactions take over inertia and gravity in determining the motion of objects. As a result, the working principles of chemical engines are fundamentally different from those at the macroscopic scale, and controlling their motion precisely is impracticable. This has become even clearer after a series of fundamental experimental evidences that revealed how various kind of chemical engines are ubiquitous in living cells and intrinsically stochastic. Biological chemical engines include motor proteins like myosins [3], kinesins [4] and dyneins [5] for muscles contraction and intracellular linear transportation [6–8], ATP synthase [9, 10] and ribosomes [11] for synthesis of ATP and proteins, and vacuolar ATPases [12] and other ion pumps [13, 14] for generating concentration and electrochemical gradients across membranes. One common trait underlying many biological chemical engines is their ability to catalyze the conversion of energy-rich chemical species (e.g., purine nucleoside triphosphates like ATP and GTP) to low-energy products. Such a fuel-to-waste conversion is what powers their functioning and makes possible to bias their stochastic dynamics towards on-average useful behavior. Other fueling mechanisms are also possible like concentration gradients in space as in the aforementioned ATP-synthase [15], light as in the photosynthetic machinery [16, 17] and electrical potentials as in bio-electrochemical systems performing microbial electrosynthesis [18, 19]. Additionally, biological chemical engines utilize architectures that restrict their degrees of freedom and/or the possible interactions they can experience (e.g., tubulin tracks guiding linear motion and membranes where ion pumps are embedded).

The variety and elaborateness of systems found in biology have inspired synthetic chemists to develop artificial chemical engines reproducing the basic features of biological ones. The field of chemistry that probably contributed most to this attempt is supramolecular chemistry, namely “the chemistry beyond the molecule, bearing on organized entities of higher complexity that result from the association of two or more chemical species held together by intermolecular forces”, as it was defined by Jean-Marie Lehn [20, 21]. Supramolecular chemists developed effective bottom-up ap-

proaches [22] to build multi-component molecular architectures with characteristic structures and shapes (e.g. molecular knots [23]) exhibiting peculiar properties and performing specific and targeted functions resembling those found in nature. In the years, it has become more and more evident that such an approach can make invaluable contributions to a better understanding of molecular-level aspects of the extremely complicated chemistry that is responsible for some of the aforementioned cellular processes. Indeed, while probing the same physical principles, supramolecular systems are much simpler than biological ones in terms of the chemical complexity of the molecules involved, the sophistication of the tasks they perform (often proofs-of-concept for targeted behaviors) and in the number of species taking part to the mechanisms. Moreover, the majority of the experiments are carried out in solution, thus avoiding cumbersome *in vivo* measurements or sample preparation. Therefore, supramolecular systems can be experimentally characterized in great detail, thus offering a unique test-bed for any theoretical framework aimed at disclosing the working principles of complex nonequilibrium chemical systems.

In the following, a non-exhaustive selection of the most important examples of synthetic chemical engines is presented, which spans over various tasks that have been experimentally demonstrated. The focus is on those systems whose ideal working regime is a nonequilibrium steady state sustained by a continuous power input.

1.2.1 Molecular aeolipiles

We first discuss a class of synthetic chemical engines of great relevance as – just like Hero’s engine – they proved the concept that it is possible to obtain an on-average directed rotary motion at the molecular level. This provided that power is constantly supplied and that a certain degree of *kinetic asymmetry* is introduced in the system’s design. This kind of behavior in molecular systems, often denoted as directed *mechanical* motion to stress that their movement happens *in space*, has been achieved with various different strategies, which gave birth to the field of synthetic molecular motors.

Overcrowded alkene-based light-driven molecular motors

These kind of systems are considered the first working examples of autonomous molecular motors artificially synthesized, as explicitly recognized by the Nobel committee by awarding Bernard Feringa one third of the 2016 Noble Prize in chemistry for getting “a molecular rotor blade to spin continually in the same direction” in 1999 (the prize was shared with Jean-Pierre Sauvage and Fraser Stoddart). Various prototypes sharing the same main working principles have been realized in the last two decades [24–28], and a representative example is reported in Figure 3 [29].

Overcrowded alkene-based light-driven molecular motors use ultraviolet and/or visible light as energy source, which induces photoisomerisation of an alkene moiety linking two molecular blades (two substituted helicenes in the original experiment [24]). The motors also undergo thermal helix inversion between the axial and the equatorial isomers of each photoisomer. As a result, these are four-states motors where a cycle through the states corresponds to rotation in space of one part of the motors with respect to the other along the alkene moiety. In this kind of light-driven systems, asymmetry in the cycling direction can be introduced in various ways, as it will be

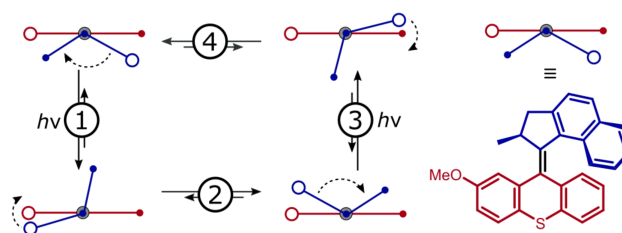


Figure 3: Exemplary overcrowded alkene-based light-driven molecular motor. Reaction scheme and molecular structure of a second-generation light-driven rotary motor [29]. Dashed arrows show the directed movement of the blue part of the motor with respect to the red one for the anticlockwise cycle.

systematically analyzed in Chapter 3. Usually, kinetic selection in the photoisomerisation steps is introduced such that photochemical reactions happens preferentially in one direction and/or thermal relaxations are tuned such that they are both energetically downhill in one cycling direction [28, 30]. Under appropriate conditions, this kinds of systems are capable of MHz rotation [25].

Imine-based light-driven molecular motors

In 2006, Jean Marie Lehn proposed to consider compounds containing imines or their derivatives a minimalist class of unidirectional molecular rotors itself [31]. As it will be detailed in Chapter 3, photoisomerization reactions happening along a certain reaction coordinate and through electronically excited states of the molecule may coexist with a ground-state thermal pathway connecting the same two isomers, possibly along a different reaction coordinate. In imines and related compounds, it may happen that the photochemical pathway draws an out-of-plane rotation around the carbon-nitrogen double bond, while the thermal pathway carries out an in-plane nitrogen inversion. Therefore, when the two isomerization processes occur in sequence, a full rotation is accomplished. This behaviour has been experimentally probed in various systems [32, 33], and a model accounting for its thermodynamic aspects will be given in Chapter 3.

A catenane-based chemically-driven molecular motor

This system was experimentally realized by the Leigh group in 2016 [34], and it is the first and currently only² autonomous chemically-fuelled synthetic molecular motor. As shown in Figure 4, it comprises a circular track with two binding sites of equal stability for a macrocycle. The latter can shuttle between the two stations, but this mechanical movement can be prevented by two fluorenylmethoxycarbonyl (Fmoc) barriers, which are continuously added and removed via the consumption of a chemical fuel F (Fmoc-Cl). Directionality emerges under continuous fuel consumption because the barrier-forming reactions are slower when the macrocycle is proximal to the barrier-formation site, most likely due to steric crowding effects. This effect is indicated as kinetic gating and is the molecular origin of kinetic asymmetry for this system. Instead, the rate constants of all the other reactions are unaffected by the location of the macrocycle. Thus, on average, shuttling in the clockwise direction by overcoming the proximal barrier-forming site

² The Leigh group has successfully synthesized a 4-state autonomous chemically-driven single bond rotary molecular motor too, but results are still unpublished (private communication).

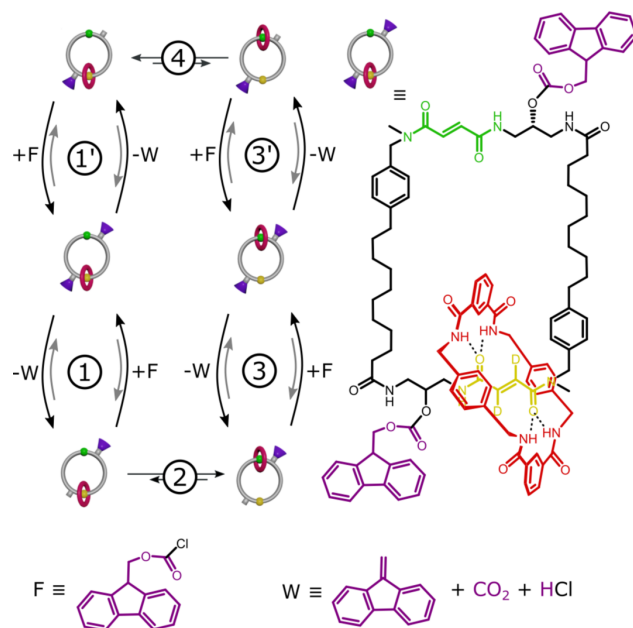


Figure 4: A catenane-based chemically-driven molecular motor. Reaction scheme and molecular structure of the rotary motor [34].

is more likely than shuttling in the counterclockwise direction by overcoming the distal barrier-forming site. To perform a directional cycle, a motor needs to invert which barrier is in place (vertical reactions in Figure 4). Under the employed experimental conditions, this process occurs via a state having two barriers simultaneously present, affording the pictured six-state scheme. Due to its uniqueness and minimalist design, this motor will be extensively analyzed in Chapter 4. It has also inspired an innovative computational approach to supramolecular chemistry based on grand canonical monte carlo simulations [35].

Nonautonomous molecular motors

Even if they will not be explicitly analyzed as case studies in this thesis, we mention here that directed rotary motion has been obtained in nonautonomous systems too. Their functioning is based on a periodic sequence of external stimuli of different nature (chemical, photochemical, electrochemical) which need to be repeated for each cycle. Examples include catenane-based systems reported by the Leigh group [36–38], the molecular motors based on palladium redox cycle [39] and lactone formation [40] reported by the Feringa group, and the lactone-based biaryl rotary motor reported in 2005 by the Branchaud group [41].

1.2.2 Molecular pumps

As will become clearer in the next Section 1.3, the fact that they display directed motion is enough to conclude that all the above systems work in an out-of-equilibrium regime, eventually reaching nonequilibrium steady states. As a consequence, they store some energy which, just like we saw for Hero's engine, is released and can be in principle used when the power input is stopped. However, this energy storage capability is more of a side-

effect of the main cause those systems serve, that is demonstrating directed mechanical motion in synthetic molecular systems.

Experimental designs explicitly aimed at inducing molecular mechanical motion such to bias the equilibrium distribution as long as input power is supplied are often denoted “molecular pumps”. In these systems, input power is used to reach and sustain a nonequilibrium steady state with respect to a mechanical degree of freedom, i.e., the out-of-equilibrium feature – and therefore the energy storage capability – is concerned with the position in space of an object, analogous to macroscopic pumps lifting water buckets. Before discussing some examples drawn from supramolecular chemistry, we stress the difference between molecular pumps – where the out-of-equilibrium steady state can be in principle maintained as long as the power input is supplied – and kinetic resolution experiments based on the Curtin-Hammett principle, which often involve mechanical motion but where the out-of-equilibrium regime is just transient – i.e., systems are such that they eventually relax to equilibrium in the long run.

A molecular information ratchet

A prominent example of an autonomous synthetic molecular pump is a light-driven rotaxane-based system reported in 2007 by the Leigh group [42] and represented in Figure 5. This system consists of a macrocycle mechan-

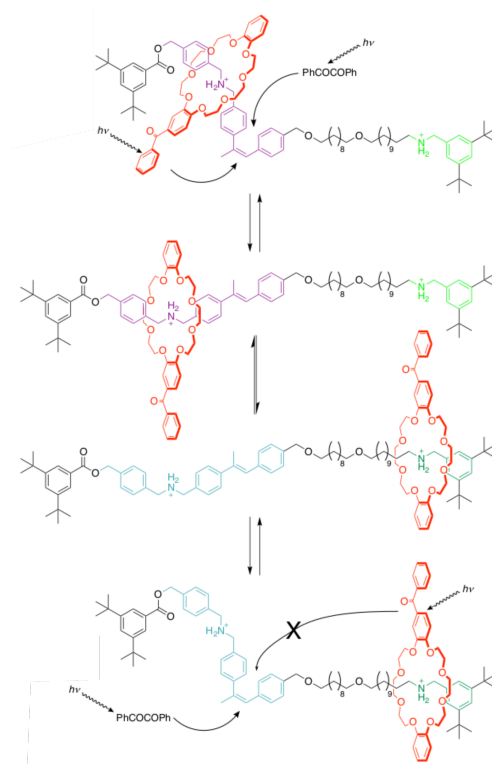


Figure 5: A rotaxane-based light-driven molecular pump. Reaction scheme and molecular structure of the molecular pump in Ref. [42]. For clarity, the energy transfers from the photosensitizer PhCOCOPh are only shown for the Z-isomers of the rotaxane. The same processes occur for the equivalent E-isomers.

ically interlocked on a linear track that has a stilbene unit displaying E/Z photoisomerization and two binding sites for the macrocycle: a dibenzyl am-

monium (DBA) and a monobenzyl ammonium (MBA). When the stilbene is in the E-configuration, the macrocycle can shuttle between the two stations, but when the stilbene is in the Z-configuration, the macrocycle shuttling is blocked. Under constant irradiation at 350 nm and in the presence of a light sensitizer (PhCOCOPh), the distribution of the macrocycle between the two stations is driven out of equilibrium. In particular, the distribution of the macrocycle between two binding sites shifts from DBA:MBA = 65:35 (equilibrium) to DBA:MBA = 45:55 (nonequilibrium steady state), showing that the macrocycle gets *pumped* from the DBA to the MBA station when light energy is supplied. Kinetic asymmetry in this system is introduced thanks to what is usually called an *information ratchet mechanism*, i.e., a kinetic bias in the photoisomerization rate of the stilbene unit which depends on the position of the macrocycle. Indeed, thanks to a photoinduced energy transfer from the macrocycle, the photoisomerization rate is increased when the latter resides on the DBA site, thus ratcheting the position of the macrocycle towards the MBA station (see Figure 5). As will be detailed in Chapter 4, the mechanism just described involves information transfer between the degrees of freedom of the system (macrocycle position and track's state), and implements a Maxwell's demon-like mechanism.

Light-powered bimolecular pumps

In the previous example, the pumping is intra-molecular, as it happens along an internal degree of freedom of a single (supra-) molecule. A step beyond in the field has been the autonomous inter-molecular pumping, where a molecule is pumped through another one. A first instance of it has been developed in 2015 by the Credi group, with a second-generation prototype reported in 2021 [43–45]. As can be seen from Figure 6, the system is directly inspired to the one previously described but, instead of being interlocked, here a macrocycle can thread both sides of an asymmetric axle. In fact,

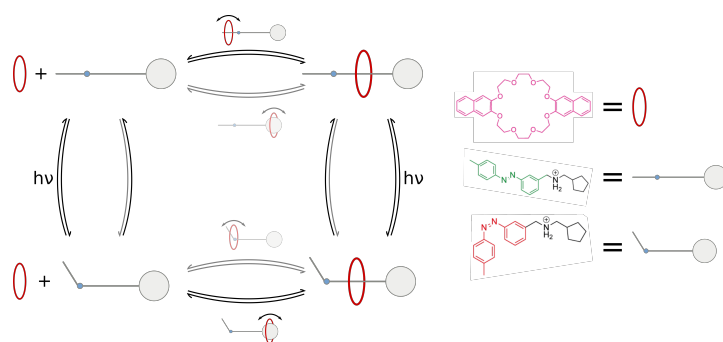


Figure 6: A light-powered bimolecular pump. Reaction scheme and molecular structure of the first generation supramolecular light-driven pump [43]. Transparency indicates kinetically unfavorable processes.

rather than the energy storage which nevertheless happens, the focus of the original experiment was more on the linear autonomous unidirectional motion between the two molecular components. The latter nonequilibrium behavior happens at steady state thanks to an information ratchet mechanism similar to the one described in the previous example: indeed, the macrocycle enhances the photoisomerization of the axle when threaded in. In addition, threaded species are energetically favored over not-threaded ones, but the macrocycle gets slightly destabilized when the axle switches from E to Z conformation. This kind of effect (which does not depend on

the macrocycle's position) is often recognized as an *energy ratchet mechanism* in the literature of light-driven molecular motors, and together with the information ratchet mechanism provides the kinetic asymmetry allowing for the motor's functioning. As a consequence, threading of the macrocycles into the axes happens preferentially in the E-conformation (and through the azobenzene side due to steric hindrance), while de-threading happens preferentially in the Z-conformation (and through the cyclopentyl side due to steric hindrance), thus favouring net directed motion of the macrocycles with respect to the axes when the system is brought out of equilibrium by light. This class of systems, in particular the second-generation prototype, will be extensively analyzed in Chapter 3 based on original experimental data allowing to quantitatively determine the energy storage capability of the pump in various regimes. Moreover, data from the first prototype will be used in Chapter 4 to qualitatively and quantitatively disclose the relationship between kinetic concepts like information and energy ratcheting and thermodynamic quantities.

A chemically-powered bimolecular pump

In the previous example, a single axle can host at most one single macrocycle, and its ability to store it is unavoidably decreased by the fact that the latter can escape from both sides (a feature necessary to obtain linear unidirectional motion though). The challenge to autonomously pump more than one macrocycle on a single axle and to store them in a nonequilibrium steady state as long as power is supplied to the system has been recently accomplished by the Leigh group [46]. Moreover, as shown in Figure 7, their system is chemically-fueled, thus adding up to the few examples of not light-driven autonomous artificial chemical engines. The working principle

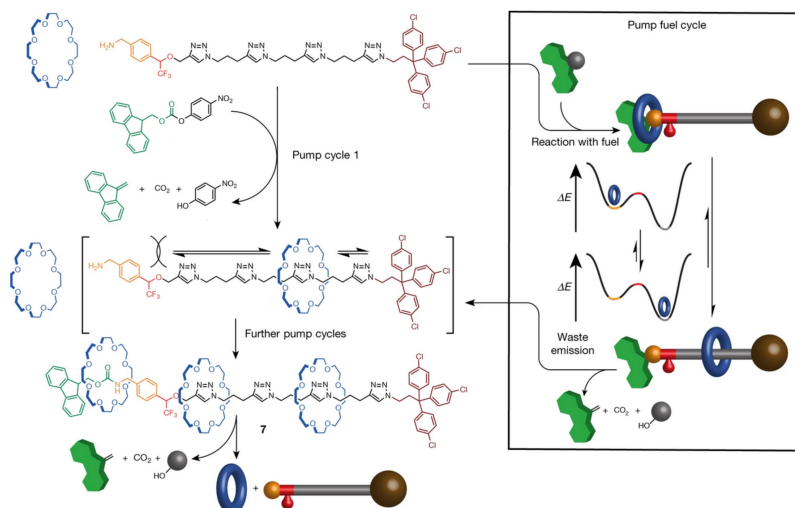


Figure 7: A chemically-powered bimolecular pump. Reaction scheme and molecular structure (adapted from Ref. [46]).

of this system is somehow similar to that of the catenane-based molecular motor described in Section 1.2.1, but here the macrocycle has the double function to promote barrier attachment behind it upon threading and to suppress subsequent barrier removal until migration to a catchment region happens. Out of equilibrium steady states with up to 3 macrocycles pumped on a single axle from a bulk solution as long as power input is provided have

been experimentally characterized. It is worth noticing that the possibility of having more than one macrocycle threaded on an axle (which can itself have or not a barrier in place) significantly increases the number of species in the system with respect to all the previous examples. As one can expect, while experiments progress, the models describing them get more and more complex very quickly. This anticipates the natural rising of a branch of supramolecular chemistry known as systems chemistry, which will be briefly introduced below (see Section 1.2.4).

An electrically-powered bimolecular pump

As it emerges from the examples discussed so far, autonomous electrically-driven molecular machines are significantly underdeveloped in comparison with light-driven and chemically-driven ones. One reason is that, while photochemical and chemical power inputs can be realized in homogeneous solutions, generating an electrical potential gradient requires interfacing the system with two electrodes operating at different potentials. As shown by Huskens and coworkers, the simultaneous but spatially distinct oxidation and reduction of redox-active species can be realized with a bipotentiostat enabling the independent control of two electrodes, thus offering a general strategy to operate electrochemically-driven systems [47]. Recently, a similar setup was successfully employed to demonstrate the autonomous operation of the redox-driven system depicted in Figure 8 [48]. The scheme describ-

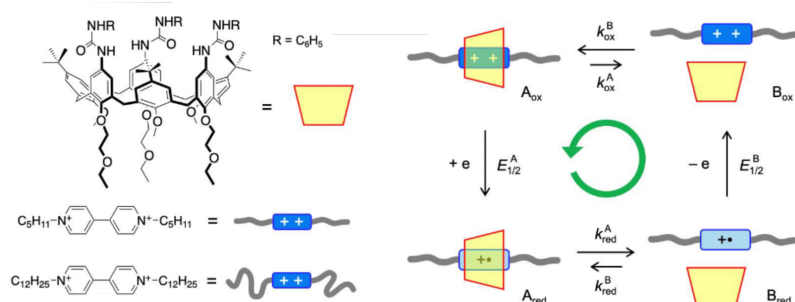


Figure 8: An electrically-powered bimolecular pump. Reaction scheme and molecular structure (adapted from Ref. [48]).

ing its functioning resembles that of the light-driven bimolecular pumps described above, but here oxidation and reduction steps coupling the systems to the electrodes happen preferentially in separate regions of space. Such a design can properly be considered an electrically driven molecular pump according to the definition we gave above, as threading and dethreading motions of the species take place autonomously in solution, reaching a nonequilibrium steady state where the concentration of the threaded reduced species A_{red} is promoted.

Nonautonomus molecular pumps

Similar to the case of directed rotary motion, molecular pumping has been achieved in nonautonomous systems too [49]. Among the most notorious examples, we mention the series of rotaxane-based molecular pumps developed by the Stoddart group starting from 2015³ [50–52] and the one

³ The first prototype of this series is often recognized as “the first artificial molecular pump” [49] according to a stricter – and more common – definition than the one used here. A broader

reported by the Leigh group in 2017 [53]. Both the approaches were recently applied to demonstrate mechanisorption, i.e. the nonequilibrium pumping to form mechanical bonds between an adsorbent and an adsorbate [54, 55].

1.2.3 Rotary and linear macroscopic work done by chemical engines

Since they were starting to be conceived, one of the great expectations from chemical engines has always been that they could one day be employed to perform useful macroscopic work. Even if far from technological applications, progress in this direction is significant. Early experimental demonstrations of the ability of light-powered molecular systems to move macroscopic objects comprise millimetre-scale directional transport of a liquid on a surface [56] and the rotation of objects exceeding the size of a motor molecule by a factor of 10^4 [57]. In 2015, the Giuseppone group reported on the macroscopic contraction of a gel induced by overcrowded alkene-based light-driven molecular motors, thus demonstrating the conversion of light energy into volume work mediated by molecular systems [58]. Two years later, molecular modulators were added to the material's design to allow for repetitive contraction and expansion respectively controlled by two different irradiation wavelengths, thus mimicking the mechanism with which cytoskeleton and titin proteins control sarcomere units in animal muscles [59]. Very recently, Feringa's molecular motors were also used to drive a chemical process out of equilibrium, thus realizing transduction from light to chemical energy through continuous mechanical motion [60].

Experiments probing the possibility to exert macroscopic mechanical work by using non light-driven molecular system have been done in the realm of nonautonomous systems [61].

1.2.4 Systems Chemistry

Subcellular and cellular processes – and life by extension – are based on the chemical conversion of molecules and the storage and transport of energy. Not coincidentally, cells are often referred as chemical factories, and biotechnology scaled up this feature towards the use of various kind of cells in industrial application, both as synthesizers and scavengers. At the core of a cell's functioning lies metabolism: a finely-regulated system of coupled biochemical reactions burning sugars and enabling transformations of the environment that would hardly occur spontaneously. Metabolism can be itself thought of as a chemical engine, or as composed of various chemical engines using energy to perform different tasks (digestion, cellular respiration, energy storage, synthesis of metabolites, regulation, etc.). Contrary to the examples discussed above, which are based on few species and reactions, chemical engines in metabolism are more complex, as they involve large networks of species and reactions, possibly interdependent and compartmentalized. The latter are the study objects of systems chemistry, a research area rooted in biochemistry and supramolecular chemistry which expanded their scope towards the understanding and the design of complex chemical reaction networks displaying life-like behaviors, i.e., chemical engines [62–65].

Exciting progress has been made at the intersection of systems chemistry and prebiotic chemistry by the reproduction of some of the key metabolic

definition was chosen in this thesis to stress the energetic aspects involved in pumping and to better fit the narrative.

reactions in the absence of enzymes by exploiting metal ions [66, 67]. These have to be considered groundbreaking experiments of our times, as they provide strong evidences for studying how life originated alongside with demonstrating the possibility to artificially reproduce complex energy transduction processes such as the Krebs cycle. Beyond literally trying to reproduce life, research in systems chemistry is expanding the frontiers of nanoscience, especially in the domain of soft material with life-like properties such as self-assembly [68], periodic oscillation [69] and adaptation to the environments both at and out-of-equilibrium [70, 71].

Driven self-assembly

One example of life-like behavior that has been challenging synthetic chemists is the ability of living systems to control in time and space the assembly – very often referred as self-assembly to stress that there's no external manipulation at the single molecule level – and disassembly of supramolecular polymers. Prominent examples of this kind of behavior are the formation of microtubules out of tubulin dimers fueled by guanosine 5'-triphosphate (GTP) [72, 73] and the ATP-driven self-assembly of actin filaments [74]. In particular, ATP is constantly refreshed in the cytosol by mitochondria and waste products are removed, thus sustaining steady states with peculiar properties.

Artificial supramolecular polymers have been developed over the years, but most of the examples realized so far switches between equilibrium states in response to external stimuli or reach transient nonequilibrium states eventually dying off [75–79]. In 2017, the Hermans group reported on the experimental setup shown in Figure 9 enabling to power chemical systems by a continuous influx of chemical fuel and outflux of waste products [70]. Thanks to it, they could keep a supramolecular polymer in various nonequi-

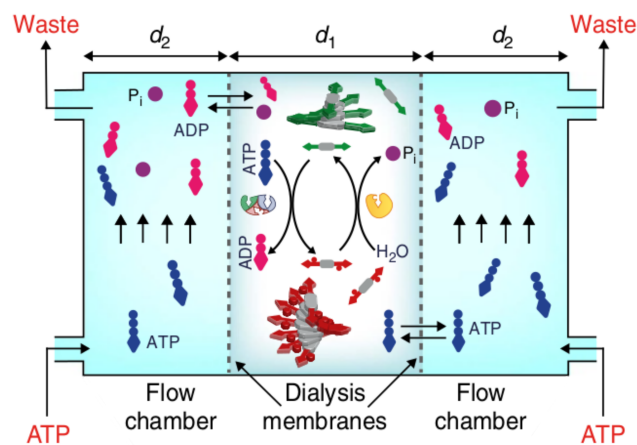


Figure 9: Experimental implementation of a chemostating mechanism. Experimental setup enabling to externally control the concentration of ATP, ADP, and phosphate ions P_i in the system (adapted from Ref. [70]).

librium steady states where ATP is used to add a charged phosphate group to the monomers, resulting in polymer growth and switching of supramolecular chirality. The concentration distribution of the species in the system is modified depending on the influx of the fuel ATP, and relaxes to equilibrium when the power input stops.

Figure 10 shows a minimalist model describing the fuel-driven self-assembly of monomers (single spheres) into dimers (pairs of spheres). In particular,

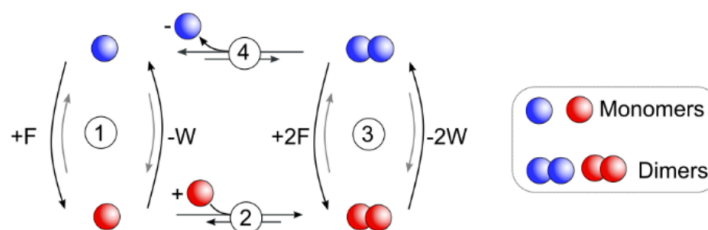


Figure 10: A model of fuel-driven self-assembly. Reaction scheme for the chemically driven self-assembly of dimers [80]. Shorter arrows denote unfavorable reactions.

the scheme comprises a low-energy monomer (blue sphere) that can self-assemble to give a high-energy dimer (blue pair of spheres) through reaction 4 in Figure 10. Both the monomer and the dimer also exist in their bound forms (red spheres), formed and unformed by two couples of reactions involving a chemical fuel F and a waste product W (vertical reactions 1 and 3 in Figure 10). Bound species take part in a self-assembling step too (reaction 2 in Figure 10), possibly favoring the dimeric state. By properly fixing the concentrations of F and W , a nonequilibrium stationary state rich in the blue dimer (target species) can be achieved. We notice that this process can be described both as a nonequilibrium synthesis – where, like in metabolism, energy is used to promote the formation of high energy species – and as a more general form of energy storage with respect to the molecular pumps mentioned above (see Section 1.2.2). Indeed, energy storage by chemical engines is not necessarily coupled to a mechanical motion, but can be more broadly understood as a shift of the equilibrium concentration distribution of the species in a system.

While simple, the model in Figure 10 epitomizes the conceptual elements of driven self-assembly. For this reason, it has been frequently employed to discuss the basic principles underlining this phenomenon [68, 80, 81]. This thesis is no exception, and the model in Figure 10 will be extensively used in Chapter 2 to discuss possible measures of thermodynamic efficiency in chemical engines performing energy storage and synthesis.

1.2.5 Final considerations

We are now in the position to clarify in which sense the current situation of artificial chemical engines shares some similarities with the one of steam engines at the beginning of the 19th century described in Section 1.1.4. Of course, it would be untenable to claim that all the advancements mentioned above are more technological than scientific. In fact, the very same objects and concepts which chemical engines are based on (molecules and reactions), together with the “various intelligent gadgets to recognize them without seeing them”, are undoubtedly products of science as it developed in the past century. Moreover, the economical impetus to develop chemical engines is almost negligible if compared to the industrial driving force which pushed steam engines’ development. Nevertheless, it remains widely acknowledged that up to now attempts to design chemical engines have been led by chemical intuition and chance, with little opportunity to reliably judge the effectiveness of a machine’s design or performance unless and until it has been realized experimentally. In particular, a quantitative

thermodynamic understanding of chemical engines is absent in the experimental literature.

One of the strongest beliefs from which this thesis started is that, at odds with the situation of steam engines in early 19th century, the thermodynamic concepts needed to develop a quantitative understanding of the processes that drive synthetic chemical engines have already been almost entirely developed by physicists. However, the strong compartmentalization of fields and languages which is typical of current scientific research has led to a situation in which chemists are generally not conversant with contemporary developments in theoretical physics and, vice versa, physicists are not trained to fully grasp experimental works in the realm of chemistry. Exactly one year before the beginning of the research work presented in this thesis, the editorial board of *Nature Nanotechnology* advocated for a “new generation of physical chemists” able “to connect supramolecular chemistry to the out-of-equilibrium thermodynamic concepts being developed by theoretical physicists” [82]. It is hard to find better words to describe the goal of this thesis.

REFERENCES FOR CHAPTER 1.2

- [1] S. AMANO et al., “Chemical engines: driving systems away from equilibrium through catalyst reaction cycles”, *Nat. Nanotechnol.* **16.10** (2021), 1057–1067.
- [2] R. P. FEYNMAN, “There’s Plenty of Room at the Bottom: An Invitation to Enter a New Field of Physics”, *Caltech Engineering and Science* **23.5** (1960), 22–36.
- [3] J. ROBERT-PAGANIN et al., “Force Generation by Myosin Motors: A Structural Perspective”, *Chem. Rev.* **120.1** (2020), 5–35.
- [4] S. M. BLOCK, “Kinesin Motor Mechanics: Binding, Stepping, Tracking, Gating, and Limping”, *Biophys. J.* **92.** (2007), 2986–2995.
- [5] G. BHABHA et al., “How Dynein Moves Along Microtubules”, *Trends Biochem. Sci.* **41.** (2015), 94–105.
- [6] R. D. VALE and R. A. MILLIGAN, “The Way Things Move: Looking Under the Hood of Molecular Motor Proteins”, *Science* **288.5463** (2000), 88–95.
- [7] M. SCHLIWA and G. WOEHLEKE, “Molecular motors”, *Nature* **422.6933** (2003), 759–765.
- [8] L. A. AMOS, “Molecular motors: not quite like clockwork”, *Cell. Mol. Life Sci.* **65.4** (2008), 509.
- [9] P. D. BOYER, “The ATP Synthase —A Splendid Molecular Machine”, *Annu. Rev. Biochem.* **66.1** (1997), 717–749.
- [10] C. von BALLMOOS, A. WIEDENMANN and P. DIMROTH, “Essentials for ATP Synthesis by F₁F_o ATP Synthases”, *Annu. Rev. Biochem.* **78.1** (2009), 649–672.
- [11] A. YONATH, “Hibernating Bears, Antibiotics, and the Evolving Ribosome (Nobel Lecture)”, *Angew. Chem. Int. Ed.* **49.26** (2010), 4340–4354.
- [12] M. FORGAC, “Vacuolar ATPases: rotary proton pumps in physiology and pathophysiology”, *Nat. Rev. Mol. Cell Biol.* **8.11** (2007), 917–929.

- [13] P. L. JORGENSEN, K. O. HÅKANSSON and S. J. D. KARLISH, "Structure and Mechanism of Na,K-ATPase: Functional Sites and Their Interactions", *Annu. Rev. Physiol.* **65.1** (2003), 817–849.
- [14] E. GOUAUX and R. MACKINNON, "Principles of Selective Ion Transport in Channels and Pumps", *Science* **310**,5753 (2005), 1461–1465.
- [15] N. SOGA et al., "Perfect chemomechanical coupling of FoF₁-ATP synthase", *Proc. Natl. Acad. Sci. U.S.A.* **114.19** (2017), 4960–4965.
- [16] J. BARBER, "Photosynthetic energy conversion: natural and artificial", *Chem. Soc. Rev.* **38**. (2009), 185–196.
- [17] R. E. BLANKENSHIP, *Molecular Mechanisms of Photosynthesis*, Wiley, 2014.
- [18] K. RABAËY et al., "Microbial ecology meets electrochemistry: electricity-driven and driving communities", *The ISME Journal* **1.1** (2007), 9–18.
- [19] K. RABAËY and R. A. ROZENDAL, "Microbial electrosynthesis — revisiting the electrical route for microbial production", *Nat. Rev. Microbiol.* **8.10** (2010), 706–716.
- [20] J.-M. LEHN, *Supramolecular Chemistry: Concepts and Perspective*, VCH, Weinheim, 1995.
- [21] V. BALZANI, A. CREDI and M. VENTURI, *Molecular Devices and Machines*, Wiley-VCH, 2008.
- [22] V. BALZANI, A. CREDI and M. VENTURI, "The Bottom-Up Approach to Molecular-Level Devices and Machines", *Chem. Eur. J.* **8.24** (2002), 5524–5532.
- [23] S. D. P. FIELDEN, D. A. LEIGH and S. L. WOLTERING, "Molecular Knots", *Angew. Chem. Int. Ed.* **56.37** (2017), 11166–11194.
- [24] N. KOUMURA et al., "Light-driven monodirectional molecular rotor", *Nature* **401.6749** (1999), 152–155.
- [25] M. KLOK et al., "MHz Unidirectional Rotation of Molecular Rotary Motors", *J. Am. Chem. Soc.* **130.32** (2008), 10484–10485.
- [26] N. RUANGSUPAPICHAT et al., "Reversing the direction in a light-driven rotary molecular motor", *Nat. Chem.* **3.1** (2011), 53–60.
- [27] G. B. BOURSALIAN et al., "All-Photochemical Rotation of Molecular Motors with a Phosphorus Stereoelement", *J. Am. Chem. Soc.* **142.39** (2020), 16868–16876.
- [28] D. R. S. POOLER et al., "Designing light-driven rotary molecular motors", *Chem. Sci.* **12**. (2021), 14964–14986.
- [29] E. M. GEERTSEMA et al., "Optimizing rotary processes in synthetic molecular motors", *Proc. Natl. Acad. Sci. U.S.A.* **106.40** (2009), 16919–16924.
- [30] R. D. ASTUMIAN, "Optical vs. chemical driving for molecular machines", *Faraday Discuss.* **195**. (2016), 583–597.
- [31] J.-M. LEHN, "Conjecture: Imines as Unidirectional Photodriven Molecular Motors—Motional and Constitutional Dynamic Devices", *Chem. Eur. J.* **12.23** (2006), 5910–5915.
- [32] L. GREB and J.-M. LEHN, "Light-Driven Molecular Motors: Imines as Four-Step or Two-Step Unidirectional Rotors", *J. Am. Chem. Soc.* **136.38** (2014), 13114–13117.

- [33] L. GREB, A. EICHHÖFER and J.-M. LEHN, "Synthetic Molecular Motors: Thermal N Inversion and Directional Photoinduced C=N Bond Rotation of Camphorquinone Imines", *Angew. Chem. Int. Ed.* **54.48** (2015), 14345–14348.
- [34] M. R. WILSON et al., "An autonomous chemically fuelled small-molecule motor", *Nature* **534**.7606 (2016), 235–240.
- [35] A. ALBAUGH and T. R. GINGRICH, "Simulating a chemically fueled molecular motor with nonequilibrium molecular dynamics", *Nat. Commun.* **13**. (2022), 2204.
- [36] D. A. LEIGH et al., "Unidirectional rotation in a mechanically interlocked molecular rotor", *Nature* **424**.6945 (2003), 174–179.
- [37] J. V. HERNÁNDEZ, E. R. KAY and D. A. LEIGH, "A Reversible Synthetic Rotary Molecular Motor", *Science* **306**.5701 (2004), 1532–1537.
- [38] S. ERBAS-CAKMAK et al., "Rotary and linear molecular motors driven by pulses of a chemical fuel", *Science* **358**.6361 (2017), 340–343.
- [39] B. S. L. COLLINS et al., "A chemically powered unidirectional rotary molecular motor based on a palladium redox cycle", *Nat. Chem.* **8.9** (2016), 860–866.
- [40] Y. ZHANG et al., "A Chemically Driven Rotary Molecular Motor Based on Reversible Lactone Formation with Perfect Unidirectionality", *Chem* **6.9** (2020), 2420–2429.
- [41] Y. LIN, B. J. DAHL and B. P. BRANCHAUD, "Net directed 180° aryl–aryl bond rotation in a prototypical achiral biaryl lactone synthetic molecular motor", *Tetrahedron Lett.* **46.48** (2005), 8359–8362.
- [42] V. SERRELI et al., "A molecular information ratchet", *Nature* **445**.7127 (2007), 523–527.
- [43] G. RAGAZZON et al., "Light-powered autonomous and directional molecular motion of a dissipative self-assembling system", *Nat. Nanotechnol.* **10.1** (2015), 70–75.
- [44] M. CANTON et al., "Second-Generation Light-Fueled Supramolecular Pump", *J. Am. Chem. Soc.* **143.29** (2021), 10890–10894.
- [45] S. CORRA et al., "Artificial Supramolecular Pumps Powered by Light", *Chem. Eur. J.* **27.43** (2021), 11076–11083.
- [46] S. AMANO, S. D. P. FIELDEN and D. A. LEIGH, "A catalysis-driven artificial molecular pump", *Nature* **594**.7864 (2021), 529–534.
- [47] S. O. KRABENBORG, J. VEERBEEK and J. HUSKENS, "Spatially Controlled Out-of-Equilibrium Host–Guest System under Electrochemical Control", *Chem. Eur. J.* **21.27** (2015), 9638–9644.
- [48] G. RAGAZZON et al., "Autonomous use of electrical energy by an artificial molecular machine", *ChemRxiv* **2021-16399**. (2021).
- [49] Y. QIU et al., "Pumps through the Ages", *Chem* **6.8** (2020), 1952–1977.
- [50] C. CHENG et al., "An artificial molecular pump", *Nat. Nanotechnol.* **10.6** (2015), 547–553.
- [51] C. PEZZATO et al., "An efficient artificial molecular pump", *Tetrahedron* **73.33** (2017), 4849–4857.
- [52] C. PEZZATO et al., "Controlling Dual Molecular Pumps Electrochemically", *Angew. Chem. Int. Ed.* **57.30** (2018), 9325–9329.

- [53] S. ERBAS-CAKMAK et al., "Rotary and linear molecular motors driven by pulses of a chemical fuel", *Science* **358**.6361 (2017), 340–343.
- [54] L. FENG et al., "Active mechanisorption driven by pumping cassettes", *Science* **374**.6572 (2021), 1215–1221.
- [55] D. THOMAS et al., "Pumping Between Phases With a Pulsed-Fuel Molecular Ratchet", *ChemRxiv* **2021-fl7tv**. (2021).
- [56] J. BERNÁ et al., "Macroscopic transport by synthetic molecular machines", *Nat. Mater.* **4**.9 (2005), 704–710.
- [57] R. EELKEMA et al., "Nanomotor rotates microscale objects", *Nature* **440**.7081 (2006), 163–163.
- [58] Q. LI et al., "Macroscopic contraction of a gel induced by the integrated motion of light-driven molecular motors", *Nat. Nanotechnol.* **10**.2 (2015), 161–165.
- [59] J. T. FOY et al., "Dual-light control of nanomachines that integrate motor and modulator subunits", *Nat. Nanotechnol.* **12**.6 (2017), 540–545.
- [60] M. KATHAN et al., "A light-fuelled nanoratchet shifts a coupled chemical equilibrium", *Nat. Nanotechnol.* (2021).
- [61] Y. LIU et al., "Linear Artificial Molecular Muscles", *J. Am. Chem. Soc.* **127**.27 (2005), 9745–9759.
- [62] R. F. LUDLOW and S. OTTO, "Systems chemistry", *Chem. Soc. Rev.* **37**. (2008), 101–108.
- [63] J. F. STODDART, "Editorial: From Supramolecular to Systems Chemistry: Complexity Emerging out of Simplicity", *Angew. Chem. Int. Ed.* **51**.52 (2012), 12902–12903.
- [64] E. MATTIA and S. OTTO, "Supramolecular systems chemistry", *Nat. Nanotechnol.* **10**.2 (2015), 111–119.
- [65] G. ASHKENASY et al., "Systems chemistry", *Chem. Soc. Rev.* **46**. (2017), 2543–2554.
- [66] K. B. MUCHOWSKA, S. J. VARMA and J. MORAN, "Synthesis and breakdown of universal metabolic precursors promoted by iron", *Nature* **569**.7754 (2019), 104–107.
- [67] K. B. MUCHOWSKA, S. J. VARMA and J. MORAN, "Nonenzymatic Metabolic Reactions and Life's Origins", *Chem. Rev.* **120**.15 (2020), 7708–7744.
- [68] K. DAS, L. GABRIELLI and L. J. PRINS, "Chemically Fueled Self-Assembly in Biology and Chemistry", *Angew. Chem. Int. Ed.* **60**.37 (2021), 20120–20143.
- [69] J. LEIRA-IGLESIAS et al., "Oscillations, travelling fronts and patterns in a supramolecular system", *Nat. Nanotechnol.* **13**.11 (2018), 1021–1027.
- [70] A. SORRENTI et al., "Non-equilibrium steady states in supramolecular polymerization", *Nat. Commun.* **8**.1 (2017), 15899.
- [71] A. SORRENTI et al., "Non-equilibrium supramolecular polymerization", *Chem. Soc. Rev.* **46**. (2017), 5476–5490.
- [72] A. DESAI and T. J. MITCHISON, "Microtubule Polymerization Dynamics", *Annu. Rev. Cell Dev. Biol.* **13**.1 (1997), 83–117.
- [73] H. HESS and J. L. ROSS, "Non-equilibrium assembly of microtubules: from molecules to autonomous chemical robots", *Chem. Soc. Rev.* **46**. (2017), 5570–5587.

- [74] J. HOWARD, *Mechanics of Motor Proteins and the Cytoskeleton*, Sunderland, MA: Sinauer Associates, 2001.
- [75] Z. YANG et al., "Using a Kinase/Phosphatase Switch to Regulate a Supramolecular Hydrogel and Forming the Supramolecular Hydrogel in Vivo", *J. Am. Chem. Soc.* **128.9** (2006), 3038–3043.
- [76] C. PEZZATO and L. J. PRINS, "Transient signal generation in a self-assembled nanosystem fueled by ATP", *Nat. Commun.* **6.1** (2015), 7790.
- [77] S. MAITI et al., "Dissipative self-assembly of vesicular nanoreactors", *Nat. Chem.* **8.7** (2016), 725–731.
- [78] J. LEIRA-IGLESIAS et al., "Supramolecular pathway selection of perylene-dimides mediated by chemical fuels", *Chem. Commun.* **52.** (2016), 9009–9012.
- [79] S. DHIMAN, A. JAIN and S. J. GEORGE, "Transient Helicity: Fuel-Driven Temporal Control over Conformational Switching in a Supramolecular Polymer", *Angew. Chem. Int. Ed.* **56.5** (2017), 1329–1333.
- [80] G. RAGAZZON and L. J. PRINS, "Energy consumption in chemical fuel-driven self-assembly", *Nat. Nanotechnol.* **13.10** (2018), 882–889.
- [81] R. D. ASTUMIAN, "Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet", *Nat. Commun.* **10.1** (2019), 3837.
- [82] "Self-assembling life", *Nat. Nanotechnol.* **11.11** (2016), 909–909.

1.3 THERMODYNAMICS OF CHEMICAL SYSTEMS

Imagine an A intimately united with a B, so that no force is able to sunder them; imagine a C likewise related to a D; now bring the two couples into contact: A will throw itself at D, C at B, without our being able to say which first deserted its partner, which first embraced the other's partner.

— Johann Wolfgang von Goethe, *Elective Affinities*
(Penguin Group, 1971. Original version in German: 1809)

The present thesis started from asking whether it is possible to provide a thermodynamic analysis of the chemical engines discussed in Section 1.2 that is similar in its extent to the one that classical thermodynamics formulated for macroscopic engines, which we discussed at the end of Section 1.1. Therefore, what we seek is a theoretical framework able to deal with chemical reaction networks of various complexity – from small molecular motors to metabolism – which exchange and process energy and/or matter in different forms (chemicals, light, heat). Ideally, we also want to recover a picture where chemical engines are powered by virtue of some forces externally imposed by the reservoir and that one can in principle control as with the temperature gradient of an heat engine. This will allow us to define the power input of chemical engines and, together with the notion of work delivered by a chemical system, to define efficiencies on the basis of which to evaluate their performance.

Historically, the driving force that powers chemical reactions has been referred to as “affinity”. To use the words⁴ of René Marcellin (1885-1914), “the notion of affinity first appeared in science with a metaphysical character” [1], and the word has been used for centuries to name different observables. This probably contributes to a certain aversion for the use of the term in the current scientific literature. In this thesis, the contemporary concept of affinity [2] as it developed in chemical thermodynamics will be extensively used. Actually, it can be argued that the approach followed in this work is rooted into the many efforts to quantify the affinity and, therefore, to predict the direction of a chemical reaction.

As mentioned at the end of the previous Section, most of the crucial steps needed to conceptually treat chemical engines at the thermodynamic level have been accomplished. In the following, we review the key assumptions and results which will be used throughout the following chapters. This will also allow us to systematically introduce the notation and further contextualize the original contributions of the thesis. Notice that logical order is preferred to the historical one when necessary and, especially when reviewing the oldest results, their modern formulation will be used in order to give a consistent and easier-to-follow presentation.

1.3.1 Mass action kinetics

Consider a generic double replacement reaction as it typically occurs when an acid reacts with a base or when two salts react together in solution – we will not go as far as Goethe in saying that this process can even represent love affairs:



⁴ Freely translated from French.

The reaction mechanism will be considered *elementary*, that is both the forward and the backward reactions happen in one step passing through the same fast intermediate state (the activated state) [3–5]. When the number of reacting molecules is high, they are well mixed and their concentrations are small compared to the solvent, the reaction current in isothermal conditions is usually very well described by *mass action kinetics*:

$$J = \underbrace{k_+ z_{AB} z_{CD}}_{J_+} - \underbrace{k_- z_{AD} z_{CB}}_{J_-}, \quad (8)$$

where the z_σ 's denote concentrations (moles per unit of volume) and the k 's are reaction-specific rate constants – originally called “affinity constants” [6] – which are positive and depend on the temperature T . Their expression was first formulated by Svante Arrhenius (1859–1927) based on experimental data [7]

$$k_\pm = C_A e^{\frac{E_\pm}{RT}} \quad (9)$$

where C_A is a factor common to both the forward and the backward reaction, and E_\pm is called activation energy of the reaction. The current J is typically expressed in units of $\text{molV}^{-1}\text{s}^{-1}$, and measures the net rate at which reactants are converted into products. For instance, the concentrations of the species in reaction (7) will vary in time according to

$$-d_t^i z_{AB} = -d_t^i z_{CD} = d_t^i z_{AD} = d_t^i z_{CB} = J, \quad (10)$$

with the superscript “i” indicating that concentrations change due to processes that are internal to the system and not due to exchanges with external reservoirs of matter. As we will see in Chapter 2, other expressions may be possible for reaction currents which better fit particular experimental conditions, but from now on mass action kinetics will be assumed unless otherwise specified. This is consistent with the *ideal dilute solution model*, which treats solutions as a mixture of ideal gases. In fact, mass action kinetics as in equation (8) was first inferred from experimental data [6], and only later justified with a microscopic hypothesis rooted into collision theory [8]. As easily checked via equation (10), the following quantities are constant over time as reaction (7) takes place:

$$L_A = z_{AB} + z_{AD}, \quad L_C = z_{CD} + z_{CB}, \quad L_D = z_{AD} + z_{CD} \quad (11)$$

and account for the fact that the total amounts of moieties A, C, and D are independently conserved whatever the system's state. Note that the total amount of B is conserved too, but the corresponding *conservation law* is already encoded in equations 11 (indeed: $L_A + L_C - L_D = z_{AB} + z_{CB}$).

We are now in the position to define the concept of *chemical equilibrium* from a kinetic viewpoint, that is the condition of null current:

$$J^{(\text{eq})} = \underbrace{k_+ z_{AB}^{(\text{eq})} z_{CD}^{(\text{eq})}}_{J_+^{(\text{eq})}} - \underbrace{k_- z_{AD}^{(\text{eq})} z_{CB}^{(\text{eq})}}_{J_-^{(\text{eq})}} = 0. \quad (12)$$

The above condition ($J_+ = J_-$) is also known as *detailed balance* [9] or *principle of microscopic reversibility*, first introduced by Richard Tolman (1881–1948) as “a new postulate which perhaps cannot yet be stated in its final form, but which requires in a general way in the case of a system in thermodynamic equilibrium not only that the total number of molecules leaving a given

state in unit time shall on the average equal the number arriving in that state in unit time, but also that the number leaving by any particular path $[(J_+)]$ shall on the average be equal to the number arriving by the reverse of that particular path $[(J_-)]$ " [4]. Since equations (11) and (12) are four independent constraints that the four concentrations must satisfy, given the values of the conserved quantities the chemical equilibrium is unique for mass action kinetics. It follows that, at equilibrium:

$$Q^{(\text{eq})} \equiv \frac{z_{\text{AD}}^{(\text{eq})} z_{\text{CB}}^{(\text{eq})}}{z_{\text{AB}}^{(\text{eq})} z_{\text{CD}}^{(\text{eq})}} = \frac{k_+}{k_-} \equiv K_{\text{eq}}. \quad (13)$$

The ratio of kinetic rates (K_{eq}) is called equilibrium constant of the reaction, while the ratio on the left hand side ($Q = z_{\text{AD}}z_{\text{CB}}/z_{\text{AB}}z_{\text{CD}}$) is called the reaction quotient. As shown by Lars Onsager (1903-1976) with his celebrated reciprocal relations [10], the principle of microscopic reversibility has consequences which go beyond the equilibrium state and which we will further explore in Section 1.3.5.

In contrast with the terminology of the original works introducing mass action kinetics – in which the affinity which causes reactants (products) to react is quantified as J_+ (J_-) –, it is useful to introduce here the modern expression for the chemical affinity of reaction from the kinetic viewpoint:

$$A \equiv RT \ln K_{\text{eq}} - RT \ln Q = RT \ln \frac{J_+}{J_-}, \quad (14)$$

where R is the gas constant. Whenever $A \neq 0$ ($Q \neq K_{\text{eq}}$), the reaction is in a nonequilibrium state evolving forward towards the product ($J > 0$) if $A > 0$ ($Q < K_{\text{eq}}$) or backward towards the reactants ($J < 0$) if $A < 0$ ($Q > K_{\text{eq}}$).

1.3.2 Chemical potentials

The current formulation of the laws of thermodynamics for chemical systems is mainly due to Josiah Willard Gibbs (1839-1903). In his seminal work "On the Equilibrium of Heterogeneous Substances" [11] published between 1875 and 1878, he considers – among other examples – a solution of chemical species as a system in contact with a single reservoir at temperature T and pressure P (e.g., a solution in a beaker surrounded by air). For such a system, the balances of internal energy (equation (1)) and entropy (equation (2)) introduced in Section 1.1 boil down to

$$d_t U = \dot{Q} + \dot{W} \quad (15)$$

and

$$d_t S = \frac{\dot{Q}}{T} + \dot{\Sigma}, \quad (16)$$

which can be combined into the following equation

$$d_t U = T d_t S + \dot{W} - T \dot{\Sigma}. \quad (17)$$

One of the key ideas of Gibbs has been to express \dot{W} in terms of chemical contributions adding up to the mechanical one:

$$\dot{W} = \underbrace{-P d_t V}_{\dot{W}_{\text{mech}}} + \underbrace{\sum_{\sigma} \mu_{\sigma} d_t^c N_{\sigma}}_{\dot{W}_{\text{chem}}}, \quad (18)$$

where V is the volume of the solution, while N_σ and μ_σ are respectively the number of moles and the *chemical potential* of species σ . It should be stressed here that Gibbs did not account for the possibility that the amount of a species changes due to chemical reactions – which are nonequilibrium processes not controllable from the exterior –, but he only considered variations to external additions or extractions, as denoted by the superscript “e” in $d_t^e N_\sigma$. For a quasi-static equilibrium process ($\dot{\Sigma} = 0$), we can therefore write

$$dU^{(\text{eq})} = TdS^{(\text{eq})} - PdV^{(\text{eq})} + \sum_{\sigma} \mu_{\sigma}^{(\text{eq})} d^e N_{\sigma}^{(\text{eq})}. \quad (19)$$

In the above expression and in the following, the superscript “(eq)” stresses that equation (19) is only valid for those idealized processes where the system is at any time in equilibrium with the reservoirs. Since such a transformation can only be realized in an infinite amount of time, we also switched from time derivatives d_t to differentials d , as we are now looking to how U varies for an infinitesimal variation of the extensive variables. To fix ideas, as an example of equilibrium process one can imagine that an infinitesimal amount of solvent S (assumed as chemically inert) is very slowly (i.e., quasi-statically) added to the solution such to cause an infinitesimal variation of the volume (dV) and of its own mole number (dN_S). In such a process, $-PdV$ and $\mu_S dN_S$ are respectively the mechanical and the chemical infinitesimal work performed on the system by the external manipulation (delivered by the system when negative). From equation (19), $\mu_{\sigma}^{(\text{eq})}$ can be formally defined as

$$\mu_{\sigma}^{(\text{eq})} = \left(\frac{\partial U^{(\text{eq})}}{\partial N_{\sigma}} \right)_{S, V, N_{j \neq \sigma}}. \quad (20)$$

To obtain a more useful expression of the chemical potentials, we first notice that the internal energy U is homogeneous of degree one in the extensive variables S , V and $\{N_{\sigma}\}$. By applying Euler’s theorem for homogeneous functions [12], we can therefore write:

$$U^{(\text{eq})} = TS^{(\text{eq})} - PV^{(\text{eq})} + \sum_{\sigma} \mu_{\sigma}^{(\text{eq})} N_{\sigma}^{(\text{eq})}, \quad (21)$$

from which we get

$$S^{(\text{eq})} = \frac{U^{(\text{eq})}}{T} + \frac{PV^{(\text{eq})}}{T} - \sum_{\sigma} \frac{\mu_{\sigma}^{(\text{eq})} N_{\sigma}^{(\text{eq})}}{T}. \quad (22)$$

If we now differentiate the last expression and compare it with equation (19), we deduce one possible representation of the well known Gibbs-Duhem relation:

$$U^{(\text{eq})} d \left(\frac{1}{T} \right) + V d \left(\frac{P}{T} \right) - \sum_{\sigma} N_{\sigma}^{(\text{eq})} d \left(\frac{\mu_{\sigma}^{(\text{eq})}}{T} \right) = 0, \quad (23)$$

which bounds the variations of chemical potentials in the system with the quasi-static (equilibrium) manipulation of T and P . We now resort to the ideal dilute solution model, thus considering each species σ to independently behave as an ideal gas obeying the following well known equations of state:

$$P_{\sigma}^{(\text{eq})} V = N_{\sigma}^{(\text{eq})} RT \quad (24a)$$

$$U_{\sigma}^{(\text{eq})} = \frac{f_{\sigma}}{2} N_{\sigma}^{(\text{eq})} RT, \quad (24b)$$

where $P_\sigma^{(\text{eq})} = PN_\sigma^{(\text{eq})} / \sum_\sigma N_\sigma^{(\text{eq})}$ is the partial pressure of species σ and f_σ is the numbers of nuclear degrees of freedom of the molecular species constituting the gas. Notice that, by virtue of the independence assumption we have $U^{(\text{eq})} = \sum_\sigma U_\sigma^{(\text{eq})}$, $S^{(\text{eq})} = \sum_\sigma S_\sigma^{(\text{eq})}$, $P = \sum_\sigma P_\sigma^{(\text{eq})}$ and $N^{(\text{eq})} = \sum_\sigma N_\sigma^{(\text{eq})}$, and the above equations (15) and (16) can be specialized for each species σ . By using relations (23), we can therefore write, for a single species, the following equation

$$N_\sigma^{(\text{eq})} d\left(\frac{\mu_\sigma^{(\text{eq})}}{T}\right) = U_\sigma^{(\text{eq})} d\left(\frac{1}{T}\right) + V^{(\text{eq})} d\left(\frac{P_\sigma^{(\text{eq})}}{T}\right), \quad (25)$$

which, upon substitution of the equations of state (24), leads to

$$d\left(\frac{\mu_\sigma^{(\text{eq})}}{T}\right) = \frac{f_\sigma}{2} RT d\left(\ln \frac{1}{T}\right) + RT d\left(\frac{\ln P_\sigma^{(\text{eq})}}{T}\right). \quad (26)$$

The above equation can now be integrated to obtain

$$\left(\frac{\mu_\sigma^{(\text{eq})}}{T}\right) = \frac{f_\sigma}{2} RT \left(\ln \frac{T_0}{T}\right) + RT \left(\ln \frac{P_\sigma^{(\text{eq})} T_0}{P_{0,\sigma}^{(\text{eq})} T}\right) + \left(\frac{\mu_{0,\sigma}^{(\text{eq})}}{T}\right), \quad (27)$$

where the subscript 0 denotes an arbitrary reference equilibrium state from which the integration started. By recalling that $P_\sigma^{(\text{eq})} = PN_\sigma^{(\text{eq})} / \sum_\sigma N_\sigma^{(\text{eq})}$, we can nicely recast the equilibrium chemical potential of a species σ as

$$\mu_\sigma^{(\text{eq})} = \mu_\sigma^0 + RT \ln \frac{N_\sigma^{(\text{eq})}}{\sum_\sigma N_\sigma^{(\text{eq})}}, \quad (28)$$

where μ_σ^0 is an arbitrary reference chemical potential which only depend on T , P and on the chosen reference state.

1.3.3 Gibbs free energy

Often in chemistry, systems are operated in practically isothermal, isobaric, and isochoric conditions. By using another idea of Gibbs, we define the Legendre transform of the internal energy named *Gibbs free energy* (G), whose balance equation reads

$$d_t G = -S d_t T + V d_t P + \sum_\sigma \mu_\sigma d_t^e N_\sigma - T \dot{\Sigma}. \quad (29)$$

Its meaningfulness for chemistry is readily demonstrated by considering a typical process happening at costant temperature and pressure in finite time ($d_t T = d_t P = 0$):

$$\dot{W}_{\text{chem}} - d_t G = T \dot{\Sigma} \geq 0. \quad (30)$$

In words, the time variation of G corresponds to the amount of chemical power ($\dot{W}_{\text{chem}} = \sum_\sigma \mu_\sigma d_t^e N_\sigma$) injected in the system which does not get dissipated as $T \dot{\Sigma}$. In the limit of an equilibrium process ($\dot{\Sigma} = 0$), all the chemical work is used to shift the Gibbs free energy and, vice versa, the time variation of the Gibbs free energy corresponds to the maximum chemical power that can be extracted from the system:

$$\sum_\sigma \mu_\sigma^{(\text{eq})} d^e N_\sigma^{(\text{eq})} = dG^{(\text{eq})}. \quad (31)$$

Since the Gibbs free energy is a homogeneous function of degree one in $\{N_\sigma\}$, its equilibrium expression can be found by applying Euler's theorem to equation (31):

$$G^{(\text{eq})} = \sum_{\sigma} \mu_{\sigma}^{(\text{eq})} N_{\sigma}^{(\text{eq})} = \sum_{\sigma} \left(\mu_{\sigma}^0 + RT \ln \frac{N_{\sigma}^{(\text{eq})}}{\sum_{\sigma} N_{\sigma}^{(\text{eq})}} \right) N_{\sigma}^{(\text{eq})}. \quad (32)$$

In the literature about synthetic chemical engines, systems are often characterized in terms of species concentrations $z_{\sigma} = N_{\sigma}/V$. It will be therefore more convenient to work with the Gibbs free energy per unit of volume ($G_V^{(\text{eq})} = G^{(\text{eq})}/V$):

$$G_V^{(\text{eq})} = \sum_{\sigma} \left(\mu_{\sigma}^0 + RT \ln \frac{z_{\sigma}^{(\text{eq})}}{\sum_{\sigma} z_{\sigma}^{(\text{eq})}} \right) z_{\sigma}^{(\text{eq})}. \quad (33)$$

We can now further exploit the ideal dilute solution model by considering that $\sum_{\sigma} z_{\sigma} \approx z_S$ (we remind that S denotes the solvent), hence:

$$G_V^{(\text{eq})} = \sum_{\sigma \neq S} \left(\mu_{\sigma}^0 + RT \ln \frac{z_{\sigma}^{(\text{eq})}}{z_S} \right) z_{\sigma}^{(\text{eq})} + \left(\mu_S^0 - RT \ln \frac{\sum_{\sigma \neq S} z_{\sigma}^{(\text{eq})} + z_S}{z_S} \right) z_S. \quad (34)$$

Notice that, as the volume V of the solution is dominated by the volume of the solvent S itself, and the latter is assumed to be chemically inert, z_S is a property of the pure solvent which only depends on T and P . As $\sum_{\sigma \neq S} z_{\sigma}^{(\text{eq})} \ll z_S$, the last equation simplifies to:

$$\begin{aligned} G_V^{(\text{eq})} &= \sum_{\sigma \neq S} \left(\mu_{\sigma}^0 + RT \ln \frac{z_{\sigma}^{(\text{eq})}}{z_S} \right) z_{\sigma}^{(\text{eq})} + \mu_S^0 z_S - RT \sum_{\sigma \neq S} z_{\sigma}^{(\text{eq})} \\ &= \sum_{\sigma \neq S} \left(\mu_{\sigma}^0 + RT \ln z_{\sigma}^{(\text{eq})} - RT \right) z_{\sigma}^{(\text{eq})} + \mu_S^0 z_S \end{aligned} \quad (35)$$

where $\mu_{\sigma}^{\circ} = \mu_{\sigma}^0 - RT \ln z_S$ is called *standard chemical potential* of the species σ . Notice that both μ_{σ}° and $\mu_S^0 z_S$ are solvent-specific quantities, but they do not depend on the current chemical equilibrium. In particular, the term $\mu_{\sigma}^0 z_S$ cancels out in most of the practical applications, where differences in the Gibbs free energy along isothermal, isobaric, and isochoric transformations conserving the total amount of solvent are considered.

1.3.4 Local equilibrium

Expressions like equation (35) are strictly valid only when the system is at thermodynamic equilibrium. Despite equilibrium thermodynamics – which could be called *thermostatistics* as all dynamic aspects are made infinitely slow such that time is not an explicit variable [13] – being of extreme theoretical importance and practical usefulness, we are here interested in studying chemical reactions, which are dynamic processes occurring in finite time among generic nonequilibrium states of the systems. As pioneered by Théophile de Donder (1872-1957), Lars Onsager and Ilya Prigogine (1917-2003), the results obtained above can be extended to nonequilibrium processes via the so-called assumption of *local equilibrium* [14–18]. For well mixed ideal dilute solutions of reacting chemical species, assuming

local equilibrium is analogous to assuming that the nonequilibrium nature of the thermodynamic description is solely due to the concentrations' distribution. Thus, what is assumed is that thermal equilibration and spatial equilibration following any reaction event are much faster than any reaction time scale [19–21]. As a consequence, at any time the system is locally in equilibrium with respect to the reservoirs, and all of the corresponding intensive thermodynamic variables are well defined and equal everywhere in the system. This suggests that nonequilibrium states can be described by thermodynamic functions (e.g., G_V , μ_σ) whose expressions are the equilibrium ones, but where equilibrium concentrations are replaced by the actual ones [12, 16, 17]. For instance, the nonequilibrium Gibbs free energy density reads:

$$G_V = \sum_{\sigma \neq S} (\mu_\sigma - RT) z_\sigma + \mu_S^0 z_S = \sum_{\sigma \neq S} (\mu_\sigma^0 + RT \ln z_\sigma - RT) z_\sigma + \mu_S^0 z_S, \quad (36)$$

and the chemical potential for species σ which are not the solvent reads

$$\mu_\sigma = \mu_\sigma^0 + RT \ln z_\sigma. \quad (37)$$

To check the consistency of the local equilibrium assumption and its consequences, consider the simple example of an ideal dilute solution closed to matter exchanges where T , P and V are constant and just reaction (7) takes place. As the system is closed ($d_t^c z_\sigma = 0$), no chemical work is performed and the only changes in time of the concentrations are due to reaction (7) ($d_t z_\sigma = d_t^i z_\sigma$). By computing the time derivative of G_V in equation (36) via equation (10) and plugging it into the general balance (29), we obtain one of the main results of De Donder [14] specialized to the example that we are considering:

$$d_t G_V = \sum_{\sigma \neq S} \mu_\sigma d_t^i z_\sigma = (-\mu_{AB} - \mu_{CD} + \mu_{AD} + \mu_{CB})J = -T\dot{\Sigma}_V \leq 0. \quad (38)$$

Its importance lies in the connection that it draws between the current J of the chemical reaction (a dynamic quantity) and the dissipation rate per unit volume $T\dot{\Sigma}_V$. The proportionality constant can be interpreted as the thermodynamic force driving the reaction, i.e. its affinity from the thermodynamic viewpoint:

$$A \equiv \mu_{AB} + \mu_{CD} - \mu_{AD} - \mu_{CB}. \quad (39)$$

From equation (39), we can give a thermodynamic definition of chemical equilibrium as a state with null affinity:

$$A^{(eq)} = \mu_{AB}^{(eq)} + \mu_{CD}^{(eq)} - \mu_{AD}^{(eq)} - \mu_{CB}^{(eq)} = 0. \quad (40)$$

Another important consequence of equation (38) is that the time derivative of the Gibbs free energy must be always negative except when the system is at equilibrium. We stress that this is a condition on G_V and not a property of the expression we assumed based on the local equilibrium. We will therefore need to verify or make sure that this condition holds, a problem which is discussed below (see Section 1.3.5). We now show that G_V as defined in equation (36) has the property to be lower bounded by its equilibrium value G_V^{eq} when the system is closed. Indeed:

$$G_V - G_V^{(eq)} = \sum_{\sigma \neq S} \mu_\sigma z_\sigma - \mu_\sigma^{(eq)} z_\sigma^{(eq)} - RT(z_\sigma - z_\sigma^{(eq)}), \quad (41)$$

which follows directly from equation (36). To proceed, we exploit the conservation laws in equations (11) to prove that

$$\begin{aligned}
 \sum_{\sigma \neq S} \mu_{\sigma}^{(\text{eq})} z_{\sigma}^{(\text{eq})} &= \mu_{\text{AB}}^{(\text{eq})} z_{\text{AB}}^{(\text{eq})} + \mu_{\text{CD}}^{(\text{eq})} z_{\text{CD}}^{(\text{eq})} + \mu_{\text{AD}}^{(\text{eq})} z_{\text{AD}}^{(\text{eq})} + \mu_{\text{CB}}^{(\text{eq})} z_{\text{CB}}^{(\text{eq})} = \\
 &= L_{\text{A}} \mu_{\text{AB}}^{(\text{eq})} + L_{\text{D}} \mu_{\text{CD}}^{(\text{eq})} + (L_{\text{C}} - L_{\text{D}}) \mu_{\text{CB}}^{(\text{eq})} + z_{\text{AD}}^{(\text{eq})} \underbrace{(\mu_{\text{AD}}^{(\text{eq})} + \mu_{\text{CB}}^{(\text{eq})} - \mu_{\text{AB}}^{(\text{eq})} - \mu_{\text{CD}}^{(\text{eq})})}_{=0} = \\
 &= \sum_{\sigma \neq S} \mu_{\sigma}^{(\text{eq})} z_{\sigma}. \tag{42}
 \end{aligned}$$

By plugging the above result into equation (41), we find

$$G_{\text{V}} - G_{\text{V}}^{(\text{eq})} = RT \sum_{\sigma \neq S} z_{\sigma} \ln \frac{z_{\sigma}}{z_{\sigma}^{(\text{eq})}} - (z_{\sigma} - z_{\sigma}^{(\text{eq})}), \tag{43}$$

where we also used equation (37). The difference in equation (43) is never negative by virtue of the logarithm inequality ($\ln x \leq x - 1$), thus proving that $G_{\text{V}}^{(\text{eq})}$ is the minimum possible value of G_{V} as long as the system stays closed to matter exchanges. As we will see in Chapter 2, this result is not restricted to this specific example and is valid as long as ideal dilute solutions are considered. Moreover, equation (43) and equation (38) together show that the Gibbs free-energy density G_{V} acts as a Lyapunov function for the systems' dynamics [22–24], meaning that the time evolution of the concentrations distribution always minimizes G_{V} – this, as we will see, as long as the system remains closed to matter exchanges. For this reason, G_{V} is said to be the proper *thermodynamic potential* for closed systems [12].

We conclude this section by noticing that, by virtue of equation (30) integrated over time, the quantity $G_{\text{V}} - G_{\text{V}}^{(\text{eq})}$ can be identified as the free energy stored per unit volume in a system with a certain concentrations distribution $\{z_{\sigma}\}$. Indeed, this quantity coincides with the free-energy released by the system per unit volume while relaxing from $\{z_{\sigma}\}$ to $\{z_{\sigma}^{(\text{eq})}\}$, which in turn quantifies the maximum amount of chemical work that can in principle be extracted from the process.

1.3.5 Local Detailed Balance

As anticipated in the previous section, we still need to verify or make sure that the expression in equation (36) that we picked for G_{V} satisfies equation (38) ($d_{\text{t}} G_{\text{V}} = -T \dot{\Sigma}_{\text{V}} < 0$, except at equilibrium), i.e., it is consistent with the second principle of thermodynamics. To do so, based on the empirical evidence, we require that the kinetic and thermodynamic definition of chemical equilibrium coincide, that is equivalent – in the local equilibrium assumption – to impose that the kinetic definition of affinity in equation (14) coincides with the thermodynamic definition in equation (39). For reaction (7) undergoing mass action kinetics, due to the principle of microscopic reversibility the latter requirement is satisfied if and only if

$$RT \ln K_{\text{eq}} = RT \ln \frac{k_{+}}{k_{-}} = \mu_{\text{AB}}^{\circ} + \mu_{\text{CD}}^{\circ} - \mu_{\text{AD}}^{\circ} - \mu_{\text{CB}}^{\circ}. \tag{44}$$

The above condition guarantees that the system relaxes to thermodynamic equilibrium when closed, and that the latter is exactly the unique detailed

balance state from a kinetic viewpoint. Indeed, by combining equations (38) and (44), we get

$$d_t G_V = -T \dot{\Sigma}_V = -(J_+ - J_-) RT \ln \frac{J_+}{J_-} \leq 0, \quad (45)$$

where the inequality is automatically verified.

The condition in equation (44) exemplifies the general idea that, in order for a model to be thermodynamically consistent, the ratio of forward and backward rate constants of a single reaction must obey a precise constraint. The first person who realized this was probably Marcelin, who gave a formulation of the constraint in equation (44) in terms of what it is now called the standard Gibbs energy of activation⁵, which he introduced as the free energy of the common activated state that both the forward and the backward pathway of an elementary reaction have to overcome in order to happen. Indeed, based on an argument of thermodynamic consistency with a result obtained by Jacobus van 't Hoff (1852-1911) about the temperature dependence of the equilibrium constant, Marcelin concluded that the energy of the molecules in the activated state has to be the same regardless of the direction in which the reaction is taking place [1, 4]. Marcelin also introduced the concepts of potential energy surface and reaction coordinates to describe a chemical reaction in terms of a point moving in a state space [25]. His work can be considered as the first theoretical treatment of the rates of a chemical reaction justifying Arrhenius equation and as the groundwork for the development of transition state theory in the thirties of the 20th century [26].

From condition (44), we can also derive the following expression for the affinity of a reaction

$$A = RT \ln \frac{J_+}{J_-}, \quad (46)$$

which is sometimes called the Marcelin-De Donder equation [27, 28] or flux-force relation [29]. It holds for all elementary reactions following mass-action kinetics, but it has been shown that under certain conditions also coarse-grained non-elementary reactions can satisfy it [29, 30]. Actually, it has been proposed to redefine the notion of elementary reaction on a thermodynamic ground – instead of on a mechanistic one as it is traditionally done [5] – as those reactions whose fluxes obey equation (46) [27, 31].

Equation (46) will be largely exploited in the following chapters and it is central in the most recent development of chemical thermodynamics. It clearly shows that for elementary reactions the sign of the current always coincides with that of the affinity. Surprisingly, to the best of our knowledge De Donder did not use it up to 1936 even if he had all the elements to write it, while Prigogine recognized in 1955 that “this equation expresses fairly generally the relation between reaction rate and affinity” [32]. However, up to 1967 Prigogine was also admitting the possibility that the product of the reaction affinity and the associated current ($A_\rho J_\rho$, where the subscript ρ denoted a specific reaction) can be negative in case the reaction is coupled with another one, without specifying that this is not true for elementary reactions (commonly assumed in theoretical works) [33]. This kind of misleading claims disappeared in later works [12], suggesting that the relevance of equation (46) emerged with time. More recently, equation (46) started to be broadly assumed in part of the biochemical literature [34–36] even if its

⁵ Marcelin used the term “equilibrium affinity” for it, which we avoid here for obvious reasons of clarity.

validity is not granted for non elementary reactions and always needs to be proven [29, 30, 37].

In this thesis, condition (44) will be denoted as *local detailed balance*. Such terminology has been introduced for the first time in a 1983 paper by Katz, Lebowitz and Spohn [38], but the concept is rooted in the attempts made by the statistical mechanics community to model generic nonequilibrium systems. According to Christian Maes [39], the idea of local detailed balance traces back in the works from Bergmann and Lebowitz in 1955 [40]. In that context, local detailed balance is the name given to a constructive principle for physically sound models describing nonequilibrium phenomena.

For cyclic chemical processes, local detailed balance implies a constraint derived by Onsager in 1931 as a direct consequence of microscopic reversibility [10]. This is the reason why the two concepts are sometimes superimposed in the literature. Here, we will keep them distinct by considering microscopic reversibility as a mechanistic constraints and local detailed balance as a thermodynamic constraint. Such a distinction will be relevant in Chapter 3, where photochemical systems will be analyzed. Indeed, one can argue that elementary photophysical processes do not obey microscopic reversibility in the original sense [41], but as we will show they are still constrained by local detailed balance.

1.3.6 Open systems

Until now, our thermodynamic description of nonequilibrium chemical processes has been limited to closed systems in contact with a unique thermal reservoir, whose only possible long-time behavior is a relaxation towards equilibrium. However, chemical engines as those described in Section 1.2 are systems which receive some input power from the environment in the form of light or matter and process it to maintain nonequilibrium steady states. As anticipated, the problem of how to include light in the description will be faced in Chapter 3. Here, we introduce the basic concepts needed to treat open chemical systems continuously exchanging matter with the exterior, where terms like $d_t^e z_\sigma$ are not null. In particular, we are interested in those systems where the exchange of matter is such to fix the chemical potentials of some of the species in solution (and therefore their concentrations in the ideal dilute solution model, see equation (37)). This kind of openness is widespread in the biological context, where homeostatic processes tune the concentrations of many species (e.g., the pH window is kept quite narrow), but it is very common also in laboratory experiments. The continuous flow reactor mentioned in Section 1.2.4 (see Figure 9) is an example of chemiosmotic regulation through semipermeable membranes, while other mechanisms allowing to control concentrations are buffer solutions, the continuous addition of a species to compensate for its depletion and the continuous removal of a species to prevent its accumulation (this can be done with scavenger species or by evaporation when the species to eliminate are volatile). In some cases, when a species is very abundant, its concentration can also be considered as constant since variations in its concentration are negligible on the time scale of interest. All these cases will be modeled with the concept of *chemostats*: large chemical reservoirs characterized by a chemical potential in contact and constantly in equilibrium with the system. A species which is exchanged is therefore said to be chemostatted, and its concentration (z_c) is controlled according to:

$$\pi_c = d_t z_c = d_t^i z_c + d_t^e z_c = d_t^i z_c + I_c, \quad (47)$$

where $I_c = \pi_c - d_t^1 z_c$ is the exchange current with the corresponding chemostat. The term π_c defines the protocol by which the concentration z_c is regulated, and we will always consider $\pi_c = 0$ (constant chemostatting) unless otherwise specified.

When the system is open to matter exchanges, the Gibbs free energy density as defined in equation (36) ceases in general to be a proper thermodynamic potential. As pointed out by Alberty [42–45], this is analogous to what happens in statistical mechanics when passing from the canonical to the grand canonical-ensemble. There, the grand-canonical partition function proves to be the right thermodynamic potential. In open chemical systems, as usually only some of the species are exchanged, the proper thermodynamic potential is commonly called *semi-grand Gibbs free energy density*, here denoted \mathcal{G}_V . As an illustration, consider the reaction (7) in a solution where the species AB is chemostatted. From equation (30), we have:

$$\mu_{AB} I_{AB} - d_t G_V = T \dot{\Sigma}_V \geq 0, \quad (48)$$

where $I_{AB} = -J$ by virtue of equations (47) and (10), as $d_t z_{AB} = -J + I_{AB}$. From equation (48), it is not granted anymore that G_V monotonically decreases in time. As a consequence, the Gibbs free energy density does not play the role of a Lyapunov function for the open system dynamics. However, by noticing that $d_t L_A = I_{AB}$, we can write

$$-d_t \mathcal{G}_V = -d_t (G_V - \mu_{AB} L_A) = T \dot{\Sigma}_V \geq 0, \quad (49)$$

from which it follows that the semi-grand Gibbs free energy density $\mathcal{G}_V = G_V - \mu_{AB} L_A$ is a state function whose time derivative is never positive. Moreover, based on what has been discussed in Section 1.3.4, it is straightforward that

$$\mathcal{G}_V - \mathcal{G}_V^{(eq)} = RT \sum_{\sigma \neq S, AB} z_\sigma \ln \frac{z_\sigma}{z_\sigma^{(eq)}} - (z_\sigma - z_\sigma^{(eq)}). \quad (50)$$

It follows that \mathcal{G}_V is lower bounded by its equilibrium value, and therefore it is a Lyapunov function for the open dynamics. With this, we proved the non-trivial result that the open system will relax to the equilibrium state minimizing \mathcal{G}_V in the long time limit, which acts as a proper thermodynamic potential. At equilibrium, the concentrations will satisfy

$$\frac{z_{AD}^{(eq)} z_{CB}^{(eq)}}{z_{CD}^{(eq)}} = \frac{k_+}{k_-} z_{AB}, \quad (51)$$

showing that the chemostatted concentration z_{AB} acts as an external parameter tuning the equilibrium distribution. Indeed, by changing z_{AB} , one can for instance shift the equilibrium distribution towards high or low values of $z_{CD}^{(eq)}$. This is a working principle of many molecular machines (called *molecular switches* in the literature) and regulatory mechanisms, which for instance respond to a sudden modification of the pH by relaxing to a new equilibrium. As we will show more generally in Chapter 2, this kind of behavior holds for open systems as long as each chemostatted species can be separately associated to a conserved quantity which is *broken* with respect to the closed system (in our example, L_A gets broken due to the chemostatting of AB). As a consequence, all chemical systems with only one chemostatted species will eventually relax towards equilibrium, as chemostatting always breaks the total mass conservation with respect to the closed system.

In order to drive a chemical system towards nonequilibrium steady states, chemostating must be such as to allow for a continuous flow of matter from one or more input chemostats to one or more output ones mediated by chemical reactions in the system, in analogy with a heat engine mediating a continuous heat flow from the hot reservoir to the cold one. The conditions under which open chemical systems can be maintained in nonequilibrium steady states will be systematically analyzed in Chapter 2. Here, we conclude by noticing that in case of driving of a chemostatted concentration ($\pi_c \neq 0$ in equation (47)), even a single chemostat may be effective in preventing the system to reach equilibrium. In particular, periodic nonequilibrium steady states can be reached when the driving itself is periodic, as in the nonautonomous chemical engines mentioned in Section 1.2. Despite the fact that such kind of experiments will not be specifically analyzed in this thesis, the framework presented in Chapter 2 considers the possibility of nonautonomous driving and is therefore directly applicable to nonautonomous open systems.

1.3.7 Chemical reaction networks

Until now, we discussed general concepts by referring to a single elementary reaction as an example. However, the chemical engines that we aim to study in this thesis are systems based on many species and reactions, namely *chemical reaction networks*. Even though the system-specific approach that we followed to derive the above results can be in principle applied to study more complex systems, it soon becomes very cumbersome, and a general approach to systematically analyze chemical reaction networks is therefore desirable.

In this thesis, we will leverage a rigorous mathematical description of chemical reaction networks mainly formulated by Feinberg [46], Horn, Jackson [47], and Oster [48], who established connections between the topological properties of the network of reactions and its dynamical behaviour. Crucial contributions then came from Terrell Hill (1917-2014) and Schnakenberg, who employed graph-theoretical techniques to characterize dynamic and thermodynamic properties of nonequilibrium steady states in terms of cycles, i.e., cyclic sequences of reactions which leaves unaltered the state of the system [49, 50]. Their works lie at the basis of the current formulation of the thermodynamics of chemical reaction networks from which this thesis got underway [51, 52]. The main advantage of this approach is that it provides general algebraic recipes to express conservation laws, thermodynamic constraints and thermodynamic quantities of in principle any chemical system, provided that the full network is known.

1.3.8 Stochastic thermodynamics

In the following chapters, we will exclusively focus on the so-called *deterministic* chemical reaction networks, namely chemical systems where the number of species is large enough to make it reasonable to focus on the average dynamics. However, it is worth mentioning that starting from the second half of the 20th century a growing interest for small chemical systems appeared, especially due to new experiments probing biochemical processes. This prompted the physical chemistry community to develop stochastic descriptions of the chemical dynamics, since when low numbers of molecules are concerned a deterministic description where concentrations follow mass-

action kinetics loses its predictive power. Instead, small systems are well described in terms of master equations or chemical master equations, describing the time evolution of the probability of observing a molecule in a certain chemical state or the probability of observing a certain population of molecules [53–55]. Hill was among the first to investigate biochemical systems as small fluctuating engines by introducing the concept of *free energy transduction* to describe the average work performed by a chemical force to drive another chemical process against its spontaneous direction [50, 56]. However, his results were limited to linear (i.e., exclusively comprising unimolecular or pseudo-unimolecular reaction) chemical reaction networks at steady state. Nevertheless, Hill’s works directly inspired important results on the thermodynamic constraints that molecular motors must obey, which revealed the key role played by kinetic factors in their design [57, 58].

A more comprehensive thermodynamic description of the fluctuating dynamics of small chemical systems was pioneered by the so-called *Brussels school of thermodynamics*, which following in De Donder’s footsteps contributed to the foundation of what is currently known as *stochastic thermodynamics* [59–61]. Crucially, thanks also to many other contributions, nonlinear stochastic dynamics and their connection with the deterministic description were addressed, but the main focus was still on steady states and their stability [62–64]. These works drew the attention of the physics community towards investigating the laws of thermodynamics at small scales, and the last two decades in particular witnessed remarkable new developments in stochastic thermodynamics, which extended its realm of applicability to the study of generic small fluctuating systems driven arbitrarily far from equilibrium [65–69].

In the realm of stochastic thermodynamics, what is most relevant for this thesis is the recent progress done in the description of chemical reaction networks [70–72], in energy conversion in finite time [73], and in the connections that have been drawn between thermodynamics and information theory [74]. Concerning the former, the most recent stochastic thermodynamic approach to chemical reaction networks [72] is entangled with the deterministic description of chemical reaction networks developed by the research group in which this thesis was carried on [52]. The analysis of conservation laws [75], the mathematical techniques [72, 76], and the coarse-graining strategies [29, 30, 77] first developed for stochastic chemical reaction networks will all be central tools in the following chapters. Concerning energy conversion, we will avail of results from the so called finite time thermodynamics [73, 78] about the study of the trade-offs between efficiency and power [79–82]. In particular, ideas from stochastic thermodynamics have been largely employed to study free-energy transduction in simple models directly inspired to biological chemical engines such as motor proteins and pumps [83–93]. Apart from some tailored analyses [94–96], most of those studies are still far from being quantitatively comparable with the experiments in the biological realm, but nonetheless they will here provide inspirations and comparison when free energy transduction in synthetic chemical engines will be characterized. Finally, the connections with information theory [97] gave birth to information thermodynamics [74, 98], which treats information as a physical quantity and shows how it relates to other thermodynamic quantities such as free energy and entropy. This turned out to be crucial to solve apparent thermodynamic paradoxes as those involving Maxwell’s demons [99]. Recently, information thermodynamics has been proving useful in the understanding of some aspects of free energy trans-

duction in stochastic chemical systems [100, 101] and biological molecular motors in particular [102–104].

1.3.9 Final considerations

One thing that emerges from the above excursus is the centrality of chemistry in the development of thermodynamics over the last century. On the one hand, this is partially due to the importance of chemistry in natural sciences and industrial applications, which gradually took over steam engines as a challenging subject matter for energetic studies. On the other hand, as soon as chemical reactions started to be understood in terms of the kinetic theory of gases, chemistry became a fertile ground for many attempts to reconcile thermodynamics – a theory embracing the temporal asymmetry of phenomena – with a microscopic theory of matter rooted in the time-symmetric classical mechanics. Moreover, while idealized quasi-static transformations purified equilibrium thermodynamics from dissipation, the problem of predicting the direction of chemical reactions forced scientists to directly deal with out-of-equilibrium systems. In this perspective, establishing the validity of local equilibrium was crucial, as it allows to obtain an expression – and not just an inequality – for the entropy production.

Another fruitful peculiarity of chemical systems is that, at odds with other transport phenomena (e.g., the Fourier law for heat conduction) and mechanical engines, they commonly violate the so-called *linear regime*, which assumes fluxes like the chemical current J to be linearly proportional to conjugate thermodynamic forces like the affinity A . In the linear regime, special theorems can be proven such as Onsager’s reciprocal relations and Prigogine’s minimal entropy production variational principle. The fact that even elementary chemical reactions violate the linear regime assumptions pushed people to develop thermodynamics beyond the linear regime, an enterprise which is still ongoing and which this thesis can be ascribed to.

In the following chapters, we will have the opportunity to further review some of the most recent results which will be largely exploited.

REFERENCES FOR CHAPTER 1.3

- [1] MARCELIN, M. R., “Contribution à l’étude de la cinétique physico-chimique”, *Ann. Phys.* **9.3** (1915), 120–231.
- [2] IUPAC, *Quantities, units and symbols in physical chemistry*, Blackwell Science, 1993, p. 50.
- [3] P. KOHNSTAMM and F. SCHEFFER, “Thermodynamic potential and velocities of reaction”, *Proc. Amst. Acad.* **13.** (1911), 789–804.
- [4] R. C. TOLMAN, “The Principle of Microscopic Reversibility”, *Proc. Natl. Acad. Sci. U.S.A.* **11.7** (1925), 436–439.
- [5] P. MULLER, “Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)”, *Pure and Applied Chemistry* **66.5** (1994), 1111.
- [6] C. GULDBERG and P. WAAGE, *Études sur les affinités chimiques*, Brøgger & Christie, 1967.

- [7] K. J. LAIDLER, "The development of the Arrhenius equation", *J. Chem. Educ.* **61.6** (1984), 494.
- [8] C. M. GULDBERG and P. WAAGE, "Ueber die chemische Affinität. § 1. Einleitung", *Journal für Praktische Chemie* **19.1** (1879), 69–114.
- [9] P. W. BRIDGMAN, "Note on the Principle of Detailed Balancing", *Phys. Rev.* **31**. (1928), 101–102.
- [10] L. ONSAGER, "Reciprocal Relations in Irreversible Processes. I." *Phys. Rev.* **37**. (1931), 405–426.
- [11] J. GIBBS, *The Scientific Papers of J. Willard Gibbs, Vol.1: Thermodynamics*, Dover Publications, 1961.
- [12] D. KONDEPUDI and I. PRIGOGINE, *Modern Thermodynamics. From heat engines to dissipative structures*, Wiley, 2015.
- [13] M. TRIBUS, *Thermostatistics and Thermodynamics: An Introduction to Energy, Information and States of Matter, with Engineering Applications*, Van Nostrand, 1961.
- [14] T. DE DONDER and V. RYSSELBERGHE, *L’Affinité*, GauthierVillars, 1936.
- [15] I. PRIGOGINE and R. DEFAY, *Chemical thermodynamics*, Prentice Hall Press, 1954.
- [16] S. R. de GROOT and P. MAZUR, *Non-Equilibrium Thermodynamics*, Dover, 1984.
- [17] J. KEIZER, *Statistical Thermodynamics of Nonequilibrium Processes*, Srpinger-Verlag, 1987.
- [18] R. LEFEVER, "The rehabilitation of irreversible processes and dissipative structures' 50th anniversary", *Philos. Trans. R. Soc. A* **376.2124** (2018), 20170365.
- [19] I. PRIGOGINE, "Le domaine de validité de la thermodynamique des phénomènes irréversibles", *Physica* **15.1** (1949), 272–284.
- [20] I. PRIGOGINE and E. XHROUET, "On the perturbation of maxwell distribution function by chemical reactions in gases", *Physica* **15.11** (1949), 913–932.
- [21] I. PRIGOGINE and M. MAHIEU, "Sur La Perturbation De La Distribution De Maxwell Par Des Réactions Chimiques En Phase Gazeuse", *Physica* **16.1** (1950), 51–64.
- [22] D. SHEAR, "An analog of the Boltzmann H-theorem (a Liapunov function) for systems of coupled chemical reactions", *J. Theor. Biol.* **16.2** (1967), 212–228.
- [23] J. HIGGINS, "Some remarks on Shear's Liapunov function for systems of chemical reactions", *J. Theor. Biol.* **21.3** (1968), 293–304.
- [24] D. F. ANDERSON et al., "Lyapunov Functions, Stationary Distributions, and Non-equilibrium Potential for Reaction Networks", *Bull. Math. Biol.* **77.9** (2015), 1744–1767.
- [25] K. J. LAIDLER, "Rene Marcelin (1885-1914), a short-lived genius of chemical kinetics", *J. Chem. Educ.* **62.11** (1985), 1012.
- [26] K. J. LAIDLER and M. C. KING, "The development of transition-state theory", *J. Phys. Chem* **87.15** (1983), 2657–2664.
- [27] P. VAN RYSSELBERGHE, "Reaction Rates and Affinities", *J. Chem. Phys.* **29.3** (1958), 640–642.

- [28] V. POVAROV, O. SOKOLOVA and I. KARPOVA, "Modified Marcelin-de Donder equations for the kinetics of one-and two-step reversible chemical reactions in nonideal solutions", *Russ. J. Phys. Chem.* **80.11** (2006), 1874–1879.
- [29] A. WACHTEL, R. RAO and M. ESPOSITO, "Thermodynamically consistent coarse graining of biocatalysts beyond Michaelis–Menten", *New J. Phys.* **20.4** (2018), 042002.
- [30] F. AVANZINI, G. FALASCO and M. ESPOSITO, "Thermodynamics of non-elementary chemical reaction networks", *New J. Phys.* **22.9** (2020), 093040.
- [31] C. A. HOLLINGSWORTH, "Kinetics and Equilibria of Complex Reactions", *J. Chem. Phys.* **27.6** (1957), 1346–1348.
- [32] I. PRIGOGINE, *Introduction to Thermodynamics of Irreversible Processes*, Interscience Publishers, New York, 1955.
- [33] I. PRIGOGINE, *Introduction to Thermodynamics of Irreversible Processes*, Interscience Publishers, New York, 1967.
- [34] D. A. BEARD and H. QIAN, "Relationship between Thermodynamic Driving Force and One-Way Fluxes in Reversible Processes", *PLOS ONE* **2.1** (2007), 1–4.
- [35] A. FLAMHOLZ et al., "Glycolytic strategy as a tradeoff between energy yield and protein cost", *Proc. Natl. Acad. Sci. U.S.A.* **110.24** (2013), 10039–10044.
- [36] E. NOOR et al., "Pathway Thermodynamics Highlights Kinetic Obstacles in Central Metabolism", *PLoS Comput. Biol.* **10.2** (2014), 1–12.
- [37] R. FLEMING et al., "Integrated stoichiometric, thermodynamic and kinetic modelling of steady state metabolism", *J. Theor. Biol.* **264.3** (2010), 683–692.
- [38] S. KATZ, J. L. LEBOWITZ and H. SPOHN, "Phase transitions in stationary nonequilibrium states of model lattice systems", *Phys. Rev. B* **28** (1983), 1655–1658.
- [39] C. MAES, "Local detailed balance", *SciPost Phys. Lect. Notes* **32** (2021), 1–17.
- [40] J. L. LEBOWITZ and P. BERGMANN, "New Approach to Nonequilibrium Process", *Physical Review* **99** (1955), 578–587.
- [41] P. MULLER, "Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)", *Pure and Applied Chemistry* **66.5** (1994), 1141.
- [42] R. A. ALBERTY and I. OPPENHEIM, "Fundamental equation for systems in chemical equilibrium", *J. Chem. Phys.* **89.6** (1988), 3689–3693.
- [43] R. A. ALBERTY and I. OPPENHEIM, "Use of semigrand ensembles in chemical equilibrium calculations on complex organic systems", *J. Chem. Phys.* **91.3** (1989), 1824–1828.
- [44] R. A. ALBERTY, "Biochemical reaction equilibria from the point of view of a semigrand partition function", *J. Chem. Phys.* **114.19** (2001), 8270–8274.
- [45] R. A. ALBERTY, *Thermodynamics of Biochemical Reactions*, Wiley, 2003.
- [46] M. FEINBERG, "Complex balancing in general kinetic systems", *Arch. Ration. Mech. Anal.* **49.3** (1972), 187–194.

- [47] F. HORN and R. JACKSON, “General mass action kinetics”, *Arch. Ration. Mech. Anal.* **47.2** (1972), 81–116.
- [48] G. OSTER and A. PERELSON, “Chemical reaction networks”, *IEEE Transactions on Circuits and Systems* **21.6** (1974), 709–721.
- [49] J. SCHNAKENBERG, “Network theory of microscopic and macroscopic behavior of master equation systems”, *Rev. Mod. Phys.* **48.** (1976), 571–585.
- [50] T. L. HILL, *Free energy transduction in biology*, Academic Press, New York, 1977.
- [51] M. POLETTINI and M. ESPOSITO, “Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws”, *J. Chem. Phys.* **141.2** (2014), 024117.
- [52] R. RAO and M. ESPOSITO, “Nonequilibrium Thermodynamics of Chemical Reaction Networks: Wisdom from Stochastic Thermodynamics”, *Phys. Rev. X* **6.** (2016), 041064.
- [53] D. A. MCQUARRIE, “Stochastic approach to chemical kinetics”, *J. Appl. Probab.* **4.3** (1967), 413–478.
- [54] D. T. GILLESPIE, “A rigorous derivation of the chemical master equation”, *Phys. A: Stat. Mech. Appl.* **188.1** (1992), 404–425.
- [55] N. G. van KAMPEN, *Stochastic Processes in Physics and Chemistry*, Elsevier, 2007.
- [56] T. L. HILL, *Free Energy Transduction and Biochemical Cycle Kinetics*, SpringerVerlag New York, 1989.
- [57] R. ASTUMIAN and M. BIER, “Mechanochemical coupling of the motion of molecular motors to ATP hydrolysis”, *Biophys. J.* **70.2** (1996), 637–653.
- [58] R. D. ASTUMIAN, “Microscopic reversibility as the organizing principle of molecular machines”, *Nat. Nanotechnol.* **7.11** (2012), 684–688.
- [59] P. GLANSDORFF and I. PRIGOGINE, *Thermodynamic theory of structure, stability and fluctuations*, Wiley-Interscience, 1971.
- [60] C. VAN DEN BROECK, “Stochastic Thermodynamics”, *Selforganization by Nonlinear Irreversible Processes*, ed. by W. EBELING and H. ULBRICHT, Springer Berlin Heidelberg, 1986, pp. 57–61.
- [61] C. Y. MOU, J. LUO and G. NICOLIS, “Stochastic thermodynamics of nonequilibrium steady states in chemical reaction systems”, *J. Chem. Phys.* **84.12** (1986), 7011–7017.
- [62] L. JIU-LI, C. VAN DEN BROECK and G. NICOLIS, “Stability criteria and fluctuations around nonequilibrium states”, *Z. Physik B - Condensed Matter* **56.2** (1984), 165–170.
- [63] Q. ZHENG and J. ROSS, “Comparison of deterministic and stochastic kinetics for nonlinear systems”, *J. Chem. Phys.* **94.5** (1991), 3644–3648.
- [64] M. O. VLAD and J. ROSS, “Fluctuation–dissipation relations for chemical systems far from equilibrium”, *J. Chem. Phys.* **100.10** (1994), 7268–7278.
- [65] C. JARZYNSKI, “Equalities and Inequalities: Irreversibility and the Second Law of Thermodynamics at the Nanoscale”, *Annu. Rev. Condens. Matter Phys.* **2.1** (2011), 329–351.

- [66] U. SEIFERT, “Stochastic thermodynamics, fluctuation theorems and molecular machines”, *Rep. Prog. Phys.* **75.12** (2012), 126001.
- [67] C. V. den BROECK and M. ESPOSITO, “Ensemble and trajectory thermodynamics: A brief introduction”, *Phys. A: Stat. Mech. Appl.* **418**, (2015), 6–16.
- [68] S. CILIBERTO, “Experiments in Stochastic Thermodynamics: Short History and Perspectives”, *Phys. Rev. X* **7**, (2017), 021051.
- [69] L. PELITI and S. PIGOLOTTI, *Stochastic Thermodynamics: An Introduction*, Princeton University Press, Princeton, 2021.
- [70] H. QIAN and D. A. BEARD, “Thermodynamics of stoichiometric biochemical networks in living systems far from equilibrium”, *Biophys. Chem.* **114.2** (2005), 213–220.
- [71] T. SCHMIEDL and U. SEIFERT, “Stochastic thermodynamics of chemical reaction networks”, *J. Chem. Phys.* **126.4** (2007), 044101.
- [72] R. RAO and M. ESPOSITO, “Conservation laws and work fluctuation relations in chemical reaction networks”, *J. Chem. Phys.* **149.24** (2018), 245101.
- [73] G. BENENTI et al., “Fundamental aspects of steady-state conversion of heat to work at the nanoscale”, *Phys. Rep.* **694**, (2017), 1–124.
- [74] J. M. R. PARRONDO, J. M. HOROWITZ and T. SAGAWA, “Thermodynamics of information”, *Nat. Phys.* **11.2** (2015), 131–139.
- [75] R. RAO and M. ESPOSITO, “Conservation laws shape dissipation”, *New J. Phys.* **20.2** (2018), 023007.
- [76] M. POLETTINI, A. WACHTEL and M. ESPOSITO, “Dissipation in noisy chemical networks: The role of deficiency”, *J. Chem. Phys.* **143.18** (2015), 184103.
- [77] M. ESPOSITO, “Stochastic thermodynamics under coarse graining”, *Phys. Rev. E* **85**, (2012), 041125.
- [78] B. ANDRESEN, “Current Trends in Finite-Time Thermodynamics”, *Angew. Chem. Int. Ed.* **50.12** (2011), 2690–2704.
- [79] T. SCHMIEDL and U. SEIFERT, “Optimal Finite-Time Processes In Stochastic Thermodynamics”, *Phys. Rev. Lett.* **98**, (2007), 108301.
- [80] T. SCHMIEDL and U. SEIFERT, “Efficiency at maximum power: An analytically solvable model for stochastic heat engines”, *EPL (Europhysics Letters)* **81.2** (2007), 20003.
- [81] M. ESPOSITO, K. LINDENBERG and C. VAN DEN BROECK, “Universality of Efficiency at Maximum Power”, *Phys. Rev. Lett.* **102**, (2009), 130602.
- [82] M. ESPOSITO et al., “Efficiency at Maximum Power of Low-Dissipation Carnot Engines”, *Phys. Rev. Lett.* **105**, (2010), 150603.
- [83] F. JÜLICHER, A. AJDARI and J. PROST, “Modeling molecular motors”, *Rev. Mod. Phys.* **69**, (1997), 1269–1282.
- [84] H. QIAN, “The mathematical theory of molecular motor movement and chemomechanical energy transduction”, *J. Math. Chem.* **27.3** (2000), 219–234.
- [85] J. M. R. PARRONDO and B. J. de CISNEROS, “Energetics of Brownian motors: a review”, *Appl. Phys. A* **75.2** (2002), 179–191.
- [86] C. MAES and M. H. van WIEREN, “A Markov Model for Kinesin”, *J. Stat. Phys.* **112.1** (2003), 329–355.

- [87] D. ANDRIEUX and P. GASPARD, “Fluctuation theorems and the nonequilibrium thermodynamics of molecular motors”, *Phys. Rev. E* **74**. (2006), 011906.
- [88] R. LIPOWSKY and S. LIEPELT, “Chemomechanical Coupling of Molecular Motors: Thermodynamics, Network Representations, and Balance Conditions”, *J. Stat. Phys.* **130**.1 (2008), 39–67.
- [89] U. SEIFERT, “Stochastic thermodynamics of single enzymes and molecular motors”, *Eur. Phys. J. E* **34**.3 (2011), 26.
- [90] A. I. BROWN and D. A. SIVAK, “Allocating and Splitting Free Energy to Maximize Molecular Machine Flux”, *J. Phys. Chem. B* **122**.4 (2018), 1387–1393.
- [91] A. I. BROWN and D. A. SIVAK, “Theory of Nonequilibrium Free Energy Transduction by Molecular Machines”, *Chem. Rev.* **120**.1 (2020), 434–459.
- [92] S. J. LARGE, J. EHRLICH and D. A. SIVAK, “Free-energy transduction within autonomous systems”, *Phys. Rev. E* **103**. (2021), 022140.
- [93] M. P. LEIGHTON and D. A. SIVAK, “Performance scaling and trade-offs for collective motor-driven transport”, *New J. Phys.* **24**.1 (2022), 013009.
- [94] T. ARIGA, M. TOMISHIGE and D. MIZUNO, “Nonequilibrium Energetics of Molecular Motor Kinesin”, *Phys. Rev. Lett.* **121**. (2018), 218101.
- [95] T. SUMI and S. KLUMPP, “Is F₁-ATPase a Rotary Motor with Nearly 100% Efficiency? Quantitative Analysis of Chemomechanical Coupling and Mechanical Slip”, *Nano Letters* **19**.5 (2019), 3370–3378.
- [96] M. L. MUGNAI et al., “Theoretical perspectives on biological machines”, *Rev. Mod. Phys.* **92**. (2020), 025001.
- [97] T. M. COVER and J. A. THOMAS, *Elements of Information Theory*, John Wiley & Sons, Inc., Hoboken, 2012.
- [98] J. M. HOROWITZ and M. ESPOSITO, “Thermodynamics with Continuous Information Flow”, *Phys. Rev. X* **4**. (2014), 031015.
- [99] H. S. LEFF and A. F. REX, eds., *Maxwell’s Demon: Entropy, Information, Computing*, Princeton University Press, Princeton, 1990.
- [100] D. LOUTCHKO, M. EISBACH and A. S. MIKHAILOV, “Stochastic thermodynamics of a chemical nanomachine: The channeling enzyme tryptophan synthase”, *J. Chem. Phys.* **146**.2 (2017), 025101.
- [101] T. MCGRATH et al., “Biochemical Machines for the Interconversion of Mutual Information and Work”, *Phys. Rev. Lett.* **118**. (2017), 028101.
- [102] J. M. HOROWITZ, T. SAGAWA and J. M. R. PARRONDO, “Imitating Chemical Motors with Optimal Information Motors”, *Phys. Rev. Lett.* **111**. (2013), 010602.
- [103] R. TAKAKI, M. L. MUGNAI and D. THIRUMALAI, “Information flow, Gating, and Energetics in dimeric molecular motors”, *BioRxiv* **2021.12.30.474541**. (2021).
- [104] E. LATHOUWERS and D. A. SIVAK, “Internal energy and information flows mediate input and output power in bipartite molecular machines”, *Phys. Rev. E* **105**. (2022), 024136.

1.4 CONTRIBUTIONS OF THIS THESIS

In this group, you may not discover new things, but you learn not to say bullshit.

— Riccardo Rao

In the last Section, we presented the theoretical framework and experimental results which the original research presented in the coming chapters took inspiration from. We conclude this Chapter by contextualizing the major achievements of this thesis in the light of the above presentation.

In Chapter 2, we overcome one major limitation of the thermodynamics of chemical reaction networks as developed by the group before the start of this thesis, namely the ideal dilute solution assumption introduced in Section 1.3.1. This new nonideal framework extends the applicability to elementary reactions which do not obey mass action kinetics and allows to include mean-field effects. Examples of models which require such nonideal treatment are reacting mixtures of van der Waals gases and solutions of electrolytes following Debye-Hückel theory. Even if they are not specifically analyzed in this thesis, this nonideal framework could be particularly relevant for energetic considerations in electrochemical engines coupled to electrodes and biochemical processes involving ion gradients across membranes. Indeed, these are typical nonideal systems where species are affected by electric potentials in ways which are not taken into account by the ideal solution model. Chapter 2 also serves to systematically introduce the chemical reaction network approach used throughout and various notions of efficiency to evaluate the performance of a chemical engine. The latter are illustrated by using an ideal model of driven self-assembly as a case study, which is the first chemical engine that we will characterize. The methods developed in this chapter can in principle be applied to any open chemically-driven reaction network, encompassing both ideal and nonideal systems. We thus provide the basis for future performance studies and optimal design of fuel-driven chemical engines. This Chapter consists of two reprinted articles: Ref. [1] where the theory is developed, and Ref. [2] where the case study is analyzed.

In Chapter 3, we extend the theory to include incoherent light as a source of free energy. This includes in the framework photophysical and photochemical processes, which have been overlooked in the most recent advancements of chemical thermodynamics. In particular, such results allow for the thermodynamic analysis of light-driven chemical engines, which represent the majority of synthetic prototypes realized until now, as illustrated in Section 1.2. In the framework of a collaboration with the Credi group, their 2021 second-generation light-powered bimolecular pump discussed in Section 1.2 is analyzed as a case study in the light of new experimental data. This collaboration provided a test-bed for the theory, which turned out to answer new questions in the field such as how efficiently can energy be harvested by light-driven chemical engines, or how does the capability of performing work correlate with the light intensity. Results from this Chapter constitute a step forward towards a unified nonequilibrium thermodynamics for chemical systems powered by different free energy sources, such as chemostats and light. They also pave the way to performance studies on far-from-equilibrium free energy transduction and storage in finite time from different sources, which were inaccessible before. This Chapter consists of two reprinted articles: Ref. [3] where the theory is developed, and Ref. [4] where the case study is analyzed.

In Chapter 4, information thermodynamics is extended to deterministic *bipartite* chemical reaction networks. As we will argue, most of the existing chemical engines fall in this class and therefore can be studied within the new framework. In particular, the 2016 catenane-based chemically-driven molecular motor from the Leigh group and the 2015 light-powered bimolecular pump from the Credi group (both discussed in Section 1.2) are analyzed against experimental data. The model of driven self-assembly analyzed in Chapter 2 is also re-discussed in the light of the information thermodynamics framework, which allows for a refinement of how its efficiency can be evaluated. Beyond chemical engines, the results presented in Chapter 4 have theoretical relevance as they unlock information-thermodynamic analyses at the level of concentrations, while until now information in physical systems was only defined in terms of probability distributions. This Chapter consists of two reprinted articles: Ref. [5] where the theory is developed and two case studies are discussed, and Ref. [6] where the catenane-based molecular motor is analyzed in the framework of a collaboration with the Leigh group.

We warn that the notation in the papers may slightly change with respect to the one introduced in this Chapter. However, each paper is self-contained and variations to the standard notation are always specified in the text.

REFERENCES FOR CHAPTER 1.4

- [1] F. AVANZINI, E. PENOCCHIO, G. FALASCO and M. ESPOSITO, “Nonequilibrium thermodynamics of non-ideal chemical reaction networks”, *J. Chem. Phys.* **154**. (2021), 094114.
- [2] E. PENOCCHIO, R. RAO and M. ESPOSITO, “Thermodynamic efficiency in dissipative chemistry”, *Nat. Commun.* **10**. (2019), 3865.
- [3] E. PENOCCHIO, R. RAO and M. ESPOSITO, “Nonequilibrium thermodynamics of light-induced reactions”, *J. Chem. Phys.* **155**. (2021), 114101.
- [4] S. CORRA, M. T. BAKIĆ, J. GROPPI, M. BARONCINI, S. SILVI, E. PENOCCHIO, M. ESPOSITO and A. CREDI, “Kinetic and energetic insights into the dissipative non-equilibrium operation of an autonomous light-powered supramolecular pump”, *Nature Nanotechnology* (2022).
- [5] E. PENOCCHIO, F. AVANZINI and M. ESPOSITO, “Information Flows in Deterministic Chemical Reaction Networks”, *arXiv* **2204.02815**. (2022).
- [6] S. AMANO, M. ESPOSITO, E. KREIDT, D. A. LEIGH, E. PENOCCHIO and B. M. W. ROBERTS, “Insights from an information thermodynamics analysis of a synthetic molecular motor”, *Nat. Chem.* **14**. (2022), 530–537.

2

THE CHEMICAL REACTION NETWORK APPROACH

The following article is reprinted from
[F. AVANZINI, E. PENOCCHIO, G. FALASCO and M. ESPOSITO, *J. Chem. Phys.*
154, (2021), 094114]
under the conditions of AIP Publishing license¹.

The page numbers placed in the outer margins provide a continuous pagination throughout the thesis.

¹ <https://publishing.aip.org/wp-content/uploads/2019/10/AIPP-Author-License.pdf>

Nonequilibrium thermodynamics of non-ideal chemical reaction networks

Cite as: *J. Chem. Phys.* **154**, 094114 (2021); doi: 10.1063/5.0041225

Submitted: 21 December 2020 • Accepted: 15 February 2021 •

Published Online: 2 March 2021



View Online



Export Citation



CrossMark

Francesco Avanzini,^{a)} Emanuele Penocchio,^{b)} Gianmaria Falasco,^{c)} and Massimiliano Esposito^{d)}

AFFILIATIONS

Complex Systems and Statistical Mechanics, Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg City, Luxembourg

^{a)} Author to whom correspondence should be addressed: francesco.avanzini@uni.lu

^{b)} Electronic mail: emanuele.penocchio@uni.lu

^{c)} Electronic mail: gianmaria.falasco@uni.lu

^{d)} Electronic mail: massimiliano.esposito@uni.lu

ABSTRACT

All current formulations of nonequilibrium thermodynamics of open chemical reaction networks rely on the assumption of non-interacting species. We develop a general theory that accounts for interactions between chemical species within a mean-field approach using activity coefficients. Thermodynamic consistency requires that rate equations do not obey standard mass-action kinetics but account for the interactions with concentration dependent kinetic constants. Many features of the ideal formulations are recovered. Crucially, the thermodynamic potential and the forces driving non-ideal chemical systems out of equilibrium are identified. Our theory is general and holds for any mean-field expression of the interactions leading to lower bounded free energies.

Published under license by AIP Publishing. <https://doi.org/10.1063/5.0041225>

I. INTRODUCTION

Thermodynamics studies the interconversions of energy. It was originally formulated as an equilibrium theory in the 19th century. Phenomenological extensions close to equilibrium, in the so-called linear regime, were introduced in the first half of the 20th century.^{1,2} The earliest formulations beyond the linear regime were developed for stochastic chemical reactions in the second half of the 20th century.^{3–7} However, it is only in the 21st century that stochastic thermodynamics was systematized to characterize dissipative processes occurring arbitrarily far from equilibrium and their fluctuations.^{8–10} In recent years, building on the work of Prigogine and co-workers,¹¹ stochastic thermodynamics was formulated for stochastic^{12–14} and extended to deterministic^{15–17} chemical reaction networks (CRNs) maintained in a nonequilibrium regime with particle reservoirs called chemostats.

Nonequilibrium thermodynamics of CRNs is a powerful tool to analyze chemical complexity. It has been used for reaction diffusion systems to quantify the energetic cost of creating patterns,¹⁸ sustaining chemical waves,¹⁹ and powering a chemical cloaking device.²⁰

The growth process of macromolecules such as copolymers^{21–23} and biomolecules²⁴ has been characterized. In addition, applications to chaotic CRNs have been considered.²⁵ Recently, the connections of the theory to information geometry has been investigated.²⁶ Furthermore, different strategies to maintain nonequilibrium regimes, such as the use of finite chemostats²⁷ and serial transfers in closed reactors,²⁸ have been explored.

All current formulations of nonequilibrium thermodynamics of CRNs are, however, based on the assumption that the chemical species do not interact, except via chemical reactions. The purpose of this paper is to generalize nonequilibrium thermodynamics to non-ideal CRNs described by deterministic rate equations. Interactions are treated with a mean-field approach using the same activity coefficients introduced in equilibrium thermodynamics²⁹ but expressed in terms of nonequilibrium concentrations. Exploiting the local detailed balance assumption, we impose thermodynamic consistency on the dynamics. The standard mass-action kinetics has to be modified, introducing concentration dependent kinetic constants to account for the effects of interactions as already recognized in the literature.^{30–32}

Crucially, we generalize the decomposition of the entropy production developed in Ref. 14 to non-ideal CRNs. This allows us to determine the thermodynamic potential and the forces driving non-ideal chemical systems out of equilibrium. The expression of the forces is remarkably similar to the corresponding one for ideal CRNs since all the effects due to the interactions are hidden in the activity coefficients. The thermodynamic potential still acts as a Lyapunov function in detailed balanced systems and is always lower bounded by its equilibrium value. However, the difference between the nonequilibrium and the equilibrium value of the potential cannot be solely expressed in terms of relative entropy as for ideal CRNs. Furthermore, the exchange currents controlling the chemostatted species cannot be defined only dynamically as in ideal CRNs, but they must account for the interactions via the activity coefficients.

To develop our theory in a self-contained way, we proceed as follows: In Sec. II, we discuss the network theory of chemical reactions. After introducing the basic notation in Subsection II A, we examine the dynamics in Subsection II B. We then define conservation laws and cycles in Subsections II C and II D, respectively. In Subsection II E, we identify under which conditions the existence of equilibrium states is granted, namely, CRNs are detailed balanced. We develop our thermodynamic theory in Sec. III. After introducing the general setup in Subsection III A, we connect the dynamics to the thermodynamics using the local detailed balance assumption in Subsection III B. The expressions for the exchange currents controlling the chemostatted species are derived in Subsection III C. We then make use of conservation laws to decompose the entropy production rate in Subsection III D. At steady state, the entropy production has a second physically meaningful decomposition based on the cycles shown in Subsection III E. The properties of the thermodynamic potential are discussed in Subsection III F, where we derive its lower bound, and in Subsection III G, where we show that it cannot be written as its equilibrium value plus a relative entropy as in ideal systems.

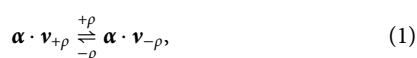
Throughout the manuscript, we use the CRN (9) to illustrate our results. This example represents a catalytic process power by a proton transfer in a system where the relevant interactions are the electrostatic interactions between charged species. We summarize our results and discuss their implications in Sec. IV.

This work should be particularly relevant for energetic considerations in biosystems, where many metabolic processes involve storing energy in ion gradients across membranes,³³ and electrochemical systems, where CRNs are coupled to electronic circuits.³⁴

II. NON-IDEAL CHEMICAL REACTION NETWORKS

A. Setup

We consider systems composed of interacting chemical species, identified by the label $\alpha \in Z$, homogeneously distributed in a constant volume V . Our description can treat gas phases and dilute solutions where the volume of the solution is overwhelmingly dominated by the solvent. The chemical species undergo elementary chemical reactions,³⁵ identified by the index $\rho \in \mathcal{R}$,



with $\alpha = (\dots, \alpha, \dots)^T$ being the vector of chemical species and $\mathbf{v}_{\pm\rho}$ being the vector of stoichiometric coefficients of the forward/backward reaction $\pm\rho$. The set of chemical reactions (1) defines the CRN. Our framework considers open CRNs. We split the set of all the species Z into two disjoint subsets: the internal species X and the chemostatted species Y . The former undergo only the chemical reactions (1). The latter too undergo the chemical reactions, but they are also externally exchanged. We will discuss the dynamical implications and the thermodynamic meaning of the chemostatting procedure in Subsections II B and III C, respectively.

B. Dynamics

The state of deterministic CRNs with constant volume is specified by the concentration vector $\mathbf{z}(t) = (\dots, [\alpha](t), \dots)^T$. Its dynamics follows the rate equation

$$d_t \mathbf{z}(t) = \mathbb{S} \mathbf{j}(\mathbf{z}(t)) + \mathbf{I}(t), \quad (2)$$

where we introduced the stoichiometric matrix \mathbb{S} , the current vector $\mathbf{j}(\mathbf{z})$, and the exchange current vector $\mathbf{I}(t)$. The first term on the rhs of Eq. (2), i.e., $\mathbb{S} \mathbf{j}(\mathbf{z}(t))$, accounts for the concentration changes due to the chemical reactions (1). The second term, i.e., $\mathbf{I}(t)$, accounts for the external matter flows.

The stoichiometric matrix \mathbb{S} codifies the topology of the CRN. Each ρ column \mathbb{S}_ρ specifies the net variation of the number of molecules for each species undergoing the ρ elementary reaction (1), $\mathbb{S}_\rho = \mathbf{v}_{-\rho} - \mathbf{v}_{+\rho}$. The current vector $\mathbf{j}(\mathbf{z}) = (\dots, j^\rho(\mathbf{z}), \dots)^T$ specifies the net reaction current for every ρ reaction as the difference between the forward $j^{+\rho}(\mathbf{z})$ and backward reaction flux $j^{-\rho}(\mathbf{z})$,

$$j^\rho(\mathbf{z}) = j^{+\rho}(\mathbf{z}) - j^{-\rho}(\mathbf{z}). \quad (3)$$

In the case of ideal CRNs, the fluxes $j^{\pm\rho}(\mathbf{z})$ are expressed in terms of mass-action kinetics,^{2,31,36}

$$j^{\pm\rho}(\mathbf{z}) = k_{\pm\rho}^{\text{id}} \mathbf{z}^{\mathbf{v}_{\pm\rho}}, \quad (4)$$

where $k_{\pm\rho}^{\text{id}}$ are the kinetic constants of the forward/backward reaction $\pm\rho$ and we used the following notation: $\mathbf{a}^{\mathbf{b}} = \prod_i a_i^{b_i}$. In the case of non-ideal CRNs, mass-action kinetics is not thermodynamically consistent as shown in Subsection III B. To take into account the effects of the interactions, we assume that the kinetic constants may depend on the concentrations $\{k_{\pm\rho}(\mathbf{z})\}$ as widely acknowledged in the literature.³⁰⁻³² Thus, the reaction fluxes are given by the general expressions

$$j^{\pm\rho}(\mathbf{z}) = k_{\pm\rho}(\mathbf{z}) \mathbf{z}^{\mathbf{v}_{\pm\rho}}. \quad (5)$$

The specific \mathbf{z} -dependence of $\{k_{\pm\rho}(\mathbf{z})\}$ derives from the particular model used to describe the interactions. We develop our theory only assuming that $k_{\pm\rho}(\mathbf{z})$ satisfy the conditions introduced in Subsections II E and III B to be thermodynamically consistent.

The exchange current vector $\mathbf{I}(t)$ specifies the external matter flows. It has null entries for the internal species, i.e., $I^\alpha(t) = 0$ for $\alpha \in X$: the concentration of the internal species changes only because of the chemical reactions (1) by definition. The entries of $\mathbf{I}(t)$

for the chemostatted species, i.e., $I^\alpha(t)$ for $\alpha \in Y$, are derived in Subsection III C because the thermodynamic meaning of the chemostating procedure must be taken into account.

We can apply the splitting of the chemical species into internal and chemostatted ones to the stoichiometric matrix,

$$\mathbb{S} = \begin{pmatrix} \mathbb{S}^X \\ \mathbb{S}^Y \end{pmatrix}, \quad (6)$$

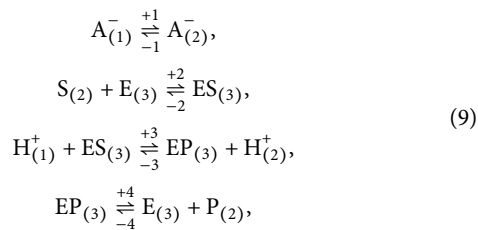
and the concentration vector $\mathbf{z} = (\mathbf{x}, \mathbf{y})$. Analogously, the rate equation (2) becomes

$$d_t \mathbf{x}(t) = \mathbb{S}^X \mathbf{j}(\mathbf{x}(t), \mathbf{y}(t)), \quad (7)$$

$$d_t \mathbf{y}(t) = \mathbb{S}^Y \mathbf{j}(\mathbf{x}(t), \mathbf{y}(t)) + \mathbf{I}^Y(t), \quad (8)$$

with $\mathbf{I}^Y(t) = (\dots, I^\alpha(t), \dots)^\top_{\alpha \in Y}$ collecting the not null entries of $\mathbf{I}(t)$. Note that Eqs. (7) and (8) are only a reformulation of Eq. (2).

Example 1. We consider the following CRN (see also Fig. 1):



where the superscripts indicate the electric charge and the subscripts indicate the compartment where a chemical species is located. Here, $\alpha \in Z = \{E_{(3)}, ES_{(3)}, EP_{(3)}, A_{(1)}^-, A_{(2)}^-, H_{(1)}^+, H_{(2)}^+, S_{(2)}, P_{(2)}\}$.

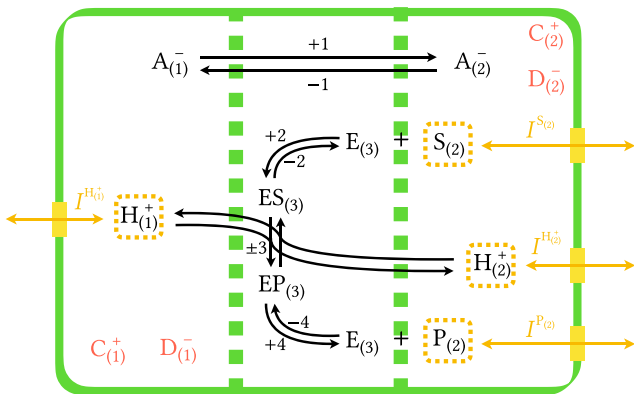


FIG. 1. Drawing of the CRN (9). A membrane (green dashed line) divides the system into compartments. A membrane enzyme E binds the substrate S. The interconversion of the substrate S into the product P is coupled to the proton H^+ transfer from compartment (1) to compartment (2). In parallel, the anions A^- move from compartment (1) to compartment (2). S, P, and H^+ are chemostatted (yellow arrows). Other cations C^+ and anions D^- are present in compartment (1) and (2).

The CRN (9) represents the transformation of the substrate S into the product P inside compartment (2) catalyzed by the membrane enzyme E. The interconversion of the complex ES into EP on the membrane [compartment (3)] is coupled to the transfer of protons H^+ from compartment (1) to compartment (2). Independently, anions A^- are free to move from compartment (1) to compartment (2). Other cations C^+ and anions D^- are present in compartment (1) and (2), but they are not involved in any chemical reaction. For this reason, the species C^+ and D^- are not included in Z. Because of the coupling between the catalytic reaction and the proton transfer, the CRN (9) resembles the working mechanism of the membrane enzyme ATP synthase.³⁷ All the protons, the substrate, and the product are chemostatted, i.e., $Y = \{H_{(1)}^+, H_{(2)}^+, S_{(2)}, P_{(2)}\}$ and $X = \{E_{(3)}, ES_{(3)}, EP_{(3)}, A_{(1)}^-, A_{(2)}^-\}$. For this kind of system, the strongest interactions are the electrostatic interactions between charged species inside the same compartment.

The stoichiometric matrix of the CRN (9) is specified as

$$\mathbb{S} = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} E_{(3)} \\ ES_{(3)} \\ EP_{(3)} \\ A_{(1)}^- \\ A_{(2)}^- \\ H_{(1)}^+ \\ H_{(2)}^+ \\ S_{(2)} \\ P_{(2)} \end{matrix} & \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}. \quad (10)$$

□

C. Conservation laws

The linearly independent vectors $\{\ell^\lambda\}$ in the cokernel of the stoichiometric matrix

$$\ell^\lambda \cdot \mathbb{S} = 0 \quad (11)$$

are the so-called conservation laws.^{16,17} All the scalars $L^\lambda(\mathbf{z}) \equiv \ell^\lambda \cdot \mathbf{z}$ would be conserved quantities if the CRN was closed, namely, $\mathbf{I}(t) = 0$. Indeed, $d_t L^\lambda(\mathbf{z}(t)) = \ell^\lambda \cdot \mathbb{S} \mathbf{j}(\mathbf{z}(t)) = 0$. In closed CRNs, the total mass is conserved, and consequently, the set $\{\ell^\lambda\}$ is never empty.

When CRNs are open, some conservation laws do not correspond anymore to conserved quantities. Hence, we split the set of conservation laws $\{\ell^\lambda\}$ into two disjoint subsets: the unbroken conservation laws $\{\ell^{\lambda_u}\}$ and the broken conservation laws $\{\ell^{\lambda_b}\}$. The unbroken conservation laws are the largest subset of conservation laws that can be written with null entries for the chemostatted species, i.e., $\ell_{\alpha}^{\lambda_u} = 0$ for $\alpha \in Y$. Their corresponding scalar quantities $L^{\lambda_u}(\mathbf{z}) = \ell^{\lambda_u} \cdot \mathbf{z}$ are still conserved,

$$d_t L^{\lambda_u}(\mathbf{z}(t)) = \ell^{\lambda_u} \cdot \mathbb{S} \mathbf{j}(\mathbf{z}(t)) + \sum_{\alpha \in Y} \ell_{\alpha}^{\lambda_u} I^\alpha(t) = 0. \quad (12)$$

The broken conservation laws are the other conservation laws $\{\ell^{\lambda_b}\} = \{\ell^\lambda\} \setminus \{\ell^{\lambda_u}\}$. Their corresponding scalar quantities $L^{\lambda_b}(\mathbf{z}) = \ell^{\lambda_b} \cdot \mathbf{z}$ are, in general, not conserved and they give rise to balance equations,

$$d_t L^{\lambda_b}(\mathbf{z}(t)) = \mathcal{L}^{\lambda_b} \cdot \mathbb{S} \mathbf{j}(\mathbf{z}(t)) + \sum_{\alpha \in Y} \ell_{\alpha}^{\lambda_b} I^{\alpha}(t) \neq 0, \quad (13)$$

where the variations of $L^{\lambda_b}(\mathbf{z}(t))$ are due to the matter exchanged through the currents $\{I^{\alpha}(t)\}_{\alpha \in Y}$. In open CRNs, the total mass is not conserved and, consequently, the set $\{\mathcal{L}^{\lambda_b}\}$ is never empty.

Chemostating a species does not always break a conservation law.^{14,17,18,38} We thus distinguish the species $Y_p \subseteq Y$ that break the conservation laws from the others $Y_f = Y - Y_p$. This classification is not unique and different choices have different physical meanings. It should be motivated by the physical role played by each chemostated species, as it will become clear in Subsection III D. We apply the same splitting to the stoichiometric matrix \mathbb{S}^Y and the matrix \mathbb{L}^b whose rows are the broken conservation laws $\{\mathcal{L}^{\lambda_b}\}$,

$$\mathbb{S}^Y = \begin{pmatrix} \mathbb{S}^{Y_f} \\ \mathbb{S}^{Y_p} \end{pmatrix}, \quad \mathbb{L}^b = (\mathbb{L}_X^b, \mathbb{L}_{Y_f}^b, \mathbb{L}_{Y_p}^b), \quad (14)$$

where \mathbb{L}_X^b , $\mathbb{L}_{Y_f}^b$, and $\mathbb{L}_{Y_p}^b$ have $\{\ell_{\alpha}^{\lambda_b}\}_{\alpha \in X}$, $\{\ell_{\alpha}^{\lambda_b}\}_{\alpha \in Y_f}$, and $\{\ell_{\alpha}^{\lambda_b}\}_{\alpha \in Y_p}$ as entries, respectively. Note that the number of Y_p species is equal to the number of broken conservation laws by definition. The matrix $\mathbb{L}_{Y_p}^b$ is then square and nonsingular, and so, it can be inverted. We will exploit these properties in Subsection III D where we decompose the entropy production rate.

Example 2. Given the stoichiometric matrix (10) of the CRN (9), there are five conservation laws,

$$\mathcal{L}^E = \begin{pmatrix} E_{(3)} \\ ES_{(3)} \\ EP_{(3)} \\ A_{(1)} \\ A_{(2)} \\ H_{(1)} \\ H_{(2)} \\ S_{(2)} \\ P_{(2)} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathcal{L}^A = \begin{pmatrix} E_{(3)} \\ ES_{(3)} \\ EP_{(3)} \\ A_{(1)} \\ A_{(2)} \\ H_{(1)} \\ H_{(2)} \\ S_{(2)} \\ P_{(2)} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathcal{L}^H = \begin{pmatrix} E_{(3)} \\ ES_{(3)} \\ EP_{(3)} \\ A_{(1)} \\ A_{(2)} \\ H_{(1)} \\ H_{(2)} \\ S_{(2)} \\ P_{(2)} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad (15)$$

$$\mathcal{L}^{HS} = \begin{pmatrix} E_{(3)} \\ ES_{(3)} \\ EP_{(3)} \\ A_{(1)} \\ A_{(2)} \\ H_{(1)} \\ H_{(2)} \\ S_{(2)} \\ P_{(2)} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathcal{L}^S = \begin{pmatrix} E_{(3)} \\ ES_{(3)} \\ EP_{(3)} \\ A_{(1)} \\ A_{(2)} \\ H_{(1)} \\ H_{(2)} \\ S_{(2)} \\ P_{(2)} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Four of them have a clear physical meaning: the total concentration of the enzyme, anions, protons, and substrate is given by $L^E = \mathcal{L}^E \cdot \mathbf{z} = [E_{(3)}] + [ES_{(3)}] + [EP_{(3)}]$, $L^A = \mathcal{L}^A \cdot \mathbf{z} = [A_{(1)}] + [A_{(2)}]$, $L^H = \mathcal{L}^H \cdot \mathbf{z} = [H_{(1)}] + [H_{(2)}]$, and $L^S = \mathcal{L}^S \cdot \mathbf{z} = [ES_{(3)}] + [EP_{(3)}] + [S_{(2)}] + [P_{(2)}]$, respectively. The conserved quantity $L^{HS} = \mathcal{L}^{HS} \cdot \mathbf{z} = [ES_{(3)}] + [S_{(2)}] + [H_{(2)}]$ comes from the coupling between the consumption of the substrate and the proton transfer.

When the species $H_{(1)}^+$, $H_{(2)}^+$, $S_{(2)}$, and $P_{(2)}$ are chemostated, the three conservation laws \mathcal{L}^H , \mathcal{L}^{HS} , and \mathcal{L}^S are broken. We identify the set Y_p as $Y_p = \{H_{(2)}^+, S_{(2)}, P_{(2)}\}$, but different choices are possible (see the discussion in Appendix A of Ref. 19). Thus, $Y_f = Y - Y_p = \{H_{(1)}^+\}$. The matrix \mathbb{L}^b whose rows are the broken conservation laws is given by

$$\mathbb{L}^b = \begin{matrix} \mathcal{L}^H \\ \mathcal{L}^{HS} \\ \mathcal{L}^S \end{matrix} \begin{pmatrix} E_{(3)} & ES_{(3)} & EP_{(3)} & A_{(1)} & A_{(2)} & H_{(1)} & H_{(2)} & S_{(2)} & P_{(2)} \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \quad (16)$$

where the gray vertical lines mark the split of \mathbb{L}^b into \mathbb{L}_X^b , $\mathbb{L}_{Y_f}^b$, and $\mathbb{L}_{Y_p}^b$. □

D. Cycles

The linearly independent vectors $\{c_i\}$ in the kernel of the stoichiometric matrix

$$\mathbb{S} c_i = 0 \quad (17)$$

are the so-called internal cycles.^{16,17} They represent sequences of reactions that upon completion leave the all concentrations unchanged. Any linear combination of internal cycles gives a steady-state current vector of closed CRNs, $d_t \mathbf{z}(t) = \mathbb{S} \bar{\mathbf{j}} = 0$, with $\bar{\mathbf{j}} = \sum_i c_i \psi^i$.

When CRNs are open, the steady-state current vector must (i) leave the concentrations of the internal species unchanged, i.e., $\mathbb{S}^X \bar{\mathbf{j}} = 0$, and (ii) be balanced by the exchanged current vector, i.e., $\mathbb{S}^Y \bar{\mathbf{j}} + \mathbf{I}^Y = 0$. Thus, $\bar{\mathbf{j}} \in \ker(\mathbb{S}^X)$ and, consequently, can be written as a linear combination of the right-null eigenvectors,

$$\mathbb{S}^X c_{\xi} = 0. \quad (18)$$

The set $\{c_{\xi}\}$ includes the internal cycles $\{c_i\}$ and, in general, other vectors called emergent cycles $\{c_e\}$. The steady-state current vector can therefore be written as

$$\bar{\mathbf{j}} = \sum_i c_i \psi^i + \sum_e c_e \psi^e. \quad (19)$$

Because of the rank-nullity theorem for the stoichiometric matrix \mathbb{S} and \mathbb{S}^X , every time a species is chemostated, either a conservation law is broken or an emergent cycle arises, namely,

$$|Y| = |\lambda_b| + |\epsilon|, \quad (20)$$

with $|Y|$, $|\lambda_b|$, and $|\epsilon|$ being the number of chemostated species, broken conservation laws, and emergent cycles, respectively. This implies that the number of Y_f species is equal to the number of emergent cycles, i.e., $|Y_f| = |\epsilon|$.

Example 3. Given the stoichiometric matrix (10) of the CRN (9), there is only one emergent cycle and no internal cycles,

$$c = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}. \quad (21)$$

□

E. Equilibrium

The equilibrium steady-state \mathbf{z}_{eq} of the rate equation (2), if it exists, is characterized by vanishing reaction currents,

$$\mathbf{j}(\mathbf{z}_{\text{eq}}) = 0. \quad (22)$$

This, together with Eq. (5), implies

$$\frac{k_{+\rho}(\mathbf{z}_{\text{eq}})}{k_{-\rho}(\mathbf{z}_{\text{eq}})} = \mathbf{z}_{\text{eq}}^{\mathbb{S}_\rho}. \quad (23)$$

If a solution \mathbf{z}_{eq} of Eq. (22) [or equivalently of Eq. (23)] exists, the CRN is said to be *detailed balanced*. Closed CRNs must be detailed balanced for thermodynamic consistency. It is not granted that open CRNs are detailed balanced. It depends on the chemostatting procedure.

Here, we show that open CRNs are necessarily detailed balanced if the number of chemostatted species is equal to the number of broken conservation laws, i.e., $Y = Y_p$, or, equivalently, there are no emergent cycles. In Subsection III D, we show that equilibrium states can exist also when $Y \neq Y_p$ and there are emergent cycles. For thermodynamic consistency, namely, the local detailed balance condition that will be introduced in Eq. (40), the product of the forward rate constants along every internal cycle (17) must equal that of the backward rate constants,

$$\prod_{\rho} \left(\frac{k_{+\rho}(\mathbf{z})}{k_{-\rho}(\mathbf{z})} \right)^{c_{\rho}^{\iota}} = 1, \quad (24)$$

where c_{ρ}^{ι} is the ρ entry of the ι internal cycle. This is known as Wegscheider's condition in the framework of ideal CRNs.³⁹ Hence, only $|\rho| - |\iota|$ ratios $k_{+\rho}(\mathbf{z})/k_{-\rho}(\mathbf{z})$ are independent (with $|\rho|$ being the number of reactions and $|\iota|$ being the number of internal cycles). This means that Eq. (23) imposes $|\rho| - |\iota|$ constraints. This is independent of whether CRNs are closed or open.

In closed CRNs, only $|\alpha| - |\lambda|$ of the equilibrium concentrations are independent (with $|\alpha|$ being the number of all the species in Z and $|\lambda|$ being the number of conservation laws). The others are fixed by the initial conditions through the conserved quantities,

$$L^{\lambda} = \mathcal{L}^{\lambda} \cdot \mathbf{z}(0) = \mathcal{L}^{\lambda} \cdot \mathbf{z}_{\text{eq}}. \quad (25)$$

Therefore, Eq. (23) imposes $|\rho| - |\iota|$ constraints for $|\alpha| - |\lambda|$ independent variables. Because of the rank-nullity theorem for the stoichiometric matrix \mathbb{S} ,

$$|\rho| - |\iota| = |\alpha| - |\lambda|, \quad (26)$$

the existence of a solution to Eq. (23) is granted. Closed CRNs must be detailed balanced, and hence, each solution of Eq. (23) must be physically meaningful, i.e., $[\alpha]_{\text{eq}} \in \mathbb{R}_{\geq 0} \forall \alpha \in Z$. This imposes constraints on the possible expressions of $\{k_{\pm\rho}(\mathbf{z})\}$.

In open CRNs, $|\lambda| - |\lambda_b|$ concentrations are fixed by the initial conditions through the unbroken conserved quantities,

$$L^{\lambda_u} = \mathcal{L}^{\lambda_u} \cdot \mathbf{z}(0) = \mathcal{L}^{\lambda_u} \cdot \mathbf{z}_{\text{eq}}. \quad (27)$$

Other $|Y|$ concentrations are constrained by the chemostatting procedure. The final number of independent concentrations is given by

$$|\alpha| - (|\lambda| - |\lambda_b|) - |Y| = |\alpha| - |\lambda| - |Y_f| = |\rho| - |\iota| - |Y_f|, \quad (28)$$

where we used $|Y| - |\lambda_b| = |Y_f|$ and Eq. (26). Therefore, Eq. (23) imposes $|\rho| - |\iota|$ constraints for $|\rho| - |\iota| - |Y_f|$ independent variables. The existence of a solution is always granted if and only if $|Y_f| = 0$, namely, all the chemostatted species break a conservation law (i.e., $Y = Y_p$). Equivalently, the existence of a solution is always granted if and only if there are no emergent cycles, $|e| = 0$. The solution is physically meaningful because it must also be a solution of the corresponding closed CRN.

These are sufficient conditions for the existence of an equilibrium state \mathbf{z}_{eq} . For ideal detailed balanced CRNs, i.e., evolving according to mass-action kinetics (4), there is a unique equilibrium state for every stoichiometric compatibility class $\Omega(\{L^{\lambda_u}\})$, namely, for every manifold in the concentration space characterized by specific and unique values of all the conserved quantities.^{39,40} For non-ideal CRNs, this property is not granted anymore, and there might be more than one equilibrium state for every stoichiometric compatibility class. The existence of multiple equilibria in the same stoichiometric compatibility class will become relevant in Subsection III F.

Example 4. It is not granted that the open CRN (9) is detailed balanced because the set Y_f is not empty, i.e., $Y_f = \{H_{(1)}^{\dagger}\}$, and there is the emergent cycle (21). □

III. THERMODYNAMICS

A. Setup

Nonequilibrium thermodynamics of CRNs presumes that all degrees of freedom other than concentrations are equilibrated. The temperature T is set by a thermal reservoir (e.g., the solvent in dilute solutions), diffusion processes are fast enough to keep the chemical species homogeneously distributed, and the fields responsible for the interactions, e.g., electrostatic fields, relax instantaneously to their mean-field values. For dilute solutions, the pressure p is set by the environment of the solution. In the absence of reactions, the CRN would be an equilibrated mixture. In this way, thermodynamic state functions can be specified by their equilibrium form but expressed in terms of nonequilibrium concentrations.

We assume that the free energy of non-ideal CRNs, for a given temperature T and volume V , is a mean-field function $G(\mathbf{z}, \mathbf{e})$ of

the concentrations \mathbf{z} and some externally controlled parameters $\mathbf{e} = (\dots, e_i, \dots)^\top$ that tune the interactions. We chose the symbol G for the free energy because it corresponds to the Gibbs free energy when dealing with dilute solutions. If we considered gas phases in a constant volume, G would be the Helmholtz free energy. For instance, Debye–Hückel theory⁴¹ derives $G(\mathbf{z}, \mathbf{e})$ for electrolyte solutions, and van der Waals theory¹¹ provides $G(\mathbf{z}, \mathbf{e})$ for a gas of interacting and finite-size molecules. Standard computational approaches are also available to compute the free energy of generic non-ideal systems.^{42,43} Without loss of generality, $G(\mathbf{z}, \mathbf{e})$ can be written as the sum of an ideal term $G^{\text{id}}(\mathbf{z})$ and a term accounting for the effects of the interactions $G^{\text{in}}(\mathbf{z}, \mathbf{e})$,

$$G(\mathbf{z}, \mathbf{e}) = G^{\text{id}}(\mathbf{z}) + G^{\text{in}}(\mathbf{z}, \mathbf{e}). \quad (29)$$

The free energy contribution carried by each species α is given by the chemical potential,⁴⁴

$$\mu_\alpha(\mathbf{z}) = \frac{\partial G(\mathbf{z}, \mathbf{e})}{\partial [\alpha]} = \mu_\alpha^{\text{id}}([\alpha]) + RT \ln \gamma_\alpha(\mathbf{z}, \mathbf{e}). \quad (30)$$

By assumption, $\mu_\alpha^{\text{id}}([\alpha])$ is the ideal chemical potential in either gas phases at constant volume V or in dilute solutions,

$$\mu_\alpha^{\text{id}}([\alpha]) = \frac{\partial G^{\text{id}}(\mathbf{z})}{\partial [\alpha]} = \mu_\alpha^\circ + RT \ln [\alpha], \quad (31)$$

with R being the gas constant and μ_α° being the standard chemical potential. The so-called activity coefficient $\gamma_\alpha(\mathbf{z}, \mathbf{e})$ accounts for the effects of the interactions,

$$RT \ln \gamma_\alpha(\mathbf{z}, \mathbf{e}) = \frac{\partial G^{\text{in}}(\mathbf{z}, \mathbf{e})}{\partial [\alpha]}. \quad (32)$$

In equilibrium thermodynamics of non-ideal CRNs, $\gamma_\alpha(\mathbf{z}, \mathbf{e})$ is often treated as a constant parameter, while in our framework, it evolves in time since the concentrations \mathbf{z} are dynamical variables. The specific expression of $\gamma_\alpha(\mathbf{z}, \mathbf{e})$, or equivalently of $G^{\text{in}}(\mathbf{z}, \mathbf{e})$, depends on the particular model used to describe the interactions. We develop our theory only assuming that $G(\mathbf{z}, \mathbf{e})$ is lower bounded and increases superlinearly as concentrations go to infinity. These constraints are necessary to guarantee thermodynamic consistency, i.e., detailed balanced CRNs relax toward an equilibrium state, as shown in Subsection III F. The free energy provided by the Debye–Hückel theory⁴¹ satisfies these conditions: as concentrations become infinitely large, $G^{\text{in}}(\mathbf{z}, \mathbf{e})$ goes to minus infinity slower than $G^{\text{id}}(\mathbf{z})$ goes to infinity.

We collect all the chemical potentials in the vector

$$\boldsymbol{\mu}(\mathbf{z}) = (\dots, \mu_\alpha(\mathbf{z}), \dots)^\top, \quad (33)$$

which can be written as

$$\begin{aligned} \boldsymbol{\mu}(\mathbf{z}) &= \boldsymbol{\mu}^\circ + RT \ln \mathbf{z} + RT \ln \boldsymbol{\gamma}(\mathbf{z}, \mathbf{e}) \\ &= \boldsymbol{\mu}^\circ + RT \ln \mathbf{z} + \nabla_{\mathbf{z}} G^{\text{in}}(\mathbf{z}, \mathbf{e}), \end{aligned} \quad (34)$$

with $\boldsymbol{\mu}^\circ = (\dots, \mu_\alpha^\circ, \dots)^\top$, $\boldsymbol{\gamma}(\mathbf{z}, \mathbf{e}) = (\dots, \gamma_\alpha(\mathbf{z}, \mathbf{e}), \dots)^\top$, and $\nabla_{\mathbf{z}} = (\dots, \partial/\partial[\alpha], \dots)^\top$.

Example 5. We assume that the only relevant interactions in the CRN (9) are the electrostatic interactions between charged species inside the same compartment. For illustrative reasons, we describe the electrostatic potential in each compartment as a linear function of the net charge density instead of using the Debye–Hückel theory.⁴¹ This is consistent with a standard model of the transmembrane potential in mitochondria.^{45–47} The interaction free energy $G^{\text{in}}(\mathbf{z}, \mathbf{e})$ is thus given by the electrostatic potential energy. It can be written as

$$G^{\text{in}}(\mathbf{z}, \mathbf{e}) = G_{(1)}^{\text{in}}(\mathbf{z}, \mathbf{e}) + G_{(2)}^{\text{in}}(\mathbf{z}, \mathbf{e}), \quad (35)$$

where $G_{(i)}^{\text{in}}(\mathbf{z}, \mathbf{e})$, i.e., the electrostatic potential energy in the i compartment, is a quadratic function of the net charge density in the i compartment,

$$G_{(i)}^{\text{in}}(\mathbf{z}, \mathbf{e}) = \frac{F\kappa}{2} \left(\begin{array}{c} -[A_{(i)}^-] + [H_{(i)}^+] + [C_{(i)}^+] - [D_{(i)}^-] \\ \equiv Q_{(i)}(\mathbf{z}, \mathbf{e}) \end{array} \right)^2, \quad (36)$$

with F being the Faraday constant and κ being a generic proportionality constant. On the membrane [compartment (3)], $G_{(3)}^{\text{in}} = 0$ since the chemical species are not charged and do not interact. Here, the parameters \mathbf{e} are the concentrations of the species that are not involved in the chemical reactions, i.e., $\mathbf{e} = ([C_{(1)}^+], [C_{(2)}^+], [D_{(1)}^-], [D_{(2)}^-])$, but interact with species α . These concentrations may be controlled with some external processes.

The chemical potentials of the species in the CRN (9) are specified as

$$\begin{aligned} \mu_{E_{(3)}}(\mathbf{z}) &= \mu_{E_{(3)}}^{\text{id}}([E_{(3)}]) = \mu_E^\circ + RT \ln [E_{(3)}], \\ \mu_{ES_{(3)}}(\mathbf{z}) &= \mu_{ES_{(3)}}^{\text{id}}([ES_{(3)}]) = \mu_{ES}^\circ + RT \ln [ES_{(3)}], \\ \mu_{EP_{(3)}}(\mathbf{z}) &= \mu_{EP_{(3)}}^{\text{id}}([EP_{(3)}]) = \mu_{EP}^\circ + RT \ln [EP_{(3)}], \\ \mu_{A_{(1)}^-}(\mathbf{z}) &= \mu_{A^-}^\circ + RT \ln [A_{(1)}^-] - F\kappa Q_{(1)}(\mathbf{z}, \mathbf{e}), \\ \mu_{A_{(2)}^-}(\mathbf{z}) &= \mu_{A^-}^\circ + RT \ln [A_{(2)}^-] - F\kappa Q_{(2)}(\mathbf{z}, \mathbf{e}), \\ \mu_{H_{(1)}^+}(\mathbf{z}) &= \mu_{H^+}^\circ + RT \ln [H_{(1)}^+] + F\kappa Q_{(1)}(\mathbf{z}, \mathbf{e}), \\ \mu_{H_{(2)}^+}(\mathbf{z}) &= \mu_{H^+}^\circ + RT \ln [H_{(2)}^+] + F\kappa Q_{(2)}(\mathbf{z}, \mathbf{e}), \\ \mu_{S_{(2)}}(\mathbf{z}) &= \mu_{S_{(2)}}^{\text{id}}([S_{(2)}]) = \mu_S^\circ + RT \ln [S_{(2)}], \\ \mu_{P_{(2)}}(\mathbf{z}) &= \mu_{P_{(2)}}^{\text{id}}([P_{(2)}]) = \mu_P^\circ + RT \ln [P_{(2)}], \end{aligned} \quad (37)$$

where we assumed that the standard chemical potentials do not depend on the compartment, and the net charge densities $\{Q_{(i)}(\mathbf{z}, \mathbf{e})\}$ are specified in Eq. (36). Note that the chemical potentials of the species $E_{(3)}$, $ES_{(3)}$, $EP_{(3)}$, $S_{(2)}$, and $P_{(2)}$ are the same as for ideal CRNs. This is a consequence of accounting only for the interactions between charge species. \square

B. Local detailed balance

Here, we build the connection between dynamics and nonequilibrium thermodynamics. We use the *local detailed balance* condition that binds the ratios between the forward and backward currents $\{j^{+\rho}(\mathbf{z})/j^{-\rho}(\mathbf{z})\}$ to the free energies of reaction $\{\Delta_\rho G(\mathbf{z})\}$,

$$RT \ln \frac{j^{+\rho}(\mathbf{z})}{j^{-\rho}(\mathbf{z})} = -\Delta_\rho G(\mathbf{z}). \quad (38)$$

The latter is given by

$$\Delta_\rho G(\mathbf{z}) = \boldsymbol{\mu}(\mathbf{z}) \cdot \mathbb{S}_\rho. \quad (39)$$

This, together with Eq. (5), implies a constraint for the kinetic constants $k_{\pm\rho}(\mathbf{z})$ of the chemical reactions,

$$\begin{aligned} RT \ln \frac{k_{+\rho}(\mathbf{z})}{k_{-\rho}(\mathbf{z})} &= -(\boldsymbol{\mu}^\circ + RT \ln \boldsymbol{\gamma}(\mathbf{z}, \mathbf{e})) \cdot \mathbb{S}_\rho \\ &= -(\boldsymbol{\mu}^\circ + \nabla_z G^{\text{in}}(\mathbf{z}, \mathbf{e})) \cdot \mathbb{S}_\rho. \end{aligned} \quad (40)$$

Equation (40) shows that the \mathbf{z} dependence of the kinetic constants, which represents the breakdown of mass-action kinetics for non-ideal CRNs, comes from the interactions. These can affect the relative stability of reagents and products and, consequently, the relative rate of the forward and backward reactions. However, the local detailed balance condition highlights only the \mathbf{z} dependence of the antisymmetric part $\sqrt{k_{+\rho}(\mathbf{z})/k_{-\rho}(\mathbf{z})}$ of the kinetic constants. This does not exclude that also the symmetric part $\sqrt{k_{+\rho}(\mathbf{z})k_{-\rho}(\mathbf{z})}$ of the kinetic constants is \mathbf{z} dependent as already discussed in the literature.³² In the limit $\boldsymbol{\gamma}(\mathbf{z}, \mathbf{e}) \rightarrow 1$, or equivalently $\nabla_z G^{\text{in}}(\mathbf{z}, \mathbf{e}) \rightarrow 0$, the local detailed balance condition for ideal CRNs is recovered,

$$RT \ln \frac{k_{+\rho}^{\text{id}}}{k_{-\rho}^{\text{id}}} = -\boldsymbol{\mu}^\circ \cdot \mathbb{S}_\rho, \quad (41)$$

and consequently, the dynamics satisfies mass-action kinetics (4).

Example 6. We specify Eq. (40) for the CRN (9) assuming that the standard chemical potentials of A^- and H^+ do not depend on the compartment [namely, $\mu_{A(1)}^\circ = \mu_{A(2)}^\circ$ and $\mu_{H(1)}^\circ = \mu_{H(2)}^\circ$],

$$RT \ln \frac{k_{+1}(\mathbf{z})}{k_{-1}(\mathbf{z})} = F\kappa(Q_{(2)}(\mathbf{z}, \mathbf{e}) - Q_{(1)}(\mathbf{z}, \mathbf{e})), \quad (42)$$

$$RT \ln \frac{k_{+2}(\mathbf{z})}{k_{-2}(\mathbf{z})} = RT \ln \frac{k_{+2}^{\text{id}}}{k_{-2}^{\text{id}}} = -(\mu_{\text{ES}(3)}^\circ - \mu_{\text{E}(3)}^\circ - \mu_{\text{S}(2)}^\circ), \quad (43)$$

$$RT \ln \frac{k_{+3}(\mathbf{z})}{k_{-3}(\mathbf{z})} = -(\mu_{\text{EP}(3)}^\circ - \mu_{\text{ES}(3)}^\circ + F\kappa(Q_{(2)}(\mathbf{z}, \mathbf{e}) - Q_{(1)}(\mathbf{z}, \mathbf{e}))), \quad (44)$$

$$RT \ln \frac{k_{+4}(\mathbf{z})}{k_{-4}(\mathbf{z})} = RT \ln \frac{k_{+4}^{\text{id}}}{k_{-4}^{\text{id}}} = -(\mu_{\text{P}(2)}^\circ + \mu_{\text{E}(3)}^\circ - \mu_{\text{EP}(3)}^\circ). \quad (45)$$

□

C. Chemostatting

Thermodynamically, the chemostatting procedure directly determines the chemical potentials $\{\mu_\alpha\}$ of the Y species. For ideal CRNs, where $\mu_\alpha(\mathbf{z}) = \mu_\alpha^{\text{id}}([\alpha])$, this means that the chemostatting procedure directly determines the concentrations, $[\alpha] = \exp((\mu_\alpha - \mu_\alpha^\circ)/RT)$. For non-ideal CRNs, the equivalence,

$$\mu_\alpha = \mu_\alpha^\circ + RT \ln[\alpha] + RT \ln \gamma_\alpha(\mathbf{z}, \mathbf{e}) \quad \forall \alpha \in Y, \quad (46)$$

does not directly determine the concentrations $\boldsymbol{\gamma}$ due to the dependence on the interactions, i.e., $[\alpha] = \exp((\mu_\alpha - \mu_\alpha^\circ)/RT)/\gamma_\alpha(\mathbf{z}, \mathbf{e}) \quad \forall \alpha \in Y$. Hence, $\mathbf{I}(t)$ in Eq. (2) [or $\mathbf{I}^Y(t)$ in Eq. (8)] describes the exchange currents such that the chemical potentials of the Y species are externally determined.

We now derive the explicit form of $\mathbf{I}^Y(t)$. To skip the details of the derivation, go directly to Eq. (51). The constraints imposed by Eq. (46) must be satisfied for every time t of the dynamics. Thus, also their time derivative must vanish,

$$\begin{aligned} \frac{RT}{[\alpha](t)} d_t[\alpha](t) + \nabla_z g_\alpha^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t)) \cdot d_t \mathbf{z}(t) \\ + \nabla_e g_\alpha^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t)) \cdot d_t \mathbf{e}(t) - d_t \mu_\alpha(t) = 0. \end{aligned} \quad (47)$$

Here, we used the compact notation $g_\alpha^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t)) = RT \ln \gamma_\alpha(\mathbf{z}(t), \mathbf{e}(t))$, and we accounted for the possible time-dependent control of the chemical potentials $\{\mu_\alpha\}$ and the parameters \mathbf{e} . The set of Eq. (47) for every $\alpha \in Y$ can be rewritten as

$$\mathbb{W}(t)(\mathbb{S}\mathbf{j}(\mathbf{z}(t)) + \mathbf{I}(t)) + \mathbf{g}_Y^{\text{in}}(t) - d_t \boldsymbol{\mu}_Y(t) = 0 \quad (48)$$

by introducing the vectors

$$\mathbf{g}_Y^{\text{in}}(t) = (\dots, \nabla_e g_\alpha^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t)) \cdot d_t \mathbf{e}(t), \dots)_{\alpha \in Y}^\top \quad (49)$$

and $\boldsymbol{\mu}_Y(t) = (\dots, \mu_\alpha(t), \dots)_{\alpha \in Y}^\top$ and the matrix $\mathbb{W}(t)$ whose entries are

$$\mathbb{W}_\beta^\alpha(t) = \frac{RT}{[\alpha](t)} \delta_{\alpha\beta} + \frac{\partial g_\alpha^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t))}{\partial [\beta]}, \quad (50)$$

with $\alpha \in Y$ being the row index and $\beta \in Z$ being the column index. We notice that the submatrix \mathbb{W}^Y , whose entries are $\{\mathbb{W}_\beta^\alpha\}_{\alpha, \beta \in Y}$, is square and not nonsingular so that it can be inverted. We thus obtain an explicit expression for $\mathbf{I}^Y(t)$,

$$\mathbf{I}^Y(t) = -(\mathbb{W}^Y(t))^{-1} (\mathbb{W}(t)\mathbb{S}\mathbf{j}(\mathbf{z}(t)) + \mathbf{g}_Y^{\text{in}}(t) - d_t \boldsymbol{\mu}_Y(t)). \quad (51)$$

Systems are said to be *autonomous* when the chemostatting procedure maintains the chemical potentials $\boldsymbol{\mu}_Y$ and the parameters \mathbf{e} constant, i.e., $d_t \boldsymbol{\mu}_Y(t) = 0$ and $\mathbf{g}_Y^{\text{in}}(t) = 0$. Otherwise, systems are said to be *nonautonomous*. The matrices $\mathbb{W}(t)$ and $\mathbb{W}^Y(t)$ link the value of the external currents to the concentrations of the internal species

that interact with the chemostatted ones. In the limit $\boldsymbol{y}(\boldsymbol{z}, \boldsymbol{e}) \rightarrow 1$, or equivalently $\nabla_{\boldsymbol{z}} G^{\text{in}}(\boldsymbol{z}, \boldsymbol{e}) \rightarrow 0$, $\boldsymbol{I}^Y(t)$ for ideal CRNs is recovered,

$$\begin{aligned} \boldsymbol{I}^Y(t) &= -\mathbb{S}^Y \boldsymbol{j}(\boldsymbol{z}(t)) + (\mathbb{W}^Y(t))^{-1} \boldsymbol{d}_t \boldsymbol{\mu}_Y^{\text{id}}(t) \\ &= -\mathbb{S}^Y \boldsymbol{j}(\boldsymbol{z}(t)) + \boldsymbol{d}_t \boldsymbol{y}(t), \end{aligned} \quad (52)$$

controlling directly the concentrations¹⁷ according to the protocol $\boldsymbol{d}_t \boldsymbol{y}(t)$.

Example 7. We specify the different terms in Eq. (51) for the CRN (9). According to Eq. (50), $\mathbb{W}(t)$ is given by

$$\mathbb{W} = \begin{array}{c} \begin{array}{ccccc} & E_{(3)} & ES_{(3)} & EP_{(3)} & A_{(1)}^- & A_{(2)}^- \\ \begin{array}{c} H_{(1)}^+ \\ H_{(2)}^+ \\ S_{(2)} \\ P_{(2)} \end{array} & \begin{pmatrix} 0 & 0 & 0 & -F\kappa & 0 \\ 0 & 0 & 0 & 0 & -F\kappa \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} & \begin{array}{c} H_{(1)}^+ \\ H_{(2)}^+ \\ S_{(2)} \\ P_{(2)} \end{array} \end{array} \left(\begin{array}{cccc} \frac{RT}{[H_{(1)}^+]} + F\kappa & 0 & 0 & 0 \\ 0 & \frac{RT}{[H_{(2)}^+]} + F\kappa & 0 & 0 \\ 0 & 0 & \frac{RT}{[S_{(2)}]} & 0 \\ 0 & 0 & 0 & \frac{RT}{[P_{(2)}]} \end{array} \right), \end{array} \quad (53)$$

and hence, $(\mathbb{W}^Y(t))^{-1}$ is given by

$$(\mathbb{W}^Y)^{-1} = \begin{array}{c} \begin{array}{cccc} & H_{(1)}^+ & H_{(2)}^+ & S_{(2)} & P_{(2)} \\ \begin{array}{c} H_{(1)}^+ \\ H_{(2)}^+ \\ S_{(2)} \\ P_{(2)} \end{array} & \begin{pmatrix} \frac{[H_{(1)}^+]}{RT+F\kappa[H_{(1)}^+]} & 0 & 0 & 0 \\ 0 & \frac{[H_{(2)}^+]}{RT+F\kappa[H_{(2)}^+]} & 0 & 0 \\ 0 & 0 & \frac{[S_{(2)}]}{RT} & 0 \\ 0 & 0 & 0 & \frac{[P_{(2)}]}{RT} \end{pmatrix} & \end{array} \end{array}. \quad (54)$$

Here, we did not write the explicit time dependence of the concentrations and the matrices for the sake of compactness. The vector $\boldsymbol{g}_Y^{\text{in}}(t)$ is specified as

$$\boldsymbol{g}_Y^{\text{in}}(t) = \begin{array}{c} \begin{array}{c} H_{(1)}^+ \\ H_{(2)}^+ \\ S_{(2)} \\ P_{(2)} \end{array} \left(\begin{array}{c} F\kappa(\boldsymbol{d}_t[C_{(1)}^+](t) - \boldsymbol{d}_t[D_{(1)}^-](t)) \\ F\kappa(\boldsymbol{d}_t[C_{(2)}^+](t) - \boldsymbol{d}_t[D_{(1)}^-](t)) \\ 0 \\ 0 \end{array} \right). \end{array} \quad (55)$$

□

D. Decomposition of the entropy production rate

We split the entropy production rate into different contributions. To do so, we adapt the decomposition of the entropy production for stochastic ideal CRNs developed in Ref. 14 to deterministic non-ideal CRNs. This is not straightforward and represents a major result of this work. To skip the details of the derivation, go directly to Eq. (68).

We start by expressing the entropy production rate of deterministic CRNs,^{16,17}

$$\dot{\Sigma}(t) = R \sum_{\rho \in \mathcal{R}} (j^{+\rho}(\boldsymbol{z}(t)) - j^{-\rho}(\boldsymbol{z}(t))) \ln \frac{j^{+\rho}(\boldsymbol{z}(t))}{j^{-\rho}(\boldsymbol{z}(t))} \geq 0. \quad (56)$$

Using the local detailed balance (38) and the specific expressions for the free energies of reaction (39), we get

$$T\dot{\Sigma}(t) = -\boldsymbol{\mu}(\boldsymbol{z}(t)) \cdot \mathbb{S} \boldsymbol{j}(\boldsymbol{z}(t)). \quad (57)$$

By explicitly considering the contributions of the X , Y_f , and Y_p species, we obtain

$$T\dot{\Sigma}(t) = -\left(\boldsymbol{\mu}_X(\boldsymbol{z}(t)) \cdot \mathbb{S}^X + \boldsymbol{\mu}_{Y_f}(t) \cdot \mathbb{S}^{Y_f} + \boldsymbol{\mu}_{Y_p}(t) \cdot \mathbb{S}^{Y_p} \right) \boldsymbol{j}(\boldsymbol{z}(t)), \quad (58)$$

with $\boldsymbol{\mu}_i(\cdot) = (\dots, \mu_\alpha(\cdot), \dots)_{\alpha \in i}^\top$ and $i = \{X, Y_f, Y_p\}$. Note that we do not explicitly write the $\boldsymbol{z}(t)$ dependence of the chemical potentials of the Y species since they are externally controlled.

A part of the free energy externally exchanged through the currents $\{I^\alpha(t)\}$ modifies the free energy of the system. Another part is transferred through the system. To proceed with the entropy production decomposition, we need to distinguish these two parts. To do so, we exploit the conservation laws. We consider

$$\mathbb{L}^b \mathbb{S} = 0 = \mathbb{L}_X^b \mathbb{S}^X + \mathbb{L}_{Y_f}^b \mathbb{S}^{Y_f} + \mathbb{L}_{Y_p}^b \mathbb{S}^{Y_p}, \quad (59)$$

and recalling that the matrix $\mathbb{L}_{Y_p}^b$ can be inverted (see Subsection II C), we write

$$\mathbb{S}^{Y_p} = -(\mathbb{L}_{Y_p}^b)^{-1} \left(\mathbb{L}_X^b \mathbb{S}^X + \mathbb{L}_{Y_f}^b \mathbb{S}^{Y_f} \right). \quad (60)$$

Equation (60) binds the net variation of the number of molecules for the Y_p species to that of the X and Y_f species. We substitute this result in Eq. (58),

$$\begin{aligned} T\dot{\Sigma}(t) &= -\left(\boldsymbol{\mu}_X(\boldsymbol{z}(t)) \cdot \mathbb{S}^X + \boldsymbol{\mu}_{Y_f}(t) \cdot \mathbb{S}^{Y_f} \right. \\ &\quad \left. - \boldsymbol{\mu}_{Y_p}(t) \cdot (\mathbb{L}_{Y_p}^b)^{-1} \left(\mathbb{L}_X^b \mathbb{S}^X + \mathbb{L}_{Y_f}^b \mathbb{S}^{Y_f} \right) \right) \boldsymbol{j}(\boldsymbol{z}(t)). \end{aligned} \quad (61)$$

By adding and subtracting $\boldsymbol{\mu}_{Y_p}(t) \cdot (\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}_{Y_p}^b \mathbb{S}^{Y_p}$, the entropy production rate becomes

$$T\dot{\Sigma}(t) = -\mathcal{F}(\boldsymbol{z}(t)) \cdot \mathbb{S} \boldsymbol{j}(\boldsymbol{z}(t)), \quad (62)$$

where

$$\mathcal{F}(\mathbf{z}(t)) = \left(\boldsymbol{\mu}(\mathbf{z}(t)) \cdot \mathbb{1} - \boldsymbol{\mu}_{Y_p}(t) \cdot (\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}^b \right)^\top, \quad (63)$$

with $\mathbb{1}$ being the identity matrix. Note that we just introduced a term in the expression of the entropy production rate, i.e., $\boldsymbol{\mu}_{Y_p}(t) \cdot (\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}^b$, whose net contribution vanishes since $\mathbb{L}^b \mathbb{S} = 0$. Introducing this term allows us to recognize the proper thermodynamic potential of open non-ideal CRNs.

To do so, we define the following state function:

$$\Psi(\mathbf{z}(t)) = \mathcal{F}(\mathbf{z}(t)) \cdot \mathbf{z}(t), \quad (64)$$

whose time derivative according to the rate equation (2) reads

$$\begin{aligned} d_t \Psi(\mathbf{z}(t)) &= \mathcal{F}(\mathbf{z}(t)) \cdot \mathbb{S}j(\mathbf{z}(t)) + \mathcal{F}_Y(t) \cdot \mathbf{I}^Y(t) \\ &\quad + RT d_t \|\mathbf{z}(t)\| - d_t \boldsymbol{\mu}_{Y_p}(t) \cdot (\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}^b \mathbf{z}(t) \\ &\quad + d_t (\nabla_z G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t))) \cdot \mathbf{z}(t), \end{aligned} \quad (65)$$

where $\mathcal{F}_Y(t) = (\dots, \mathcal{F}_\alpha(t), \dots)_{\alpha \in Y}$, $\|\mathbf{z}(t)\| = \sum_\alpha [\alpha](t)$. The time dependence of $\mathbf{e}(t)$ accounts for the (possible) time dependent manipulation of the parameters.

To proceed with the decomposition, we first recognize that the term $\mathcal{F}_Y(t) \cdot \mathbf{I}^Y(t)$ simplifies to $\mathcal{F}_{Y_f}(t) \cdot \mathbf{I}^{Y_f}(t)$ since $\mathcal{F}_\alpha = 0$ for $\alpha \in Y_p$. We then notice that the term $d_t (\nabla_z G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t))) \cdot \mathbf{z}(t)$ can be written as

$$\begin{aligned} &d_t (\nabla_z G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t))) \cdot \mathbf{z}(t) \\ &= d_t (\nabla_z G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t))) \cdot \mathbf{z}(t) - \nabla_z G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t)) \cdot d_t \mathbf{z}(t) \\ &= d_t (\nabla_z G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t))) \cdot \mathbf{z}(t) - d_t G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t)) \\ &\quad + \nabla_e G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t)) \cdot d_t \mathbf{e}(t) \end{aligned} \quad (66)$$

by using the time derivative of the interaction free energy,

$$\begin{aligned} d_t G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t)) &= \nabla_z G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t)) \cdot d_t \mathbf{z}(t) \\ &\quad + \nabla_e G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t)) \cdot d_t \mathbf{e}(t), \end{aligned} \quad (67)$$

with $\nabla_e = (\dots, \partial/\partial e_i, \dots)^\top$.

We now exploit Eq. (65) to express $\mathcal{F}(\mathbf{z}(t)) \cdot \mathbb{S}j(\mathbf{z}(t))$ as a function of the other terms, and we substitute it in Eq. (62). By doing so and collecting all the full time derivatives in a single term denoted as $d_t \mathcal{G}$, we obtain the following decomposition of the entropy production rate, which constitutes a central result of this work:

$$T \dot{\Sigma}(t) = -d_t \mathcal{G}(\mathbf{z}(t)) + \dot{w}_{\text{nc}}(t) + \dot{w}_{\text{driv}}(t) \geq 0. \quad (68)$$

Here, we introduced the semigrand free energy of non-ideal systems as

$$\begin{aligned} \mathcal{G}(\mathbf{z}) &= \boldsymbol{\mu}(\mathbf{z}) \cdot \mathbf{z} - RT \|\mathbf{z}\| - \boldsymbol{\mu}_{Y_p}(t) \cdot \mathbf{m}(\mathbf{z}) \\ &\quad + G^{\text{in}}(\mathbf{z}, \mathbf{e}) - \nabla_z G^{\text{in}}(\mathbf{z}, \mathbf{e}) \cdot \mathbf{z}, \end{aligned} \quad (69)$$

with the concentrations of the so-called moieties,

$$\mathbf{m}(\mathbf{z}) = (\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}^b \mathbf{z}. \quad (70)$$

They represent parts of (or entire) molecules that are exchanged with the environment. The semigrand free energy (69) can be rewritten in a more appealing ways as

$$\begin{aligned} \mathcal{G}(\mathbf{z}) &= \boldsymbol{\mu}^{\text{id}}(\mathbf{z}) \cdot \mathbf{z} - RT \|\mathbf{z}\| - \boldsymbol{\mu}_{Y_p}(t) \cdot \mathbf{m}(\mathbf{z}) + G^{\text{in}}(\mathbf{z}, \mathbf{e}) \\ &= G(\mathbf{z}, \mathbf{e}) - \boldsymbol{\mu}_{Y_p}(t) \cdot \mathbf{m}(\mathbf{z}) \end{aligned} \quad (71)$$

by using Eq. (34). The term $\boldsymbol{\mu}^{\text{id}}(\mathbf{z}) \cdot \mathbf{z} - RT \|\mathbf{z}\|$ is the free energy of closed ideal CRNs, i.e., $G^{\text{id}}(\mathbf{z})$, accounting for the chemical energy of every species. The term $G^{\text{in}}(\mathbf{z}, \mathbf{e})$ is the interaction free energy. Their sum gives the free energy $G(\mathbf{z}, \mathbf{e})$ in Eq. (29). The term $\boldsymbol{\mu}_{Y_p}(t) \cdot \mathbf{m}(\mathbf{z})$ is the energetic contribution of the matter exchanged with the environment. It vanishes for closed systems. Since $\mathcal{G}(\mathbf{z})$ is a state function, its time derivative vanishes at steady state.

The driving work rate $\dot{w}_{\text{driv}}(t)$ can be split into the sum of the chemical and the interaction driving work rates $\dot{w}_{\text{driv}}(t) = \dot{w}_{\text{driv}}^{\text{ch}}(t) + \dot{w}_{\text{driv}}^{\text{in}}(t)$ that are specified as

$$\dot{w}_{\text{driv}}^{\text{ch}}(t) = -d_t \boldsymbol{\mu}_{Y_p}(t) \cdot \mathbf{m}(\mathbf{z}(t)) \quad (72)$$

and

$$\dot{w}_{\text{driv}}^{\text{in}}(t) = \nabla_e G^{\text{in}}(\mathbf{z}(t), \mathbf{e}(t)) \cdot d_t \mathbf{e}(t), \quad (73)$$

respectively. The chemical driving work (72) accounts for the time dependent control of only the chemical potentials $\boldsymbol{\mu}_{Y_p}(t)$. The interaction driving work (73) accounts for the time dependent control of the interaction free energy through the time evolution of the parameters $\mathbf{e}(t)$. They represent the energetic cost of modifying the equilibrium that is defined if $Y = Y_p$ (as shown in Subsection II E, open CRNs with $Y = Y_p$ have a well-defined equilibrium state for every value of \mathbf{e}). These two terms vanish in autonomous systems.

The nonconservative work rate is given by

$$\dot{w}_{\text{nc}}(t) = \mathcal{F}_{Y_f}(t) \cdot \mathbf{I}^{Y_f}(t). \quad (74)$$

It quantifies the energetic cost to sustain fluxes of chemical species through the CRN by means of the fundamental nonconservative forces,

$$\mathcal{F}_{Y_f}(t) = \left(\boldsymbol{\mu}_{Y_f}(t) \cdot \mathbb{1} - \boldsymbol{\mu}_{Y_p}(t) \cdot (\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}_{Y_f}^b \right)^\top. \quad (75)$$

These are the forces keeping the system out of equilibrium. Indeed, they vanish when the CRN is detailed balanced. This always happens when all the chemostatted species break a conservation law $Y = Y_p$, and for this reason, the CRNs are said *unconditionally* detailed balanced. Therefore, the Y_f species are named *force* species. The Y_p species are named *potential* species since they fix the equilibrium conditions. Thus, the splitting $Y = Y_f \cup Y_p$ must be motivated

by the different thermodynamic role of the Y_f species and Y_p species. The nonconservative work accounts for autonomous mechanisms keeping the system out of equilibrium unlike the chemical driving work (72) and the interaction driving work (73). We stress that the expression of the fundamental force (75) for non-ideal CRNs is analogous to that found for ideal CRNs.¹⁴ The effects of the interactions that modify the value of the nonconservative work (74) are hidden in the expressions of the exchanged currents (51). Note also that the force (75) can vanish, and hence, the CRN becomes detailed balanced, even if $Y \neq Y_p$. Indeed, $\mathcal{F}_{Y_f} = 0$ when the chemical potentials of the Y_f species satisfy $\mu_{Y_f}^\top = \mu_{Y_p} \cdot (\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}_{Y_f}^b$. Thus, the fundamental nonconservative forces are generated by the difference between the actual values of chemical potentials of the Y_f species, i.e., μ_{Y_f} , and their equilibrium values if only the Y_p species were chemostatted, i.e., $\mu_{Y_p} \cdot (\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}_{Y_f}^b$.

Finally, we notice that for autonomous [$\dot{w}_{\text{driv}}^{\text{ch}}(t) = \dot{w}_{\text{driv}}^{\text{in}}(t) = 0$] detailed balanced [$\dot{w}_{\text{nc}}(t) = 0$] CRNs, the semigrand free energy decreases monotonously in time, i.e., $d_t \mathcal{G}(t) = -T\bar{\Sigma}(t) \leq 0$. The semigrand free energy is dissipated.

Example 8. We specify the moieties (70), the interaction driving work (73), the nonconservative forces (75), and the nonconservative work (74) for the open CRN (9). We start by giving an explicit expression for the matrix $(\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}^b$ [with \mathbb{L}^b given in Eq. (16)],

$$(\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}^b = \begin{array}{c} \text{H}_2^+ \\ \text{S}_2 \\ \text{P}_2 \end{array} \begin{array}{c|ccc|ccc} \text{E}_{(3)} & \text{ES}_{(3)} & \text{EP}_{(3)} & \text{A}_{(1)} & \text{A}_{(2)} & \text{H}_{(1)}^+ & \text{H}_{(2)}^+ & \text{S}_{(2)} & \text{P}_{(2)} \\ \hline & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{array}. \quad (76)$$

The vertical gray lines mark the split of $(\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}^b$ into $(\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}_{X}^b$, $(\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}_{Y_f}^b$, and $(\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}_{Y_p}^b = 1$. By using Eq. (76) in the definition of the concentrations of the moieties (70), we obtain that

$$\mathbf{m}(\mathbf{z}) = \begin{pmatrix} [\text{H}_{(1)}^+] + [\text{H}_{(2)}^+] \\ [\text{ES}_{(3)}] + [\text{S}_{(2)}] - [\text{H}_{(1)}^+] \\ [\text{EP}_{(3)}] + [\text{P}_{(2)}] + [\text{H}_{(1)}^+] \end{pmatrix}. \quad (77)$$

Only the first moiety has a straightforward interpretation as the protons. The others are a linear combination of broken conserved quantities accounting for the exchange of fragments of molecules through the CRN.

We now consider the interaction driving work (73). By using the expression of the interaction free energy for this example (35), we obtain

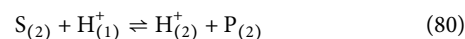
$$\dot{w}_{\text{driv}}^{\text{in}}(t) = F\kappa(Q_{(1)}(\mathbf{z}(t), \mathbf{e}(t))(d_t[C_{(1)}^+](t) - d_t[D_{(1)}^-](t)) + Q_{(2)}(\mathbf{z}(t), \mathbf{e}(t))(d_t[C_{(2)}^+](t) - d_t[D_{(2)}^-](t))), \quad (78)$$

accounting for the variation of the semigrand free energy due to possible nonautonomous processes changing the concentrations of the cations C^+ and the anions D^- in the two compartments.

Finally, we turn to the nonconservative forces (75). There is only one nonconservative force for the open CRN (9),

$$\mathcal{F}_{Y_f}(t) = \mu_{\text{H}_{(1)}^+}(t) - \mu_{Y_p}(t) \cdot (\mathbb{L}_{Y_p}^b)^{-1} \mathbb{L}_{Y_f}^b = -(\mu_{\text{P}_{(2)}}^{\text{id}}(t) + \mu_{\text{H}_{(2)}^+}(t) - \mu_{\text{H}_{(1)}^+}(t) - \mu_{\text{S}_{(2)}}^{\text{id}}(t)), \quad (79)$$

since $\mu_{Y_p}(t) = (\mu_{\text{H}_{(2)}^+}(t), \mu_{\text{S}_{(2)}}^{\text{id}}(t), \mu_{\text{P}_{(2)}}^{\text{id}}(t))^\top$ [recall that the chemical potential of $\text{S}_{(2)}$ and $\text{P}_{(2)}$ is the same as that for ideal CRNs]. It represents the thermodynamic force (i.e., the free energy of reaction) of the following effective chemical reaction:



between chemostatted species. Note that the role of the chemostatted species is precisely that of keeping this effective reaction out of equilibrium by imposing the thermodynamic force (79). The corresponding nonconservative work reads

$$\dot{w}_{\text{nc}}(t) = -(\mu_{\text{P}_{(2)}}^{\text{id}}(t) + \mu_{\text{H}_{(2)}^+}(t) - \mu_{\text{H}_{(1)}^+}(t) - \mu_{\text{S}_{(2)}}^{\text{id}}(t))I^{\text{H}_{(1)}^+}(t). \quad (81)$$

We stress that, despite Eq. (81) resembling the corresponding expression for ideal CRNs, the interactions affect the value of the exchange current, $I^{\text{H}_{(1)}^+}(t)$. \square

E. Steady-state entropy production

When autonomous CRNs reach a nonequilibrium steady state $\bar{\mathbf{z}}$, the entropy production rate (68) simplifies to

$$T\bar{\Sigma} = \bar{w}_{\text{nc}} = \mathcal{F}_{Y_f} \cdot \bar{\mathbf{I}}^{Y_f}, \quad (82)$$

highlighting the physical role of the nonconservative work: it provides the energy to balance dissipation and keep CRNs out of equilibrium.

An equivalent decomposition can be achieved using the cycle splitting of the current vector (19) in Eq. (57),

$$\begin{aligned} T\bar{\Sigma} &= -\boldsymbol{\mu} \cdot \bar{\mathbf{S}}\bar{\mathbf{j}} \\ &= -\boldsymbol{\mu} \cdot \mathbf{S} \left\{ \sum_i \mathbf{c}_i \psi^i + \sum_\epsilon \mathbf{c}_\epsilon \psi^\epsilon \right\} \\ &= -\boldsymbol{\mu}_Y \cdot \hat{\mathbf{S}}^Y \boldsymbol{\psi}, \end{aligned} \quad (83)$$

where we used Eqs. (17) and (18). The vector $\boldsymbol{\psi}$ collects the coefficients $\{\psi^\epsilon\}$, i.e., $\boldsymbol{\psi} = (\dots, \psi^\epsilon, \dots)^\top$. Each ϵ column of the effective stoichiometric matrix $\hat{\mathbf{S}}_\epsilon^Y = \mathbf{S}^Y \mathbf{c}_\epsilon$ codifies the net stoichiometry of the chemostatted species along the ϵ emergent cycle.

This alternative decomposition of the entropy production has two advantages. First, it emphasizes the role of the emergent cycles in the steady-state dissipation. They are out of equilibrium effective reactions interconverting chemostatted species according to the stoichiometry codified in $\hat{\mathbf{S}}^Y$. Second, this decomposition can

be used to develop thermodynamically consistent coarse-graining theories as it has been done for ideal CRNs.^{48,49}

Example 9. We specify the effective stoichiometric matrix \hat{S}^Y for the CRN (9) given the emergent cycle (21),

$$\hat{S}^Y = \begin{matrix} H_{(1)}^+ \\ H_{(2)}^+ \\ S_{(2)} \\ P_{(2)} \end{matrix} \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}. \quad (84)$$

Notice that the corresponding effective reaction is equal to the one we inferred from the nonconservative force in Eq. (80). This equivalence does not hold in general. \square

F. Lower bound of the semigrand free energy

We now examine how the assumptions on the free energy $G(\mathbf{z}, \mathbf{e})$ introduced in Subsection III A imply that the semigrand free energy $\mathcal{G}(\mathbf{z})$ (71) is lower bounded and thus guarantee thermodynamic consistency.

Since $G(\mathbf{z}, \mathbf{e})$ is lower bounded and increases superlinearly as concentrations go to infinity by assumption, also $\mathcal{G}(\mathbf{z})$ [see Eq. (71)] is lower bounded for finite values of the chemical potential of the chemostatted species. Indeed, as concentrations become infinitely large, $-\boldsymbol{\mu}_{Y_p} \cdot \mathbf{m}$ can go to minus infinity slower (linearly with \mathbf{z}) than $G(\mathbf{z}, \mathbf{e})$ goes to infinity (superlinearly with \mathbf{z}). Thus, $\exists \mathcal{G}_{\min}$ such that

$$\mathcal{G}(\mathbf{z}) \geq \mathcal{G}_{\min}, \quad \forall \mathbf{z}. \quad (85)$$

This ensures thermodynamic consistency: autonomous detailed balanced systems relax toward an equilibrium state. Given that (i) $d_t \mathcal{G}(\mathbf{z}(t)) = -T\dot{\Sigma}(t) \leq 0$ (see Subsection III D), (ii) $\dot{\Sigma}(t) = 0$ only at equilibrium, and (iii) $\mathcal{G}(\mathbf{z}(t))$ cannot decrease indefinitely because of Eq. (85), the system must reach an equilibrium state, $\lim_{t \rightarrow +\infty} \mathbf{z}(t) = \mathbf{z}_{\text{eq}}$, and $\mathcal{G}(\mathbf{z}(t)) \geq \mathcal{G}(\mathbf{z}_{\text{eq}})$, $\forall t$.

The specific equilibrium state \mathbf{z}_{eq} to which the system relaxes depends on the initial conditions. In ideal CRNs, there is a unique equilibrium state for each stoichiometric compatibility class, here labeled $\Omega(\{L^{\lambda_u}\})$ (see Subsection II E). In non-ideal CRNs, there might be multiple equilibria for every stoichiometric compatibility class. Thus, we introduce $\Omega(\mathbf{z}_{\text{eq}}) \subseteq \Omega(\{L^{\lambda_u}\})$ as the manifold in the concentration space such that any state in $\Omega(\mathbf{z}_{\text{eq}})$ will relax toward the specific equilibrium \mathbf{z}_{eq} . This implies that

$$\mathcal{G}(\mathbf{z}) \geq \mathcal{G}(\mathbf{z}_{\text{eq}}), \quad \forall \mathbf{z} \in \Omega(\mathbf{z}_{\text{eq}}). \quad (86)$$

Let us stress two more points. First, the bound (86) does not provide any constraint for the values of the semigrand free energy evaluated on different manifolds. Consequently, it is possible that $\exists \mathbf{z} \notin \Omega(\mathbf{z}_{\text{eq}})$ such that $\mathcal{G}(\mathbf{z}) < \mathcal{G}(\mathbf{z}_{\text{eq}})$. Second, \mathcal{G}_{\min} must be an equilibrium value of the semigrand free energy. Indeed, if there was a nonequilibrium state \mathbf{z} such that $\mathcal{G}(\mathbf{z}) = \mathcal{G}_{\min}$, the relaxation of \mathbf{z} would necessarily correspond to a decrease in the semigrand free energy, which is not possible because of Eq. (85).

We now turn to nonautonomous and/or nondetailed balanced CRNs. The expression of the semigrand free energy (71) is exactly the same as for the corresponding autonomous detailed balanced CRNs. Thus, Eq. (86) still holds with \mathbf{z}_{eq} now being the equilibrium state of the corresponding detailed balanced CRNs with $Y = Y_p$. However, the evolution of $\mathbf{z}(t)$ is constrained on a specific stoichiometric compatibility class $\Omega(\{L^{\lambda_u}\})$, but not (necessarily) on a specific manifold $\Omega(\mathbf{z}_{\text{eq}})$. This means that different bounds (86) can be identified during the same nondetailed balanced dynamics.

In conclusion, the semigrand free energy is always lower bounded by $\mathcal{G}(\mathbf{z}_{\text{eq}})$, where \mathbf{z}_{eq} is the equilibrium state to which the system would relax if it was detailed balanced. This, together with $d_t \mathcal{G}(\mathbf{z}(t)) = -T\dot{\Sigma}(t) \leq 0$ for autonomous detailed balanced CRNs, makes the semigrand free energy defined in Eq. (69) the proper thermodynamic potential of open non-ideal CRNs. In other words, $\mathcal{G}(\mathbf{z}(t)) - \mathcal{G}(\mathbf{z}_{\text{eq}})$ is a Lyapunov function.

G. Relative entropy

We show here that, because of the interactions, the difference $\mathcal{G}(\mathbf{z}(t)) - \mathcal{G}(\mathbf{z}_{\text{eq}})$ cannot be solely expressed as relative entropy. This is a significant difference with respect to previous studies on the thermodynamics of ideal CRNs^{14,17} and ideal reaction-diffusion systems.^{18,19}

To this aim, we notice that, because of the local detailed balance (38), the equilibrium chemical potentials satisfy

$$\boldsymbol{\mu}(\mathbf{z}_{\text{eq}}) \cdot \mathbb{S} = 0. \quad (87)$$

Thus, $\boldsymbol{\mu}(\mathbf{z}_{\text{eq}})$ can be written as a linear combination of conservation laws (11), $\boldsymbol{\mu}(\mathbf{z}_{\text{eq}}) = \sum_{\lambda} f_{\lambda} \boldsymbol{\ell}^{\lambda}$. Recalling that the conservation laws split into unbroken and broken conservation laws in open CRNs, we obtain

$$\boldsymbol{\mu}(\mathbf{z}_{\text{eq}}) \cdot \mathbf{z}_{\text{eq}} = \boldsymbol{\mu}(\mathbf{z}_{\text{eq}}) \cdot \mathbf{z}(t) - \sum_{\lambda_b} f_{\lambda_b} \boldsymbol{\ell}^{\lambda_b} \cdot \mathbf{z}(t) + \sum_{\lambda_b} f_{\lambda_b} \boldsymbol{\ell}^{\lambda_b} \cdot \mathbf{z}_{\text{eq}}. \quad (88)$$

By recalling that the unbroken conservation laws are those with null entries for the chemostatted species, hence $\mu_{\alpha}(\mathbf{z}_{\text{eq}}) = \sum_{\lambda_b} f_{\lambda_b} \ell_{\alpha}^{\lambda_b}$ for $\alpha \in Y$, and the expressions for the concentrations of the moieties (70), we verify that

$$\boldsymbol{\mu}_{Y_p}(t) \cdot \mathbf{m}(\mathbf{z}(t)) = \sum_{\lambda_b} f_{\lambda_b} \boldsymbol{\ell}^{\lambda_b} \cdot \mathbf{z}(t), \quad (89)$$

$$\boldsymbol{\mu}_{Y_p}(t) \cdot \mathbf{m}(\mathbf{z}_{\text{eq}}) = \sum_{\lambda_b} f_{\lambda_b} \boldsymbol{\ell}^{\lambda_b} \cdot \mathbf{z}_{\text{eq}}. \quad (90)$$

By using Eqs. (88)–(90), we find that

$$\begin{aligned} \mathcal{G}(\mathbf{z}(t)) - \mathcal{G}(\mathbf{z}_{\text{eq}}) &= (\boldsymbol{\mu}(\mathbf{z}(t)) - \boldsymbol{\mu}(\mathbf{z}_{\text{eq}})) \cdot \mathbf{z}(t) + \\ &\quad - RT(\|\mathbf{z}(t)\| - \|\mathbf{z}_{\text{eq}}\|) \\ &\quad + G^{\text{in}}(\mathbf{z}, \mathbf{e}) - \nabla_{\mathbf{z}} G^{\text{in}}(\mathbf{z}, \mathbf{e}) \cdot \mathbf{z} + \\ &\quad - G^{\text{in}}(\mathbf{z}_{\text{eq}}, \mathbf{e}) + \nabla_{\mathbf{z}} G^{\text{in}}(\mathbf{z}_{\text{eq}}, \mathbf{e}) \cdot \mathbf{z}_{\text{eq}}. \end{aligned} \quad (91)$$

Finally, by using Eq. (34) to split the ideal and the non-ideal contribution of the chemical potential, we obtain

$$\mathcal{G}(\mathbf{z}(t)) - \mathcal{G}(\mathbf{z}_{\text{eq}}) = RT\mathcal{L}(\mathbf{z}(t) \parallel \mathbf{z}_{\text{eq}}) + G^{\text{in}}(\mathbf{z}, \mathbf{e}) - G^{\text{in}}(\mathbf{z}_{\text{eq}}, \mathbf{e}) - \nabla_{\mathbf{z}} G^{\text{in}}(\mathbf{z}_{\text{eq}}, \mathbf{e}) \cdot (\mathbf{z} - \mathbf{z}_{\text{eq}}), \quad (92)$$

where we introduced the relative entropy for non-normalized concentration distributions,

$$\mathcal{L}(\mathbf{a} \parallel \mathbf{b}) = \sum_i a_i \ln\left(\frac{a_i}{b_i}\right) - (a_i - b_i) \geq 0. \quad (93)$$

Equation (92) proves that $\mathcal{G}(\mathbf{z}(t)) - \mathcal{G}(\mathbf{z}_{\text{eq}})$ cannot be written only as a relative entropy, but additional terms emerge from the interactions. It is not granted that these additional terms are always positive unless $G^{\text{in}}(\mathbf{z}, \mathbf{e})$ is a convex function of \mathbf{z} . Nevertheless, if $\mathbf{z}(t) \in \Omega(\mathbf{z}_{\text{eq}})$, $\mathcal{G}(\mathbf{z}(t)) - \mathcal{G}(\mathbf{z}_{\text{eq}}) \geq 0$ as proved in Subsection III F.

Example 10. The difference $\mathcal{G}(\mathbf{z}) - \mathcal{G}(\mathbf{z}_{\text{eq}})$ in the CRN (9), where \mathbf{z}_{eq} is the equilibrium to which the system would relax if only the Y_p species were chemostatted, can be written as

$$\mathcal{G}(\mathbf{z}) - \mathcal{G}(\mathbf{z}_{\text{eq}}) = RT\mathcal{L}(\mathbf{z}(t) \parallel \mathbf{z}_{\text{eq}}) + \frac{F\kappa}{2} (Q_{(1)}(\mathbf{z}, \mathbf{e}) - Q_{(1)}(\mathbf{z}_{\text{eq}}, \mathbf{e}))^2 + \frac{F\kappa}{2} (Q_{(2)}(\mathbf{z}, \mathbf{e}) - Q_{(2)}(\mathbf{z}_{\text{eq}}, \mathbf{e}))^2. \quad (94)$$

Since the interaction free energy (35) is a convex function, the terms in Eq. (94) emerging from the electrostatic interactions are always positive, i.e., $(Q_{(i)}(\mathbf{z}, \mathbf{e}) - Q_{(i)}(\mathbf{z}_{\text{eq}}, \mathbf{e}))^2 \geq 0$. \square

Example 11. We examine the difference $\mathcal{G}(\mathbf{z}) - \mathcal{G}(\mathbf{z}_{\text{eq}})$ for the following closed CRN:



when the electrostatic interactions are described within Debye-Hückel theory,⁴¹

$$G^{\text{in}}(\mathbf{z}) = -\frac{4F\kappa}{h^3} \left\{ \frac{h^2}{2} \mathcal{I}(\mathbf{z}) - h\sqrt{\mathcal{I}(\mathbf{z})} + \ln[h\sqrt{\mathcal{I}(\mathbf{z})} + 1] \right\}, \quad (96)$$

where $\mathcal{I}(\mathbf{z})$ is the ionic strength,

$$\mathcal{I}(\mathbf{z}) = \frac{1}{2} ([\text{A}^-] + [\text{C}^+]) \geq 0, \quad (97)$$

and κ and h are positive constants depending on the specific system (see Ref. 41 for their explicit expressions). In this case, $\mathcal{G}(\mathbf{z}) - \mathcal{G}(\mathbf{z}_{\text{eq}})$ becomes

$$\mathcal{G}(\mathbf{z}) - \mathcal{G}(\mathbf{z}_{\text{eq}}) = RT\mathcal{L}(\mathbf{z}(t) \parallel \mathbf{z}_{\text{eq}}) + (G^{\text{in}}(\mathbf{z}) - G^{\text{in}}(\mathbf{z}_{\text{eq}})) + 2F\kappa[\mathcal{I}(\mathbf{z}) - \mathcal{I}(\mathbf{z}_{\text{eq}})] \frac{\sqrt{\mathcal{I}(\mathbf{z}_{\text{eq}})}}{1 + h\sqrt{\mathcal{I}(\mathbf{z}_{\text{eq}})}}. \quad (98)$$

Here, the terms emerging from the electrostatic interactions may be negative because the interaction free energy provided by the Debye-Hückel theory is not a convex function. \square

IV. CONCLUSIONS

In this work, we developed a thermodynamic theory of non-ideal CRNs. We used activity coefficients to account for interactions within a mean-field approach. Our framework is general and holds for any (thermodynamic consistent) expressions of the interactions. Exploiting the local detailed balance assumption, we showed that the dynamics of non-ideal CRNs does not follow mass-action kinetics unlike for ideal-CRNs. Generalizing some results developed for ideal-CRNs, we used conservation laws and cycles to derive physically meaningful decompositions of the entropy production rate. The first (68) is derived exploiting the conservation laws and allowed us to determine the proper thermodynamic potential and the forces driving non-ideal CRNs out of equilibrium. The second (83) is derived exploiting the cycles and allowed us to express the steady-state entropy production using only the chemical potential of the chemostatted species and the currents along the emergent cycles. We proved that the thermodynamic potential acts as a Lyapunov function in detailed balanced CRNs.

Crucially, we showed that the thermodynamic structure of non-ideal CRNs is the same as in ideal CRNs. In particular, the conservation laws and cycles determine whether systems are detailed balanced or not and define the forces (75) maintaining nonequilibrium regimes in the same way for ideal and non-ideal CRNs.

These results are fundamental to characterize processes of energy transduction in biosystems and electrochemical systems where interacting ions are ubiquitous. As a perspective, one could exploit the same strategy we followed in this work to specialize the stochastic thermodynamics of CRNs to interacting systems. We leave these points to future investigations.

ACKNOWLEDGMENTS

This research was funded by the European Research Council project NanoThermo (ERC-2015-CoG Agreement No. 681456).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

- 1 I. Prigogine, *Introduction to Thermodynamics of Irreversible Processes*, 2nd ed. (Interscience Publishers, New York, 1961).
- 2 S. R. de Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (Dover, 1984).
- 3 T. L. Hill, *Free Energy Transduction in Biology* (Academic Press, New York, 1977).
- 4 J. Schnakenberg, "Network theory of microscopic and macroscopic behavior of master equation systems," *Rev. Mod. Phys.* **48**, 571–585 (1976).
- 5 L. Jiu-li, C. Van den Broeck, and G. Nicolis, "Stability criteria and fluctuations around nonequilibrium states," *Z. Phys. B: Condens. Matter* **56**, 165–170 (1984).

- ⁶C. Y. Mou, J. I. Luo, and G. Nicolis, "Stochastic thermodynamics of nonequilibrium steady states in chemical reaction systems," *J. Chem. Phys.* **84**, 7011–7017 (1986).
- ⁷J. Ross, K. L. C. Hunt, and P. M. Hunt, "Thermodynamics far from equilibrium: Reactions with multiple stationary states," *J. Chem. Phys.* **88**, 2719–2729 (1988).
- ⁸C. Jarzynski, "Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale," *Annu. Rev. Condens. Matter Phys.* **2**, 329–351 (2011).
- ⁹U. Seifert, "Stochastic thermodynamics, fluctuation theorems and molecular machines," *Rep. Prog. Phys.* **75**, 126001 (2012).
- ¹⁰C. V. den Broeck and M. Esposito, "Ensemble and trajectory thermodynamics: A brief introduction," *Physica A* **418**, 6–16 (2015).
- ¹¹Dilip Kondepudi and Ilya Prigogine, *Modern Thermodynamics: From Heat Engines to Dissipative Structures*, 2nd ed. (Wiley, 2015).
- ¹²P. Gaspard, "Fluctuation theorem for nonequilibrium reactions," *J. Chem. Phys.* **120**, 8898–8905 (2004).
- ¹³T. Schmiedl and U. Seifert, "Stochastic thermodynamics of chemical reaction networks," *J. Chem. Phys.* **126**, 044101 (2007).
- ¹⁴R. Rao and M. Esposito, "Conservation laws and work fluctuation relations in chemical reaction networks," *J. Chem. Phys.* **149**, 245101 (2018).
- ¹⁵H. Qian and D. A. Beard, "Thermodynamics of stoichiometric biochemical networks in living systems far from equilibrium," *Biophys. Chem.* **114**, 213–220 (2005).
- ¹⁶M. Polettini and M. Esposito, "Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws," *J. Chem. Phys.* **141**, 024117 (2014).
- ¹⁷R. Rao and M. Esposito, "Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics," *Phys. Rev. X* **6**, 041064 (2016).
- ¹⁸G. Falasco, R. Rao, and M. Esposito, "Information thermodynamics of Turing patterns," *Phys. Rev. Lett.* **121**, 108301 (2018).
- ¹⁹F. Avanzini, G. Falasco, and M. Esposito, "Thermodynamics of chemical waves," *J. Chem. Phys.* **151**, 234103 (2019).
- ²⁰F. Avanzini, G. Falasco, and M. Esposito, "Chemical cloaking," *Phys. Rev. E* **101**, 060102 (2020).
- ²¹D. Andrieux and P. Gaspard, "Nonequilibrium generation of information in copolymerization processes," *Proc. Natl. Acad. Sci. U. S. A.* **105**, 9516–9521 (2008).
- ²²A. Blokhuis and D. Lacoste, "Length and sequence relaxation of copolymers under recombination reactions," *J. Chem. Phys.* **147**, 094905 (2017).
- ²³P. Gaspard, "Template-directed growth of copolymers," *Chaos* **30**, 043114 (2020).
- ²⁴R. Rao, D. Lacoste, and M. Esposito, "Glucans monomer-exchange dynamics as an open chemical network," *J. Chem. Phys.* **143**, 244903 (2015).
- ²⁵P. Gaspard, "Stochastic approach to entropy production in chemical chaos," *Chaos* **30**, 113103 (2020).
- ²⁶K. Yoshimura and S. Ito, "Information geometric inequalities of chemical thermodynamics," [arXiv:2005.08444](https://arxiv.org/abs/2005.08444) [cond-mat.stat-mech] (2020).
- ²⁷J. H. Fritz, B. Nguyen, and U. Seifert, "Stochastic thermodynamics of chemical reactions coupled to finite reservoirs: A case study for the Brusselator," *J. Chem. Phys.* **152**, 235101 (2020).
- ²⁸A. Blokhuis, D. Lacoste, and P. Gaspard, "Reaction kinetics in open reactors and serial transfers between closed reactors," *J. Chem. Phys.* **148**, 144902 (2018).
- ²⁹R. Holyst and A. Poniewierski, *Thermodynamics for Chemists, Physicists and Engineers* (Springer, 2012).
- ³⁰M. Boudart, *Kinetics of Chemical Processes* (Prentice Hall, 1968).
- ³¹K. J. Laidler, *Chemical Kinetics* (HarperCollins Publishers, 1987).
- ³²J. B. Butt, *Reaction Kinetics and Reactor Design* (Marcel Dekker, Inc., 1999).
- ³³D. Voet and J. G. Voet, *Biochemistry* (John Wiley & Sons, 2010).
- ³⁴A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, 2nd ed. (Wiley, 2001).
- ³⁵G. Svehla, "Nomenclature of kinetic methods of analysis (IUPAC Recommendations 1993)," *Pure Appl. Chem.* **65**, 2291 (1993).
- ³⁶M. Pekař, "Thermodynamics and foundations of mass-action kinetics," *Prog. React. Kinet. Mech.* **30**, 3–113 (2005).
- ³⁷K. Kinoshita, R. Yasuda, H. Noji, and K. Adachi, "A rotary molecular motor that can work at near 100% efficiency," *Philos. Trans. R. Soc. London, Ser. B* **355**, 473–489 (2000).
- ³⁸R. Rao and M. Esposito, "Conservation laws shape dissipation," *New J. Phys.* **20**, 023007 (2018).
- ³⁹S. Schuster and R. Schuster, "A generalization of Wegscheider's condition. Implications for properties of steady states and for quasi-steady-state approximation," *J. Math. Chem.* **3**, 25–42 (1989).
- ⁴⁰F. Horn and R. Jackson, "General mass action kinetics," *Arch. Ration. Mech. Anal.* **47**, 81–116 (1972).
- ⁴¹D. McQuarrie, *Statistical Mechanics* (Harper & Row, 1975).
- ⁴²C. Chipot and A. Pohorille, *Free Energy Calculations* (Springer, Berlin, 2007).
- ⁴³T. Lelièvre, M. Rousset, and G. Stoltz, *Free Energy Computations: A Mathematical Perspective* (Imperial College Press, 2010).
- ⁴⁴R. Holyst and A. Poniewierski, *Thermodynamics for Chemists, Physicists and Engineers* (Springer, Dordrecht, The Netherlands, 2012).
- ⁴⁵G. Magnus and J. Keizer, "Minimal model of beta-cell mitochondrial Ca^{2+} handling," *Am. J. Physiol.: Cell Physiol.* **273**, C717–C733 (1997).
- ⁴⁶G. Magnus and J. Keizer, "Model of β -cell mitochondrial calcium handling and electrical activity. I. Cytoplasmic variables," *Am. J. Physiol.: Cell Physiol.* **274**, C1158–C1173 (1998).
- ⁴⁷C. P. Fall and J. E. Keizer, "Mitochondrial modulation of intracellular Ca^{2+} signaling," *J. Theor. Biol.* **210**, 151–165 (2001).
- ⁴⁸A. Wachtel, R. Rao, and M. Esposito, "Thermodynamically consistent coarse graining of biocatalysts beyond Michaelis-Menten," *New J. Phys.* **20**, 042002 (2018).
- ⁴⁹F. Avanzini, G. Falasco, and M. Esposito, "Thermodynamics of non-elementary chemical reaction networks," *New J. Phys.* **22**, 093040 (2020).

The following article is reprinted from
[E. PENOCCHIO, R. RAO and M. ESPOSITO, *Nat. Commun.* **10**. (2019), 3865]
under the conditions of the Creative Commons CC BY 4.0 license².

The page numbers placed in the outer margins provide a continuous pagination throughout the thesis.

² <https://creativecommons.org/licenses/by/4.0/>

ARTICLE

<https://doi.org/10.1038/s41467-019-11676-x>

OPEN

Thermodynamic efficiency in dissipative chemistry

Emanuele Penocchio ¹, Riccardo Rao ^{1,2} & Massimiliano Esposito¹

Chemical processes in closed systems inevitably relax to equilibrium. Living systems avoid this fate and give rise to a much richer diversity of phenomena by operating under nonequilibrium conditions. Recent experiments in dissipative self-assembly also demonstrated that by opening reaction vessels and steering certain concentrations, an ocean of opportunities for artificial synthesis and energy storage emerges. To navigate it, thermodynamic notions of energy, work and dissipation must be established for these open chemical systems. Here, we do so by building upon recent theoretical advances in nonequilibrium statistical physics. As a central outcome, we show how to quantify the efficiency of such chemical operations and lay the foundation for performance analysis of any dissipative chemical process.

¹Complex Systems and Statistical Mechanics, Physics and Materials Science Research Unit, University of Luxembourg, L-1511 Luxembourg, Luxembourg.

²Present address: The Simons Center for Systems Biology, School of Natural Sciences, Institute for Advanced Study, Princeton, NJ 08540, USA.

Correspondence and requests for materials should be addressed to M.E. (email: massimiliano.esposito@uni.lu)

Traditional chemical thermodynamics deals with closed systems, which always evolve towards equilibrium. At equilibrium, all reaction currents—defined as the forward reaction fluxes minus the backwards ($J_\rho = J_{+\rho} - J_{-\rho}$, where ρ labels the reactions)—eventually vanish. The first thermodynamic description of nonequilibrium chemical processes was achieved by the Brussels school founded by de Donder and perpetuated by Prigogine^{1,2}, but they focused on few reactions close to equilibrium in the so-called linear regime. However, processes such as fuel-driven self-assembly involve open chemical reaction networks (CRN) with many reactions operating far away from equilibrium^{3,4}. The openness arises from the presence of one or more chemostats, i.e. particle reservoirs coupled with the system which externally control the concentrations of some species—just like thermostats control temperatures—and allow for matter exchanges. Open CRN can then be thought of as thermodynamic machines powered by chemostats. Two operating regimes may be distinguished, reminiscent of stroke and steady-state engines. In the first, work is used to induce a time-dependent change in the species abundances that could never be reached at equilibrium. An example could be the accumulation of a large amount of molecules with a high free energy content as in fuel-driven self-assembly, or the depletion of some undesired species as in metabolite repair⁵. In the second, work is used to maintain the system in a nonequilibrium stationary state which continuously transduces an input work into useful output work. Beyond energy transduction within pseudo-first order reactions⁶, no framework currently exists to assess how efficient and powerful such chemical engines can be. We provide one grounded in the recently established nonequilibrium thermodynamics of CRN^{7,8}, which was born from the combination of state-of-the-art statistical mechanics^{9–14} and mathematical CRN theory^{15,16}. Establishing rigorous concepts of free energy, chemical work and dissipation valid far from equilibrium reveals crucial. They provide the basis for thermodynamically meaningful definitions of efficiencies and optimal performance in the different operating regimes. In the following, energy storage (ES) and driven synthesis (DS) are analyzed as models epitomizing the first and the second operating regime, respectively, but our findings apply to any dissipative chemical process.

Results

Energy storage. In energy storage, an open CRN initially at equilibrium with high concentrations of low-energy molecules and low concentrations of high-energy ones is brought out of equilibrium with the aim to increase the concentrations of the high-energy species. This process is reminiscent of charging a capacitor via the coupling to a voltage generator. In the context of supramolecular chemistry, the concept of ES was proposed by Ragazzon and Prins⁴. An insightful model capturing its main features is described in Fig. 1. Its thermodynamic analysis, detailed in Supplementary Note 1b, will now be outlined. Given a set of reaction rate constants, the accumulation of the high-energy species A_2 may be enabled when chemostats set a certain positive chemical potential difference of fuel and waste, i.e. $\mathcal{F}_{\text{fuel}} = \mu_F - \mu_W > 0$, by steering [F] (see Supplementary Fig. 1). This implies the injection of F molecules at a rate I_F and the extraction of W at rate I_W . The resulting power (i.e., work per unit of time) performed on the system by the fueling mechanism is $\dot{W}_{\text{fuel}} = I_F \mathcal{F}_{\text{fuel}}$ ^{7,8,17}. The proper way to quantify the energy content of an open CRN is via its nonequilibrium free energy \mathcal{G} . During the charging process, only part of the work, namely $\Delta\mathcal{G}$, is dedicated to shift the concentrations distribution and is stored as free energy in the system⁴. The remaining fraction, namely $T\Sigma$, is dissipated according to the

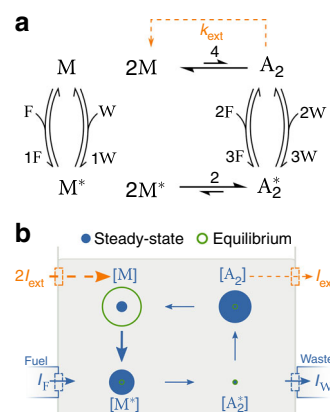


Fig. 1 Model for energy storage and driven synthesis. Without (resp. with) the orange dashed transition, the chemical reaction network models energy storage (resp. driven synthesis). The high-energy species A_2 is at low concentration at equilibrium. Powering the system by chemostating fuel (F) and waste (W) species boosts the formation of A_2 out of the monomer M via the activated species M_2 and A_2^* . **a** The chemical reaction network (forward fluxes are defined counter-clockwise). **b** Sketch of concentrations distributions (proportional to radii) and net currents (proportional to arrows thickness, see Supplementary Note 1c)

second law of thermodynamics

$$\mathcal{W}_{\text{fuel}} = \Delta\mathcal{G} + T\Sigma, \quad (1)$$

where T is temperature and $\Sigma \geq 0$ the entropy production, which only vanishes at equilibrium. The time-dependent thermodynamic efficiency of an ES process is thus the ratio

$$\eta_{\text{es}} = \frac{\Delta\mathcal{G}}{\mathcal{W}_{\text{fuel}}} = 1 - \frac{T\Sigma}{\mathcal{W}_{\text{fuel}}}. \quad (2)$$

Equation (1) has been used to derive the second equality. We emphasize that each of these contributions has an explicit expression in terms of concentrations and rate constants (see Supplementary Note 1b). For instance, the energy stored at any time with respect to equilibrium is given by the expression

$$\Delta\mathcal{G} = RT \sum_{X=M, M^*, A_2^*, A_2} \left[[X] \ln \frac{[X]}{[X]_{\text{eq}}} - [X] + [X]_{\text{eq}} \right] \geq 0, \quad (3)$$

which is reminiscent of an information theoretical measure called relative entropy¹⁸. Crucially, any concentration distribution different from the equilibrium one has a positive energy content. Equation (1) thus implies that an amount of work of at least $\Delta\mathcal{G}$ needs to be provided to reach it. It also ensures that η_{es} is bounded between zero and one.

We simulated an ES process and plotted the dynamics of concentrations as well as efficiency and its contributions in Fig. 2. The process can be divided into a charging and a maintenance phase. During the former, the system energy grows ($d_t \mathcal{G} > 0$) in a way which correlates with the accumulation of the high-energy species A_2 . The process can be quite efficient as a significant portion of the work is converted into free energy. However, in the maintenance phase, the system has reached a nonequilibrium steady state. The efficiency drops towards zero (proportional to the inverse time) as the entire work is being spent to preserve the energy previously accumulated ($d_t \mathcal{G} \simeq 0$). The maximum η_{es} is reached during the charging phase (see Supplementary Note 1b for a rigorous proof) and defines

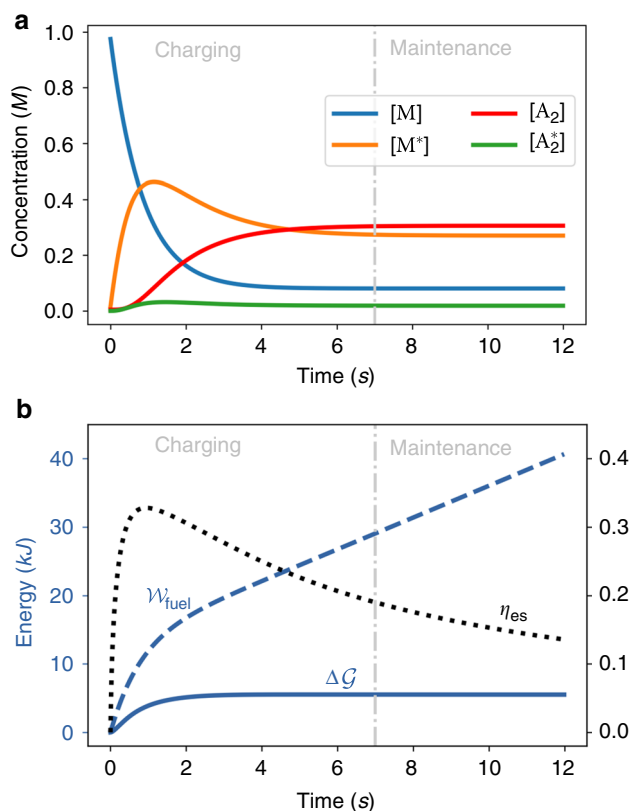


Fig. 2 Dynamics of energy storage. The system is initially prepared at thermodynamic equilibrium where $[M]_{\text{eq}} \gg [A_2]_{\text{eq}}$. At time $t = 0$, the chemical potential difference between fuel and waste is turned on at $\mathcal{F}_{\text{fuel}} = 7.5 \cdot RT$ and drives the system away from equilibrium. After a transient (charging phase), the system eventually settles into a nonequilibrium steady state (maintenance phase). **a** Species abundances. **b** Energy stored, work and efficiency (right axis, adimensional units). Kinetic constants ($\{k_{sp}\}$) and chemical potentials used for simulations are given in Supplementary Table 1

the time that minimizes the dissipation of ES. The value of the efficiency when the process enters the maintenance phase characterizes instead the performance of the ES process when the system has reached its maximum storage capacity. The best time to stop ES and start making use of it (cf. driven synthesis below) will be a tradeoff between maximizing the energy stored and minimizing dissipation. In general, the ideal situation will be the one in which η_{es} peaks as close as possible to the maintenance phase.

Figure 3 focuses on the maintenance phase for different values of $\mathcal{F}_{\text{fuel}}$. It shows that by driving the system away from equilibrium, one can reach species abundances that are very different with respect to the equilibrium ones. It also shows that the accumulation of free energy does not necessarily coincide with an increase in concentration of the most energetic species A_2 . Indeed, while at low values of $\mathcal{F}_{\text{fuel}}$ the accumulation of \mathcal{G} correlates with $[A_2]$, beyond a threshold A_2 starts to be depleted while energy continues getting stored by further moving away the concentration distribution from equilibrium. We finally note that the connection of our work to “kinetic asymmetry”^{4,19} is discussed in Supplementary Note 1c.

As we have seen, the crucial part of energy storage is the charging phase, as the maintenance phase is purely dissipative and consumes chemical work without any energy gain. In order to make use of the energy accumulated during the charging phase, a mechanism extracting the energetic species from the system

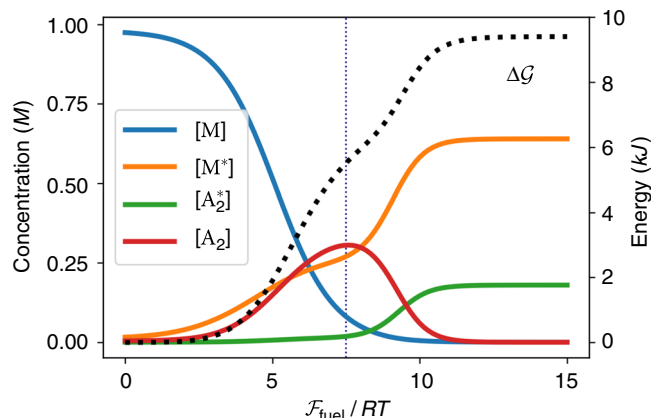


Fig. 3 Maintenance phase of energy storage. Stationary concentrations and free energy difference from equilibrium in the maintenance phase of energy storage as a function of the chemical potential difference between fuel and waste. The vertical dotted line denotes the value $\mathcal{F}_{\text{fuel}} = 7.5 \cdot RT$ used to study the charging phase in Fig. 2

must be introduced. This complementary but distinct working regime of an open CRN will now be considered.

Driven synthesis. In driven synthesis, an energetic species that accumulates thanks to a fueling process is continuously extracted from a system in a nonequilibrium steady state. One may consider for instance processes where the product either evaporates, precipitates or undergoes other fast transformations while being rapidly replaced by reactants. By building upon the above ES scheme, a simple way to model DS is to add an ideal extraction/injection mechanism to the CRN (orange dashed arrows in Fig. 1). This mechanism removes the assembled molecule A_2 and renews two M molecules at a rate $I_{\text{ext}} = k_{\text{ext}}[A_2]$. In doing so, we model the endergonic synthesis of molecules that are strongly unfavored at equilibrium, a strategy used by Nature^{20–22} and which may be within reach of supramolecular chemists^{23–25}.

From the thermodynamic standpoint detailed in Supplementary Note 2b, the input power spent by the fueling mechanism, $\dot{W}_{\text{fuel}} = I_{\text{F}} \mathcal{F}_{\text{fuel}} = I_{\text{F}}(\mu_{\text{F}} - \mu_{\text{W}})$, is now not just dissipated as $T\dot{\Sigma}$, but part of it is used to sustain the production of A_2 :

$$\dot{W}_{\text{fuel}} = -\dot{W}_{\text{ext}} + T\dot{\Sigma}. \quad (4)$$

The output power released by the extraction mechanism, $\dot{W}_{\text{ext}} = I_{\text{ext}}(2\mu_{\text{M}} - \mu_{\text{A}_2})$, is negative when DS occurs. In this context the thermodynamic efficiency is thus given by

$$\eta_{\text{ds}} = -\frac{\dot{W}_{\text{ext}}}{\dot{W}_{\text{fuel}}} = 1 - \frac{T\dot{\Sigma}}{\dot{W}_{\text{fuel}}}, \quad (5)$$

where Eq. (4) has been used to derive the second equality. It is bounded between zero and one when DS occurs.

In Fig. 4, we simulated DS for various working conditions by varying k_{ext} and $\mathcal{F}_{\text{fuel}}$. We start our analysis by considering a given value of $\mathcal{F}_{\text{fuel}}$. As k_{ext} is increased, η_{ds} first grows to an optimal value before decreasing towards negative values where the DS regime ends (see Fig. 4a). At the same time I_{ext} increases until it reaches a plateau (see Fig. 4c). This happens when k_{ext} overcomes the ability of the system to sustain high values of $[A_2]$ (Fig. 4b). Eventually the drop in $[A_2]$ is such that $2\mu_{\text{M}} - \mu_{\text{A}_2} > 0$, thus resulting in the loss of the DS regime. We now fix k_{ext} and increase $\mathcal{F}_{\text{fuel}}$. The DS regime starts at a threshold value, when $[A_2]$ becomes high enough. After that, both $[A_2]$ and the efficiency grow to an optimal value before decreasing again. This time however, the efficiency remains positive as $[M]$ drops

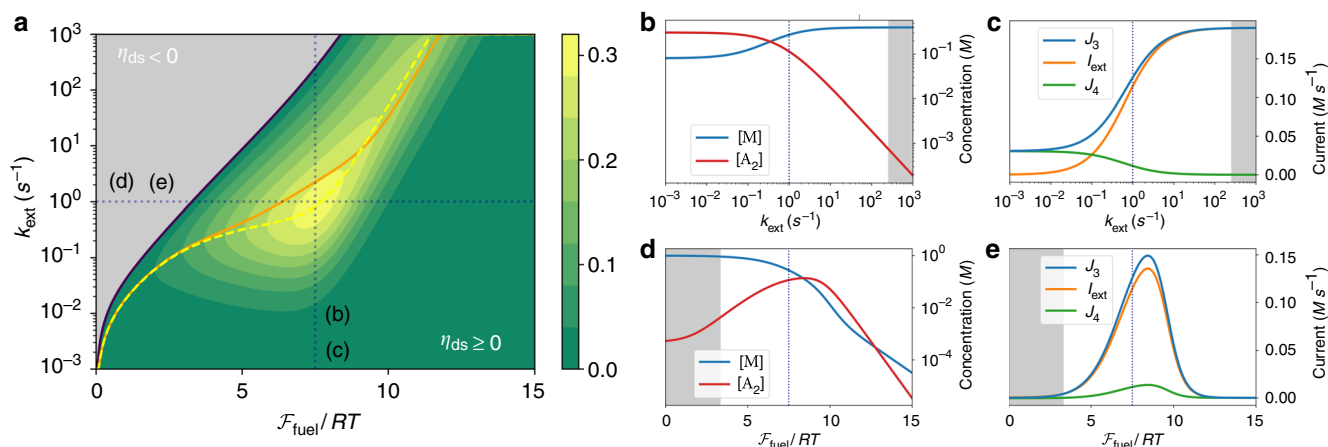


Fig. 4 Performance of driven synthesis. **a** Efficiency (η_{ds}) of driven synthesis as function of \mathcal{F}_{fuel} and k_{ext} . Regions of operating regimes that do not perform driven synthesis are colored in gray. The yellow dashed (resp. orange solid) line denotes the maximum of η_{ds} (resp. $-W_{ext}$, see also Supplementary Fig. 2) versus k_{ext} , while the black solid line corresponds to $\eta_{ds} = 0$. **b, c** (resp. **d, e**) Concentrations and currents as a function of k_{ext} (resp. \mathcal{F}_{fuel}) along the line $\mathcal{F}_{fuel} = 7.5 \cdot RT$ (resp. $k_{ext} = 1$ s $^{-1}$) highlighted by blue vertical (resp. horizontal) dotted line in plot (**a**). Kinetic constants and standard chemical potentials are the same as for ES analysis (see Supplementary Information). Note that $J_3 = I_{ext} + J_4$ always holds, where $J_3 = J_{3F} + J_{3W}$ is the net current flowing from A_2^* to A_2 .

together with $[A_2]$ (see Fig. 4d). Figure 4e shows another important feature. As \mathcal{F}_{fuel} is increased, I_{ext} first increases too, but eventually reaches a maximum after which it decreases. This phenomenon is a hallmark of far-from-equilibrium physics which could not happen in a linear regime, namely when k_{ext} and \mathcal{F}_{fuel} are small. Remarkably, the global maximum of the efficiency in Fig. 4a is reached in a region far from equilibrium. We note that it corresponds to values of \mathcal{F}_{fuel} close to the one maximizing $[A_2]$ in the maintenance phase of ES (see Fig. 3) and to values of k_{ext} of order one resulting in I_{ext} which do not overly deplete $[A_2]$. We finally turn to the lines of maximum efficiency and efficiency at maximum power in Fig. 4a, where the maximization is done with respect to k_{ext} at a given \mathcal{F}_{fuel} . Since these two lines typically do not coincide, the study of the tradeoffs is the object of a rich field called finite-time thermodynamics²⁶. Interestingly, while these two lines cannot coincide in the linear regime (see Supplementary Note 2d), we see that they do intersect far-from-equilibrium, not far from the global maximum of the efficiency. Our analysis thus allowed us to identify a region of good tradeoff between power and efficiency for the model of DS we introduced. In order to emphasize the fact that all the interesting features that we identified in DS occur far-from-equilibrium, we analyze in detail in Supplementary Note 2d the linear regime of DS. After identifying the Onsager matrix, we are able to analytically reproduce the results of the simulations in the limit of small \mathcal{F}_{fuel} and k_{ext} (bottom-left part of Fig. 4a, see Supplementary Fig. 3 for details), thus pinpointing the limit of validity of the linear regime approximation.

Discussion

Thermodynamics was born from the effort to systematize the performance of steam engines. Open CRN, which are at the core of recent efforts in artificial synthesis²⁷ and ubiquitous in living systems^{22,28,29}, can be seen as chemical engines. In the spirit of this analogy, in this article we built a chemical thermodynamic framework which enables us to systematically analyze the performance of two fundamental dissipative chemical processes. The first, energy storage, is concerned with the time-dependent accumulation of high-energy species far from equilibrium and is currently raising significant attention from supramolecular chemists. The second, driven synthesis, aims at continuously extracting the

previously obtained high-energy species and provides a simple and insightful instance of energy transduction beyond pseudo-unimolecular CRN. In doing so, we identified their optimal regimes of operation. Crucially they lie far-from-equilibrium in regions unreachable using conventional linear regime thermodynamics. We emphasize that the methods developed in this paper can in principle be applied to any open CRN and thus provide the basis for future performance studies and optimal design of dissipative chemistry. They are thus destined to play a major role in bioengineering and nanotechnologies.

Data availability

All data needed to reproduce numerical results are reported in the Supplementary Information.

Code availability

The code that generated the plots is available from the corresponding author upon request.

Received: 18 March 2019 Accepted: 18 July 2019

Published online: 27 August 2019

References

- Prigogine, I. & Defay, R. *Chemical Thermodynamics* (Longmans, Green & Co., London, 1954).
- Prigogine, I. *Introduction to Thermodynamics of Irreversible Processes* (John Wiley & Sons, New York, 1967).
- van Rossum, S. A. P., Tena-Solsona, M., van Esch, J. H., Eelkema, R. & Boekhoven, J. Dissipative out-of-equilibrium assembly of man-made supramolecular materials. *Chem. Soc. Rev.* **46**, 5519–5535 (2017).
- Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotechnol.* **13**, 882–889 (2018).
- Linster, C. L., Schaftingen, E. Van & Hanson, A. D. Metabolite damage and its repair or pre-emption. *Nat. Chem. Biol.* **9**, 72–80 (2013).
- Hill, T. L. *Free Energy Transduction in Biology* (Academic Press, New York, 1977).
- Rao, R. & Esposito, M. Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. *Phys. Rev. X* **6**, 041064 (2016).
- Falasco, G., Rao, R. & Esposito, M. Information thermodynamics of turing patterns. *Phys. Rev. Lett.* **121**, 108301 (2018).
- Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **75**, 126001 (2012).

10. Ciliberto, S. Experiments in stochastic thermodynamics: short history and perspectives. *Phys. Rev. X* **7**, 021051 (2017).
11. Jarzynski, C. Equalities and inequalities: irreversibility and the second law of thermodynamics at the nanoscale. *Annu. Rev. Condens. Matter Phys.* **2**, 329–351 (2011).
12. Zhang, X.-J., Qian, H. & Qian, M. Stochastic theory of nonequilibrium steady states and its applications. Part i. *Phys. Rep.* **510**, 1–86 (2012).
13. Ge, H., Qian, M. & Qian, H. Stochastic theory of nonequilibrium steady states and its applications. Part ii: applications in chemical biophysics. *Phys. Rep.* **510**, 87–118 (2012).
14. Van den Broeck, C. & Esposito, M. Ensemble and trajectory thermodynamics: a brief introduction. *Phys. A* **418**, 6–16 (2015).
15. Horn, F. & Jackson, R. General mass action kinetics. *Arch. Ration. Mech. An.* **47**, 81–116 (1972).
16. Feinberg, M. Complex balancing in general kinetic systems. *Arch. Ration. Mech. An.* **49**, 187–194 (1972).
17. Rao, R. & Esposito, M. Conservation laws and work fluctuation relations in chemical reaction networks. *J. Chem. Phys.* **149**, 245101 (2018).
18. Cover, T. M. & Thomas, J. A. *Elements of Information Theory* (John Wiley & Sons, New York, 2006).
19. Astumian, R. D. Stochastic pumping of non-equilibrium steady-states: how molecules adapt to a fluctuating environment. *ChemComm.* **54**, 427–444 (2018).
20. Desai, A. & Mitchison, T. J. Microtubule polymerization dynamics. *Annu. Rev. Cell. Dev. Biol.* **13**, 83–117 (1997).
21. Howard, J. *Mechanics of Motor Proteins and the Cytoskeleton* (Sinauer Associates, Sunderland, MA, 2001).
22. Hess, H. & Ross, J. L. Non-equilibrium assembly of microtubules: from molecules to autonomous chemical robots. *Chem. Soc. Rev.* **46**, 5570–5587 (2017).
23. Boekhoven, J. et al. Dissipative self-assembly of a molecular gelator by using a chemical fuel. *Angew. Chem.* **122**, 4935–4938 (2010).
24. Boekhoven, J., Hendriksen, W. E., Koper, G. J. M., Eelkema, R. & van Esch, J. H. Transient assembly of active materials fueled by a chemical reaction. *Science* **349**, 1075–1079 (2015).
25. Sorrenti, A., Leira-Iglesias, J., Markvoort, A. J., de Greef, T. F. A. & Hermans, T. M. Non-equilibrium supramolecular polymerization. *Chem. Soc. Rev.* **46**, 5476–5490 (2017).
26. Benenti, G., Casati, G., Saito, K. & Whitney, R. S. Fundamental aspects of steady-state conversion of heat to work at the nanoscale. *Phys. Rep.* **694**, 1–124 (2017).
27. Mattia, E. & Otto, S. Supramolecular systems chemistry. *Nat. Nanotechnol.* **10**, 111 (2015).
28. Zwaag, Dvander & Meijer, E. W. Fueling connections between chemistry and biology. *Science* **349**, 1056–1057 (2015).
29. Grzybowski, B. A. & Huck, W. T. S. The nanotechnology of life-inspired systems. *Nat. Nanotechnol.* **11**, 585 (2016).

Acknowledgements

This work was funded by the Luxembourg National Research Fund (AFR Ph.D. Grant 2014-2, No. 9114110) and the European Research Council project NanoThermo (ERC-2015-CoG Agreement No. 681456).

Author contributions

E.P., R.R., and M.E. all significantly contributed to conceive and realize the project as well as in writing the paper.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-019-11676-x>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Peer review information: *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019

Supplementary Information for “Thermodynamic Efficiency in Dissipative Chemistry”

Emanuele Penocchio,¹ Riccardo Rao,^{1,2} and Massimiliano Esposito^{1,*}

¹*Complex Systems and Statistical Mechanics, Physics and Materials Science
Research Unit, University of Luxembourg, L-1511 Luxembourg, G.D. Luxembourg*

²*Present Address: The Simons Center for Systems Biology, School of Natural
Sciences, Institute for Advanced Study, Princeton, 08540 New Jersey, U.S.A.*

CONTENTS

Supplementary Note 1: Details on Energy Storage	1
a. Dynamics	1
b. Thermodynamics	2
c. Cycles & kinetic symmetry	5
d. Parameters	6
Supplementary Note 2: Details on Driven Synthesis	7
a. Dynamics	7
b. Thermodynamics	7
c. Plots of $-\dot{W}_{\text{ext}}$ and \dot{W}_{fuel}	8
d. Linear Regime	8
Supplementary References	11

SUPPLEMENTARY NOTE 1: DETAILS ON ENERGY STORAGE

a. Dynamics

The evolution in time of the concentrations of the species M , M^* , A_2^* , and A_2 is ruled by the rate equations

$$\underbrace{d_t \begin{pmatrix} [M] \\ [M^*] \\ [A_2^*] \\ [A_2] \end{pmatrix}}_{[X]} = \underbrace{\begin{pmatrix} -1 & -1 & 0 & 0 & 0 & 2 \\ 1 & 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{pmatrix}}_{\mathbb{S}^X} \cdot \underbrace{\begin{pmatrix} k_{+1F}[F][M] - k_{-1F}[M^*] \\ k_{+1W}[W][M] - k_{-1W}[M^*] \\ k_{+2}[M^*]^2 - k_{-2}[A_2^*] \\ k_{+3F}[A_2^*] - k_{-1F}[A_2][F]^2 \\ k_{+3W}[A_2^*] - k_{-1W}[A_2][W]^2 \\ k_{+4}[A_2] - k_{-4}[M]^2 \end{pmatrix}}_{J = J_+ - J_-}, \quad (1)$$

* massimiliano.esposito@uni.lu

where $[F]$ and $[W]$ are the concentrations of fuel and waste species. Since these latter are externally kept constant by the chemostats, the balance equations for their concentrations read

$$\mathbf{0} = d_t \underbrace{\begin{pmatrix} [F] \\ [W] \end{pmatrix}}_{[Y]} = \underbrace{\begin{pmatrix} -1 & 0 & 0 & 2 & 0 & 0 \\ 0 & -1 & 0 & 0 & 2 & 0 \end{pmatrix}}_{\mathbb{S}^Y} \cdot \underbrace{\begin{pmatrix} k_{+1F}[F][M] - k_{-1F}[M^*] \\ k_{+1W}[W][M] - k_{-1W}[M^*] \\ k_{+2}[M^*]^2 - k_{-2}[A_2^*] \\ k_{+3F}[A_2^*] - k_{-1F}[A_2][F]^2 \\ k_{+3W}[A_2^*] - k_{-1W}[A_2][W]^2 \\ k_{+4}[A_2] - k_{-4}[M]^2 \end{pmatrix}}_{\mathbf{J} = \mathbf{J}_+ - \mathbf{J}_-} + \underbrace{\begin{pmatrix} I_F \\ I_W \end{pmatrix}}_{\mathbf{I}}, \quad (2)$$

with I_F and I_W denoting the external currents of fuel and waste flowing from the chemostats. We denote by $X = M, M^*, A_2, A_2^*$ the internal species, by $Y = F, W$ the chemostatted ones, and label by $\rho = 1F, 1W, 2, 3F, 3W, 4$ the reactions.

b. Thermodynamics

We consider an isothermal, isobaric, and well-stirred ideal dilute solution containing species undergoing elementary reactions. Each species is thermodynamically characterized by chemical potentials of the form

$$\mu_X = \mu_X^\circ + RT \ln \frac{[X]}{[0]}, \quad \mu_Y = \mu_Y^\circ + RT \ln \frac{[Y]}{[0]}, \quad (3)$$

where μ_X° and μ_Y° are standard-state chemical potentials and $[0]$ is the standard-state concentration.

Dynamics and thermodynamics are related via the hypothesis of local detailed balance, which relates the ratio of rate constants to the differences of standard-state chemical potentials along reactions

$$RT \ln \frac{k_{+\rho}}{k_{-\rho}} = - \sum_X \mu_X^\circ \mathbb{S}_\rho^X - \sum_Y \mu_Y^\circ \mathbb{S}_\rho^Y. \quad (4)$$

At equilibrium, the thermodynamic forces driving each reaction, also called affinities, vanish

$$A_\rho^{\text{eq}} = - \sum_X \mu_X^{\text{eq}} \mathbb{S}_\rho^X - \sum_Y \mu_Y^{\text{eq}} \mathbb{S}_\rho^Y = 0, \quad (5)$$

as well as all reaction currents

$$J_\rho^{\text{eq}} = J_{+\rho}^{\text{eq}} - J_{-\rho}^{\text{eq}} = 0. \quad (6)$$

The dissipation of the process is captured by the entropy production (EP) rate

$$T \dot{\Sigma} = RT \sum_\rho J_\rho \ln \frac{J_{+\rho}}{J_{-\rho}} \geq 0, \quad (7)$$

which also vanishes at equilibrium. Using the rate equations and the local detailed balance, Supplementary Equation 4, one can rewrite this quantity as

$$T \dot{\Sigma} = -d_t G + \dot{\mathcal{W}}_{\text{chem}}, \quad (8)$$

where

$$G = \sum_X [X] (\mu_X - RT) + \sum_Y [Y] (\mu_Y - RT) \quad (9)$$

3

is the Gibbs free energy, while

$$\dot{\mathcal{W}}_{\text{chem}} = \sum_Y \mu_Y I_Y = \mu_F I_F + \mu_W I_W \quad (10)$$

is the chemical work per unit time exchanged with the chemostats.

One can also show that if the CRN were closed (fuel and waste not chemostatted) it would relax to equilibrium by minimizing G [1]. Fuel and waste are however chemostatted and we need to identify the conditions for equilibrium in the open CRN. To do so we preliminary identify the topological properties of the network.

The stoichiometric matrix $\mathbb{S} \equiv (\mathbb{S}^X, \mathbb{S}^Y)^T$ (see Supplementary Equations 1 and 2) encodes the topological properties of the CRN. We can access these properties by determining its cokernel, which is spanned by

$$\boldsymbol{\ell}_M = \begin{pmatrix} & \text{M} & \text{M}^* & \text{A}_2^* & \text{A}_2 & \text{F} & \text{W} \\ 1 & 1 & 2 & 2 & 0 & 0 \end{pmatrix}, \quad (11)$$

$$\boldsymbol{\ell}_W = \begin{pmatrix} & \text{M} & \text{M}^* & \text{A}_2^* & \text{A}_2 & \text{F} & \text{W} \\ 0 & 1 & 2 & 0 & 1 & 1 \end{pmatrix}. \quad (12)$$

The first of these vectors identifies a conserved quantity

$$\begin{aligned} L_M &= \boldsymbol{\ell}_L \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [M] + [M^*] + 2[A_2^*] + 2[A_2], \\ d_t L_M &= 0 \end{aligned} \quad (13)$$

which is proved using the rate equations Supplementary Equation 1 and Supplementary Equation 2.

The second vector identifies what we call a broken conserved quantity

$$L_W = \boldsymbol{\ell}_W \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [M^*] + 2[A_2^*] + [F] + [W]. \quad (14)$$

Using again the rate equations, it can be shown that

$$d_t L_W := I_F + I_W. \quad (15)$$

Namely, L_W changes only due to the exchange of fuel and waste with the chemostats. If the CRN were closed, L_W would be constant. Using Supplementary Equation 15, we can rewrite the entropy production in Supplementary Equation 8 as

$$T\dot{\Sigma} = -d_t \mathcal{G} + \dot{\mathcal{W}}_{\text{fuel}}, \quad (16)$$

where

$$\begin{aligned} \mathcal{G} &= \sum_X [X] (\mu_X - RT) + \sum_Y [Y] (\mu_Y - RT) - \mu_W L_W \\ &= [M] \mu_M + [A_2] \mu_{A_2} + [M^*] (\mu_{M^*} - \mu_W) + [A_2^*] (\mu_{A_2^*} - 2\mu_W) + [F] (\mu_F - \mu_W) + \\ &\quad - RT ([M] + [A_2] + [M^*] + [A_2^*] + [F] + [W]) \end{aligned} \quad (17)$$

is a semigrand Gibbs potential, and

$$\dot{\mathcal{W}}_{\text{fuel}} := I_F (\mu_F - \mu_W). \quad (18)$$

is the fueling chemical work per unit of time (*i.e.*, the fueling power). The derivation of Supplementary Equation 18 for an arbitrary CRN is discussed in Supplementary References [1], [2] and [3].

If $\mu_F = \mu_W$, Supplementary Equation 16 shows that \mathcal{G} is a monotonically decreasing function in time, given that $T\dot{\Sigma} \geq 0$. Its minimum value — *i.e.*, the equilibrium value — under the constraint given by

the conservation law (Supplementary Equation 13) is found by minimizing the function $\Lambda = \mathcal{G} - \lambda L_M$, where λ is the Lagrange multiplier corresponding to L_M . The equilibrium concentrations thus satisfy the following conditions

$$\begin{aligned}
0 &= \left. \frac{d\Lambda}{d[M]} \right|_{\text{eq}} = \mu_M^{\text{eq}} - \lambda = \mu_M^\circ + RT \ln[M]_{\text{eq}} - \lambda, \\
0 &= \left. \frac{d\Lambda}{d[A_2]} \right|_{\text{eq}} = \mu_{A_2}^{\text{eq}} - 2\lambda = \mu_{A_2}^\circ + RT \ln[A_2]_{\text{eq}} - 2\lambda, \\
0 &= \left. \frac{d\Lambda}{d[M^*]} \right|_{\text{eq}} = \mu_{M^*}^{\text{eq}} - \mu_W - \lambda = \mu_{M^*}^\circ + RT \ln[M^*]_{\text{eq}} - \mu_W - \lambda, \\
0 &= \left. \frac{d\Lambda}{d[A_2^*]} \right|_{\text{eq}} = \mu_{A_2^*}^{\text{eq}} - 2\mu_W - 2\lambda = \mu_{A_2^*}^\circ + RT \ln[A_2^*]_{\text{eq}} - 2\mu_W - 2\lambda.
\end{aligned} \tag{19}$$

The equilibrium semigrand Gibbs potential reads

$$\begin{aligned}
\mathcal{G}_{\text{eq}} &= \lambda L_M - RT ([M]_{\text{eq}} + [A_2]_{\text{eq}} + [M^*]_{\text{eq}} + [A_2^*]_{\text{eq}} + [F]_{\text{eq}} + [W]_{\text{eq}}) \\
&= [M] \mu_M^{\text{eq}} + [A_2] \mu_{A_2}^{\text{eq}} + [M^*] (\mu_{M^*}^{\text{eq}} - \mu_W) + [A_2^*] (\mu_{A_2^*}^{\text{eq}} - 2\mu_W) + \\
&\quad - RT ([M]_{\text{eq}} + [A_2]_{\text{eq}} + [M^*]_{\text{eq}} + [A_2^*]_{\text{eq}} + [F]_{\text{eq}} + [W]_{\text{eq}}),
\end{aligned} \tag{20}$$

which leads by direct calculation to Equation (3) in the main text:

$$\mathcal{G} - \mathcal{G}_{\text{eq}} = RT \sum_X \left[[X] \ln \frac{[X]}{[X]_{\text{eq}}} - [X] + [X]_{\text{eq}} \right] \geq 0. \tag{21}$$

Therefore, when $\mu_F = \mu_W$, the quantity $\mathcal{G} - \mathcal{G}_{\text{eq}}$ is a Lyapunov function for the open network relaxing to equilibrium. When $\mathcal{F}_{\text{fuel}} = \mu_F - \mu_W \neq 0$, the fueling chemical work in Supplementary Equation 16 does not vanish, and the system is prevented from reaching equilibrium.

Equation (1) in the main text is obtained by integrating Supplementary Equation 16 from time $t = 0$ to a generic time t . In our simulation of energy storage, we focused on the special case in which the system at time $t = 0$ is at equilibrium ($\mathcal{F}_{\text{fuel}} = 0$).

We end this section by analytically proving that η_{es} defined in the main text goes to zero both in the short and in the long time limits.

In the short time limit $t = \delta t \ll 1$, *i.e.* immediately after tuning on $\mathcal{F}_{\text{fuel}}$ by changing $[F]$, we have

$$\begin{aligned}
\Delta \mathcal{G} &\simeq d_t \mathcal{G}|_0 \delta t + d_t^2 \mathcal{G}|_0 \delta t^2 \\
\mathcal{W}_{\text{fuel}} &\simeq \mathcal{W}_{\text{fuel}}(0) \delta t + d_t \mathcal{W}_{\text{fuel}}|_0 \delta t^2,
\end{aligned} \tag{22}$$

While $\mathcal{W}_{\text{fuel}}(0) \neq 0$, $d_t \mathcal{G}|_0 = 0$. Indeed, by using Supplementary Equation 17, we find that

$$d_t \mathcal{G}|_0 = \sum_X \hat{\mu}_X(0) d_t [X]|_0 = \sum_X \hat{\mu}_X(0) \mathbb{S}^X \mathbf{J}(0), \tag{23}$$

where $\hat{\mu}_X$ is equal to

$$\hat{\mu}_M = \mu_M, \quad \hat{\mu}_{M^*} = \mu_{M^*} - \mu_W, \quad \hat{\mu}_{A_2^*} = \mu_{A_2^*} - 2\mu_W, \quad \hat{\mu}_{A_2} = \mu_{A_2}. \tag{24}$$

At $t = 0$, the concentrations of internal species X as well as their chemical potentials μ_X are at equilibrium. By using Supplementary Equation 19, one readily sees that $\sum_X \hat{\mu}_X^{\text{eq}} \mathbb{S}^X = 0$ for all reactions, and from Supplementary Equation 23 one proves that $d_t \mathcal{G}|_0 = 0$. Therefore

$$\eta_{\text{es}}(\delta t) \simeq \frac{d_t^2 \mathcal{G}|_0 \delta t}{\mathcal{W}_{\text{fuel}}(0)} \tag{25}$$

goes to zero when δt goes to zero.

In the long time limit $t \rightarrow \infty$, the system approaches a steady state in which $d_t \mathcal{G} \rightarrow 0$, and thus $\dot{\mathcal{W}}_{\text{fuel}} \approx \dot{\Sigma} \geq 0$. Therefore, while $\Delta \mathcal{G}$ remains finite, $\mathcal{W}_{\text{fuel}}$ keeps growing, and $\eta_{\text{es}} \rightarrow 0$.

Since in ES $\Delta \mathcal{G} \geq 0$, as announced, η_{es} will start in zero, increase in time, reach a maximum, and then eventually decrease back to zero. The above proof can be generalized to arbitrary chemical reaction networks evolving towards steady states, but η_{es} might have more than one local maximum depending on the underlying dynamics.

c. Cycles & kinetic symmetry

From the thermodynamic point of view we adopted in this work, any steady state other than the equilibrium one has a non null energy content which is quantified through its concentrations distribution according to Supplementary Equation 21 (equation (3) in the main text). A condition referred to as “kinetic symmetry” is central in the literature on ES [4, 5]. This section aims to frame this concept into our theory.

In a CRN, a *cycle* is a reaction pathway which does not alter the internal state of the system. They play an important role at steady state, where chemical currents can only flow along cycles. Any possible cycle is represented by a vector in the kernel of \mathbb{S}^X , which is spanned by

$$\mathbf{c}_1^T = \begin{pmatrix} & 1F & 1W & 2 & 3F & 3W & 4 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (26)$$

$$\mathbf{c}_2^T = \begin{pmatrix} & 1F & 1W & 2 & 3F & 3W & 4 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix}, \quad (27)$$

$$\mathbf{c}_3^T = \begin{pmatrix} & 1F & 1W & 2 & 3F & 3W & 4 \\ 2 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}, \quad (28)$$

where each entry represents the number of times the corresponding reaction has to be performed in order to complete the cycle. The vector of steady-state currents can always be expressed in terms of a complete base of cycles:

$$\mathbf{J} = J^{(c_1)} \mathbf{c}_1 + J^{(c_2)} \mathbf{c}_2 + J^{(c_3)} \mathbf{c}_3 \quad (29)$$

where the coefficients are called *cycle currents* and represent the contribution of each cycle to the total current observed along each reaction. An important thing to note is that the relation $\frac{1}{2}(J_{1F} + J_{1W}) = J_2 = J_{3F} + J_{3W} = J_4$ always holds. It shows that the net current from M to M* has to be twice as much those across other steps at the stationary state, as represented through arrow thickness in Figure 1b of the main text. When the cycle current for a particular cycle is equal to zero, we refer to that cycle as being *stalled* [6, 7].

Kinetic symmetry as defined in Supplementary Reference [5] corresponds to the situation where no accumulation of A_2 occurs in the system, i.e. when no net current from monomers to assemblies can occur. This corresponds to the situation where the cycle c_3 is stalled, i.e. when $J^{(c_3)} \mathbf{c}_3 = 0$ in Supplementary Equation 29. Mathematically it implies

$$\frac{([\text{F}] \cdot k_{+1F} + [\text{W}] \cdot k_{+1W})^2 k_{+2} (k_{+3F} + k_{+3W}) k_{+4}}{(k_{-1F} + k_{-1W})^2 k_{-2} ([\text{F}]^2 \cdot k_{-3F} + [\text{W}]^2 \cdot k_{-3W}) k_{-4}} = 1, \quad (30)$$

which is the same as Equation (2) reported in Supplementary Reference [5], but with the dependence on the chemostatted species made explicit. The above equation has two solutions in [F]. One always exists and corresponds to the equilibrium state, where by definition all the cycles are stalled and no energy is stored in the system. The other one, when physical (it may be negative), corresponds to a

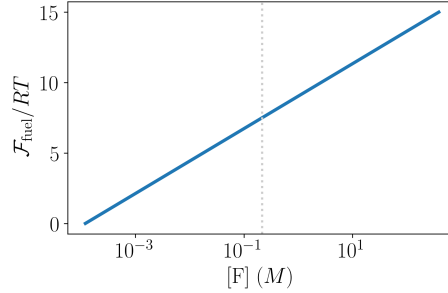
nonequilibrium steady state where A_2 does not accumulate (current will instead flow along cycles c_1 and c_2) but ES nevertheless occurs via an increase of \mathcal{G} . Therefore, ES involving the accumulation of A_2 requires to break the kinetic symmetry of the network, as happens when varying $[F]$. We note that for the choice of kinetic constants and $[W]$ we adopted in the paper (see Supplementary Note 1d below), the condition of nonequilibrium kinetic symmetry can't be realized with any value of $\mathcal{F}_{\text{fuel}}$ different from the equilibrium one.

d. Parameters

With reference to the model in Figure 1 of the main text, the following parameters were used for all the simulations:

Supplementary Table 1. Parameters used for the energy storage model depicted in Figure 1 of the main text. Values of the backward kinetic constants were obtained through Supplementary Equation 4 in order to assure thermodynamic consistency, here they are reported with 3 digits. For the sake of completeness, equilibrium constants of the various reactions ($K_\rho = k_{+\rho}/k_{-\rho}$) are reported. Note that $[W]$ is kept fixed, while we used $[F]$ to tune $\mathcal{F}_{\text{fuel}}$ in the various discussions (see Supplementary Figure 1 below).

	k_{+1F}	$5 \text{ M}^{-1} \text{ s}^{-1}$	
	k_{+1W}	$1 \cdot 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$	
	k_{+2}	$1 \text{ M}^{-1} \text{ s}^{-1}$	
μ_M°	$-2 \cdot 10^3 \text{ J mol}^{-1}$	k_{+3F}	$1 \cdot 10^{-6} \text{ s}^{-1}$
$\mu_{M^*}^\circ$	$-3 \cdot 10^3 \text{ J mol}^{-1}$	k_{+3W}	5 s^{-1}
$\mu_{A_2^*}^\circ$	$-4 \cdot 10^3 \text{ J mol}^{-1}$	k_{+4}	$1 \cdot 10^{-1} \text{ s}^{-1}$
$\mu_{A_2}^\circ$	$9 \cdot 10^3 \text{ J mol}^{-1}$	k_{-1F}	$3.63 \cdot 10^{-2} \text{ s}^{-1}$
μ_F°	$11 \cdot 10^3 \text{ J mol}^{-1}$	k_{-1W}	$6.06 \cdot 10^{-2} \text{ s}^{-1}$
μ_W°	$-11 \cdot 10^3 \text{ J mol}^{-1}$	k_{-2}	2.27 s^{-1}
L_M	1 M	k_{-3F}	$1.74 \text{ M}^{-2} \text{ s}^{-1}$
$[F]$	$[1 \cdot 10^{-4}, 4 \cdot 10^2] \text{ M}$	k_{-3W}	$1.24 \cdot 10^{-1} \text{ M}^{-2} \text{ s}^{-1}$
$[W]$	1 M	k_{-4}	$4.81 \cdot 10^{-4} \text{ M}^{-1} \text{ s}^{-1}$
		K_{1F}	$1.38 \cdot 10^2 \text{ M}^{-1}$
		K_{1W}	$1.65 \cdot 10^{-2} \text{ M}^{-1}$
		K_2	$4.40 \cdot 10^{-1} \text{ M}^{-1}$
		K_{3F}	$5.76 \cdot 10^{-7} \text{ M}^2$
		K_{3W}	$4.02 \cdot 10^1 \text{ M}^2$
		K_4	$2.08 \cdot 10^2 \text{ M}$



Supplementary Figure 1. Values of $\mathcal{F}_{\text{fuel}}$ as a function of $[F]$ (note the logarithmic scale for the x axis). The value of $[F]$ giving $\mathcal{F}_{\text{fuel}} = 7.5 \cdot RT$ is highlighted by the vertical dotted line.

SUPPLEMENTARY NOTE 2: DETAILS ON DRIVEN SYNTHESIS

a. Dynamics

With the addition of the extraction mechanism, the evolution in time of the concentrations of the species M , M^* , A_2 , and A_2^* is ruled by the following rate equations

$$\underbrace{d_t \begin{pmatrix} [M] \\ [M^*] \\ [A_2^*] \\ [A_2] \end{pmatrix}}_{[X]} = \underbrace{\begin{pmatrix} -1 & -1 & 0 & 0 & 0 & 2 \\ 1 & 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{pmatrix}}_{\mathbb{S}^X} \cdot \underbrace{\begin{pmatrix} k_{+1F}[F][M] - k_{-1F}[M^*] \\ k_{+1W}[W][M] - k_{-1W}[M^*] \\ k_{+2}[M^*]^2 - k_{-2}[A_2^*] \\ k_{+3F}[A_2^*] - k_{-3F}[A_2][F]^2 \\ k_{+3W}[A_2^*] - k_{-3W}[A_2][W]^2 \\ k_{+4}[A_2] - k_{-4}[M]^2 \end{pmatrix}}_{J = J_+ - J_-} + \begin{pmatrix} 2I_{\text{ext}} \\ 0 \\ 0 \\ -I_{\text{ext}} \end{pmatrix}, \quad (31)$$

where the current of extraction reads $I_{\text{ext}} = k_{\text{ext}}[A_2]$.

We examine this system at the steady state, in which all concentrations are stationary: $d_t[X]_{\text{ss}} = 0$ for all X . Their expressions are not analytical, but can be easily obtained numerically, thus showing that the steady state state is unique within a broad range of values for the parameters that we examined.

b. Thermodynamics

For the driven synthesis model at the steady state, Supplementary Equation 8 becomes

$$T\dot{\Sigma} = \dot{\mathcal{W}}_{\text{chem}}, \quad (32)$$

where the chemical work per unit of time now reads

$$\dot{\mathcal{W}}_{\text{chem}} = \mu_F I_F + \mu_W I_W + 2\mu_M I_{\text{ext}} - \mu_{A_2} I_{\text{ext}}. \quad (33)$$

In order to construct the entropy balance as in Equation (4) of the main text, we once again need to consider conservation vectors (11) and (12), *i.e.* a basis of the cokernel of \mathbb{S} .

$$\ell_M = \begin{pmatrix} M & M^* & A_2^* & A_2 & F & W \\ 1 & 1 & 2 & 2 & 0 & 0 \end{pmatrix}, \quad (34)$$

$$\ell_W = \begin{pmatrix} M & M^* & A_2^* & A_2 & F & W \\ 0 & 1 & 2 & 0 & 1 & 1 \end{pmatrix}. \quad (35)$$

Now, both these vectors identify broken conserved quantities. The former corresponds to the conserved quantity relative to the monomer

$$L_M = \ell_M \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [M] + [M^*] + 2[A_2^*] + 2[A_2]. \quad (36)$$

In the framework of Supplementary Reference [1], this is a broken conservation law because of the presence of the extraction mechanism. Here its value does not change by construction of the model, since every A_2 which is exchanged is readily replaced by 2 M molecules

$$d_t L_M = 2I_{\text{ext}} - 2I_{\text{ext}} = 0. \quad (37)$$

The latter represents the F/W conservation law

$$L_W = \ell_W \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [M^*] + 2[A_2^*] + [F] + [W], \quad (38)$$

which is broken by the fueling mechanism

$$d_t L_W = I_F + I_W. \quad (39)$$

At the steady state all time derivative vanish, and we can use Supplementary Equation 39 to recast the chemical work per unit of time in Supplementary Equation 33 into

$$\dot{\mathcal{W}}_{\text{chem}} = \dot{\mathcal{W}}_{\text{fuel}} + \dot{\mathcal{W}}_{\text{ext}} \quad (40)$$

where

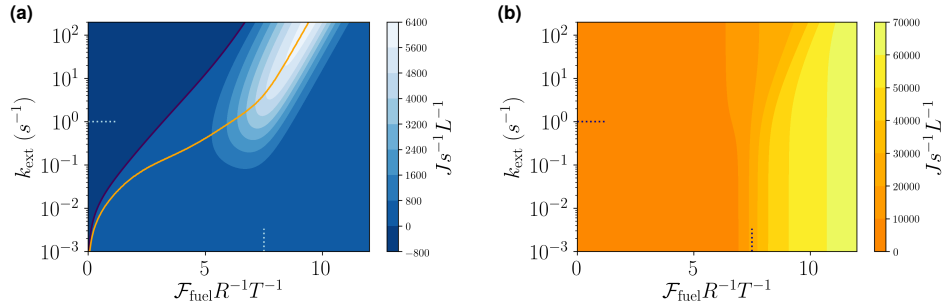
$$\dot{\mathcal{W}}_{\text{fuel}} = I_F (\mu_F - \mu_W). \quad (41)$$

is the input power, and

$$\dot{\mathcal{W}}_{\text{ext}} = I_{\text{ext}} (2\mu_M - \mu_{A_2}) \quad (42)$$

is the output power. By combining Supplementary Equation 40 with Supplementary Equation 32, we obtain Equation (4) of the main text.

c. Plots of $-\dot{\mathcal{W}}_{\text{ext}}$ and $\dot{\mathcal{W}}_{\text{fuel}}$



Supplementary Figure 2. **(a)** Minus the output power ($-\dot{\mathcal{W}}_{\text{ext}}$) and **(b)** input power ($\dot{\mathcal{W}}_{\text{fuel}}$) plotted in the same range of parameter as in Figure 4 of the main text. The efficiency is given by the ratio of the two plots, according to Equation (5) of the main text.

d. Linear Regime

For $k_{\text{ext}} = 0$ and $\mathcal{F}_{\text{fuel}} = \mu_F - \mu_W = 0$, the entropy production at the steady state vanishes, and hence the steady state is an equilibrium one ($[X]_{\text{eq}}$). For

$$k_{\text{ext}} \ll 1 \quad (43a)$$

$$\mathcal{F}_{\text{fuel}} = \mu_F - \mu_W \ll RT \quad (43b)$$

We now observe that from the definition of conservation law, the following constraint holds

$$\mathbf{0} = \mathbb{M} \boldsymbol{\ell}_M^\top = \mathbb{M}_a \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \mathbb{M}_e \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad (49)$$

which implies that

$$\mathbb{M}_M = -2\mathbb{M}_{A_2} - \mathbb{M}_a \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \quad (50)$$

This allows us to rewrite Supplementary Equations 48 as

$$\begin{aligned} 0 &= \mathcal{F}_{\text{fuel}} \mathbb{M}_F^a + \mathbb{M}_a^a \cdot \begin{pmatrix} f_{M^*} - f_M \\ f_{A_2^*} - 2f_M \end{pmatrix} + (f_{A_2} - 2f_M) \mathbb{M}_{A_2}^a \\ -I_{\text{ext}} &= \mathbb{M}_F^{A_2} \mathcal{F}_{\text{fuel}} + \mathbb{M}_a^{A_2} \cdot \begin{pmatrix} f_{M^*} - f_M \\ f_{A_2^*} - 2f_M \end{pmatrix} + (f_{A_2} - 2f_M) \mathbb{M}_{A_2}^{A_2} \\ I_F &= \mathcal{F}_{\text{fuel}} \mathbb{M}_F^F + \mathbb{M}_a^F \cdot \begin{pmatrix} f_{M^*} - f_M \\ f_{A_2^*} - 2f_M \end{pmatrix} + (f_{A_2} - 2f_M) \mathbb{M}_{A_2}^F. \end{aligned} \quad (51)$$

We now solve the first of the three equations above for the vector in parenthesis, using the fact that \mathbb{M}_a^a is nonsingular.

$$\begin{pmatrix} f_{M^*} - f_M \\ f_{A_2^*} - 2f_M \end{pmatrix} = -(\mathbb{M}_a^a)^{-1} \cdot [\mathcal{F}_{\text{fuel}} \mathbb{M}_F^a + (f_{A_2} - 2f_M) \mathbb{M}_{A_2}^a]. \quad (52)$$

This follows from the fact that \mathbb{M}_a^a is Gramian [8], and \mathbb{S}^a contains linearly independent vectors. Therefore, the last two equations in 51 can be recast into

$$\begin{aligned} -I_{\text{ext}} &= \mathcal{F}_{\text{fuel}} \left[\mathbb{M}_F^{A_2} - \mathbb{M}_a^{A_2} \cdot (\mathbb{M}_a^a)^{-1} \cdot \mathbb{M}_F^a \right] + (f_{A_2} - 2f_M) \left[\mathbb{M}_{A_2}^{A_2} - \mathbb{M}_a^{A_2} \cdot (\mathbb{M}_a^a)^{-1} \cdot \mathbb{M}_{A_2}^a \right] \\ I_F &= \mathcal{F}_{\text{fuel}} \left[\mathbb{M}_F^F - \mathbb{M}_a^F \cdot (\mathbb{M}_a^a)^{-1} \cdot \mathbb{M}_F^a \right] + (f_{A_2} - 2f_M) \left[\mathbb{M}_{A_2}^F - \mathbb{M}_a^F \cdot (\mathbb{M}_a^a)^{-1} \cdot \mathbb{M}_{A_2}^a \right]. \end{aligned} \quad (53)$$

Changing signs conveniently, we can rewrite the above equations in terms of the Onsager matrix \mathbb{L} , which expresses the linear dependence of currents from forces when the system is close to equilibrium:

$$\begin{pmatrix} I_F \\ I_{\text{ext}} \end{pmatrix} = \mathbb{L} \begin{pmatrix} \mu_F - \mu_W \\ 2\mu_M - \mu_{A_2} \end{pmatrix}. \quad (54)$$

Indeed, in the linear regime the chemical force associated to the extraction currents is $2\mu_M - \mu_{A_2} = 2f_M - f_{A_2}$. The entries of the Onsager matrix are given by

$$\mathbb{L} = \begin{pmatrix} \mathbb{M}_F^F - \mathbb{M}_a^F \cdot (\mathbb{M}_a^a)^{-1} \cdot \mathbb{M}_F^a & \mathbb{M}_a^F \cdot (\mathbb{M}_a^a)^{-1} \cdot \mathbb{M}_{A_2}^a - \mathbb{M}_{A_2}^F \\ \mathbb{M}_{A_2}^{A_2} - \mathbb{M}_a^{A_2} \cdot (\mathbb{M}_a^a)^{-1} \cdot \mathbb{M}_{A_2}^a & \mathbb{M}_a^{A_2} \cdot (\mathbb{M}_a^a)^{-1} \cdot \mathbb{M}_F^a - \mathbb{M}_F^{A_2} \end{pmatrix} := \begin{pmatrix} \mathbb{L}_{11} & \mathbb{L}_{12} \\ \mathbb{L}_{21} & \mathbb{L}_{22} \end{pmatrix}. \quad (55)$$

We can use Supplementary Equation 54 to analytically evaluate the efficiency η_{ds} introduced in Equation (5) of the main text, as well as the output power \mathcal{W}_{ext} , in terms of k_{ext} and $\mathcal{F}_{\text{fuel}}$, namely the control parameters in the model:

$$\eta_{\text{ds}} = -\frac{I_{\text{ext}}(I_{\text{ext}} - \mathcal{F}_{\text{fuel}} \mathbb{L}_{12})}{\mathcal{F}_{\text{fuel}}(I_{\text{ext}} \mathbb{L}_{12} + \mathcal{F}_{\text{fuel}} \det[\mathbb{L}])}; \quad \mathcal{W}_{\text{ext}} = \frac{I_{\text{ext}}(I_{\text{ext}} - \mathcal{F}_{\text{fuel}} \mathbb{L}_{12})}{\mathbb{L}_{11}}. \quad (56)$$

When $\mathcal{F}_{\text{fuel}}$ is kept fixed, the values of k_{ext} which extremise η_{ds} and $-\dot{W}_{\text{ext}}$ are readily found by deriving the previous expressions and look for the unique stable points:

$$\text{max efficiency : } k_{\text{ext}}^* = \frac{\sqrt{\mathbb{L}_{11}\mathbb{L}_{22}\det[\mathbb{L}] - \det[\mathbb{L}]}}{\mathbb{L}_{12}[A_2]_{\text{eq}}} \mathcal{F}_{\text{fuel}} \quad (57)$$

$$\text{max output power : } k_{\text{ext}}^* = \frac{\mathbb{L}_{12}}{2[A_2]_{\text{eq}}} \mathcal{F}_{\text{fuel}} \cdot \quad (58)$$

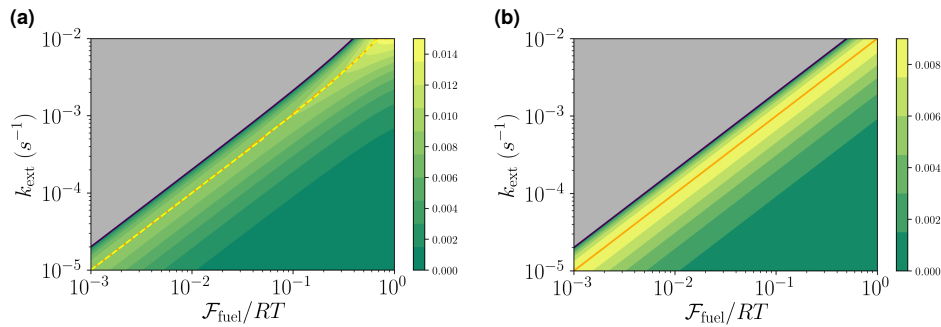
The above equations define the sets of points of maximum efficiency and efficiency at maximum power for any value of $\mathcal{F}_{\text{fuel}}$ within the linear regime. By equating the right hand sides of Supplementary Equations 57 and 58, one obtains that these two expressions coincide if and only if $\mathbb{L}_{12} = \mathbb{L}_{21} = 0$, which is never the case for coupled currents.

When evaluated using the parameters in Supplementary Table 1, Supplementary Equation 55 reads

$$\mathbb{L} = \begin{pmatrix} 17.7835 & 3.74893 \\ 3.74893 & 23.7732 \end{pmatrix} \cdot 10^{-8} \text{ mol}^2/\text{sLJ} \quad (59)$$

where the cross coefficients are equal according to the Onsager reciprocal relations.

When the analytical solution is plotted against k_{ext} and $\mathcal{F}_{\text{fuel}}$, we obtain the plot in Supplementary Figure 3b, where both maximum efficiency and efficiency at maximum power are highlighted as in Figure 3 of the main text. An enlargement of the linear region of Figure 3 of the main text is shown in Supplementary Figure 3a.



Supplementary Figure 3. Comparison between exact simulation of the full dynamics **(a)** and analytical formula obtained in the linear regime **(b)** for the efficiency in the linear regime. The log scale emphasizes the changes of magnitude of these values. For low forces and extraction rates — where Supplementary Equation 54 is a good approximation — the two plots clearly coincide. When $\mathcal{F}_{\text{fuel}}$ is of the order of 0.1 (in units of RT) and k_{ext} reaches 10^{-3} s^{-1} differences in both numerical values and shape emerge. In particular, we see that the increase in efficiency visible for high $\mathcal{F}_{\text{fuel}}$ and k_{ext} in (a) is a genuine nonequilibrium feature as it is absent in the linear regime, (b).

SUPPLEMENTARY REFERENCES

- [1] Riccardo Rao and Massimiliano Esposito, “Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics,” *Phys. Rev. X* **6**, 041064 (2016).
- [2] Gianmaria Falasco, Riccardo Rao, and Massimiliano Esposito, “Information thermodynamics of Turing patterns,” *Phys. Rev. Lett.* **121**, 108301 (2018).
- [3] Riccardo Rao and Massimiliano Esposito, “Conservation laws and work fluctuation relations in chemical reaction networks,” *J. Chem. Phys.* **149**, 245101 (2018).

- [4] R.D. Astumian, "Stochastic pumping of non-equilibrium steady-states: how molecules adapt to a fluctuating environment," *ChemComm* **54**, 427-444 (2018).
- [5] G. Ragazzon and L. J. Prins, "Energy consumption in chemical fuel-driven self-assembly," *Nat. Nanotechnol.* **13**, 882-889 (2018).
- [6] B. Altaner, M. Poletti, and M. Esposito, "Fluctuation-dissipation relations far from equilibrium," *Phys. Rev. Lett.* **117**, 180601 (2016).
- [7] Matteo Poletti and Massimiliano Esposito, "Effective fluctuation and response theory," *J. Stat. Phys.* (2019), [10.1007/s10955-019-02291-7](https://doi.org/10.1007/s10955-019-02291-7).
- [8] R.A. Horn and C.R. Johnson, *Matrix Analysis* (Cambridge University Press, 1985).

3

PHOTOCHEMICAL ENGINES

The following article is reprinted from
[E. PENOCCHIO, R. RAO and M. ESPOSITO, *J. Chem. Phys.* **155**, (2021), 114101]
under the conditions of AIP Publishing license¹.

The page numbers placed in the outer margins provide a continuous pagination throughout the thesis.

¹ <https://publishing.aip.org/wp-content/uploads/2019/10/AIPP-Author-License.pdf>

Nonequilibrium thermodynamics of light-induced reactions

Cite as: *J. Chem. Phys.* **155**, 114101 (2021); doi: 10.1063/5.0060774

Submitted: 22 June 2021 • Accepted: 17 August 2021 •

Published Online: 15 September 2021



View Online



Export Citation



CrossMark

Emanuele Penocchio,^{1,a)} Riccardo Rao,^{1,2,b)} and Massimiliano Esposito^{1,c)}

AFFILIATIONS

¹Complex Systems and Statistical Mechanics, Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg City, G. D. Luxembourg

²Simons Center for Systems Biology, School of Natural Sciences, Institute for Advanced Study, Princeton, New Jersey 08540, USA

^{a)} Author to whom correspondence should be addressed: emanuele.penocchio@uni.lu

^{b)} riccardorao@ias.edu

^{c)} massimiliano.esposito@uni.lu

ABSTRACT

Current formulations of nonequilibrium thermodynamics of open chemical reaction networks only consider chemostats as free-energy sources sustaining nonequilibrium behaviors. Here, we extend the theory to include incoherent light as a source of free energy. We do so by relying on a local equilibrium assumption to derive the chemical potential of photons relative to the system they interact with. This allows us to identify the thermodynamic potential and the thermodynamic forces driving light-reacting chemical systems out-of-equilibrium. We use this framework to treat two paradigmatic photochemical mechanisms describing light-induced unimolecular reactions—namely, the adiabatic and diabatic mechanisms—and highlight the different thermodynamics they lead to. Furthermore, using a thermodynamic coarse-graining procedure, we express our findings in terms of commonly measured experimental quantities, such as quantum yields.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0060774>

I. INTRODUCTION

The importance of physicochemical processes involving light is hard to overstate. They are ubiquitous in nature and constitute the prime mechanism driving our planet out-of-equilibrium. They power, for instance, the climate dynamics¹ as well as photosynthesis.^{2,3} The former example is mostly due to *photophysical* processes,⁴ as molecules absorb high frequency photons coming from the sun and decay back to the original ground state by emitting photons in the form of heat. Photosynthesis instead is a *photochemical* process⁴ since it involves chemical reactions powered by light. Here, the energy of photons is transduced into chemical free energy as molecules with a high chemical potential are synthesized from low-chemical-potential reactants (e.g., glucose from carbon dioxide and water).⁵ The opposite process where light is generated from chemical reactions can also happen, as in bioluminescence.⁶ Another crucial photochemical reaction senses light in animal vision: the photoisomerization of the 11-cis retinal chromophore to its all-trans form in rhodopsin.^{7,8} Nowadays, photochemical reactions are also commonly used to power synthetic molecular machines and devices.⁹

Examples include the first synthetic molecular motor;¹⁰ one of the few examples of a synthetic bimolecular motor;¹¹ early experimental demonstrations of the ability of molecular motors to move macroscopic objects;^{12,13} the first material performing macroscopic work, thanks to molecular motors;^{14,15} and the first experimental design of a molecular Maxwell-demon.¹⁶ Breakthrough experiments in the field of artificial photosynthesis also showed the possibility of using photochemical reactions to harvest free energy from a light source and transduce it into a gradient of ions across a membrane.^{17–19}

It comes therefore as no surprise that light-reacting matter has long intrigued scientists. However, it is only during the 19th Century that photophysics and photochemistry started to be systematically investigated by pioneers such as von Grothuss,²⁰ Draper,²¹ Lemoine,²² and Ciamician.²³ In the 20th Century, the development of modern physics spurred further investigations. The concept of photon allowed to formulate the so-called Stark–Einstein law,^{4,24} thus initiating the efforts to establish a correspondence between the number of reacting molecules and the number of photons absorbed in a photochemical reaction. Moreover, electronic structure theory²⁵ laid the basis for both the interpretation of absorption spectra of

organic molecules and the mechanistic understanding of chemical processes involving light.^{26,27} Nowadays, a theory describing photophysical and photochemical processes is well established^{28,29} and allows for computational studies^{30,31} and the accurate interpretation of ultra-fast spectroscopic experiments able to unveil mechanistic details.⁸

Focusing on energetic aspects, equilibrium thermodynamic descriptions of light-matter interactions were first formulated by Kirchhoff,³² Wien,³³ Lord Rayleigh,³⁴ and Planck³⁵ for black body radiation and by Einstein³⁶ for the photoelectric effect. Nonequilibrium aspects only started to be investigated in the second half of the last century,³⁷ often in the context of photosynthesis.^{38,39} The concept of a chemical potential quantifying the amount of free energy entering the system upon the absorption of a photon was, however, first introduced in the context of semiconductor solid state physics.^{40–42} Later on, this chemical potential was recognized as the maximal amount of reversible work that may be done by a photon that is absorbed by a chemical system.^{43,44} However, a nonequilibrium thermodynamics description of how light can drive photophysical and/or photochemical reactions out-of-equilibrium is still lacking. While progress has been made in recent years in describing the nonequilibrium thermodynamics of chemical reactions (in ideal^{45–51} and non-ideal solutions⁵² and even with diffusion^{53,54}), no free-energy sources other than chemical potentials gradients (which can be maintained by chemostating some species) have been considered.

In this paper, we establish the foundation of a nonequilibrium thermodynamic description of elementary chemical processes coupled to incoherent radiation. We show that the concept of chemical potential of the photons naturally emerges from the assumption of local equilibrium, which lays at the core of the modern formulations of nonequilibrium thermodynamics of chemical reactions.^{49,55,56} Einstein's relations between absorption and emission coefficients⁵⁷ provide a local detailed balance principle connecting thermodynamic quantities to dynamical ones. This leads to an explicit expression for the entropy production rate of elementary photochemical processes that allows us to unambiguously identify the thermodynamic potentials at work in photochemical systems and the forces driving them out-of-equilibrium. The outcome is a thermodynamic description of how light and reactions relax to equilibrium when interacting in a closed box or how radiation

sources can drive reactions out-of-equilibrium in an open system scenario. We also connect our theory to quantities commonly measured experimentally, such as quantum yields,⁵⁸ using a thermodynamically consistent coarse-graining.^{59,60} A key finding is that photochemical processes with the same coarse-grained reaction fluxes may have different dissipation rates depending on the underlying elementary mechanism.

This work lays the basis for quantitative energetic considerations in light-powered chemical systems. A direct application could be to evaluate the thermodynamic efficiency of light-driven molecular motors, whose performance has been assessed just from a dynamical standpoint.^{61,62} In the long run, our results pave the way for a unified thermodynamic perspective on free-energy transduction from different sources (e.g., light, electricity, and chemostats) mediated by chemical reactions.

The content of this paper is organized as follows: In Sec. II, we introduce our thermodynamic description on a simple model of elementary photophysical reaction. In Sec. III, we apply our theory to two common schemes of photochemical reactions, adiabatic and diabatic mechanisms.⁴ The practical relevance of our theoretical results is discussed in Sec. IV. Throughout this paper, we consider homogeneous ideal dilute solutions interacting with monochromatic light. This allows us to keep the mathematical treatment simple. Future extensions are outlined in Sec. V.

II. ELEMENTARY PHOTOPHYSICAL PROCESSES

The aim of this section is to build the nonequilibrium thermodynamic description of the prototypical photophysical process, which is depicted in Fig. 1. To do so, we adopt the approach developed in Refs. 47, 49, and 51 for purely thermal chemical processes. In Sec. II A, we introduce the kinetics of the model. We then discuss how thermodynamic state functions can be specified for chemical species and radiation in far-from-equilibrium regimes (Sec. II B). The connection between dynamics and thermodynamics is made in Sec. II C through the condition of local detailed balance. This condition ensures that a closed system made of light-reacting chemicals relaxes to equilibrium and does so by minimizing a suitable thermodynamic potential (Sec. II D). Instead, in the presence of external light sources, nonequilibrium steady states can be maintained by flows of free energy crossing the system (Sec. II E).

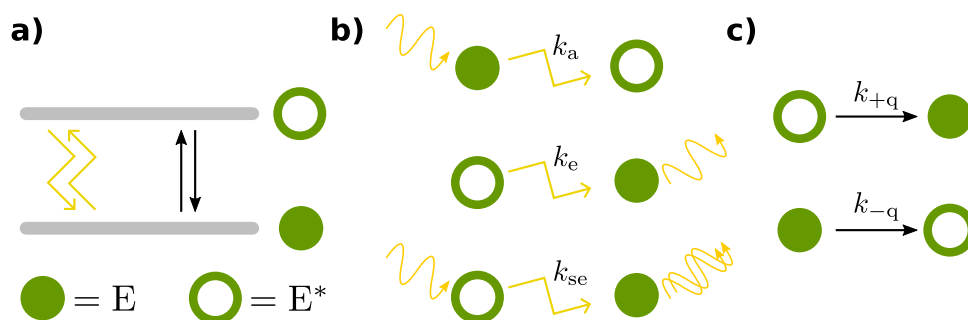
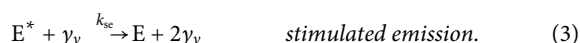
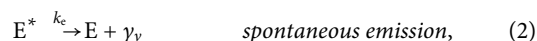
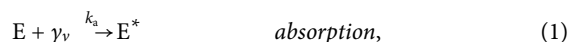


FIG. 1. Elementary photophysics of a two-level molecule. (a) Schematic representation (also known as Jablonski diagram⁴) of a two level molecule undergoing radiative and non-radiative transitions (quenching) among ground (E) and excited (E*) electronic states. (b) Absorption, spontaneous emission, and stimulated emission. (c) Quenching transitions.

A. Kinetics

We consider the *photophysical mechanism* depicted in Fig. 1. It comprises the so-called *primary events* in experimental photochemistry,^{4,63} i.e., transitions between different electronic states, here denoted as E and E*.

Among such transitions, those highlighted in Fig. 1(b) involve light in the form of a photon of frequency ν —denoted γ_ν . Their kinetics is characterized by the Einstein coefficients,^{57,64}



Taken together, these three transitions constitute the elementary *photophysical reaction*. Its net current reads

$$J_\nu = J_{+\nu} - J_{-\nu}, \quad (4)$$

where

$$J_{+\nu} = k_a n_\nu [E], \quad (5a)$$

$$J_{-\nu} = (k_e + k_{se} n_\nu) [E^*] \quad (5b)$$

denote the directed fluxes and n_ν denote the molar concentration of photons of frequency ν or angular frequency $\omega_\nu = 2\pi\nu$.

In addition to the photophysical transitions, we also consider a nonradiative thermal pathway connecting the ground and the excited states, as depicted in Fig. 1(c),



This is usually called the *quenching* reaction, and we assume it to be elementary. This implies that its current follows mass action kinetics,

$$J_q = J_{+q} - J_{-q}, \quad (7)$$

with

$$J_{+q} = k_{+q} [E^*], \quad (8a)$$

$$J_{-q} = k_{-q} [E]. \quad (8b)$$

Overall, the dynamics of the elementary photophysical and quenching reactions described in terms of concentrations reads

$$d_t [E^*] = -d_t [E] = J_\nu - J_q, \quad (9a)$$

$$d_t n_\nu = -J_\nu. \quad (9b)$$

As a consequence, the total concentration of E-molecules, denoted $E_0 \equiv [E] + [E^*]$, is conserved,

$$d_t E_0 = d_t ([E] + [E^*]) = 0 \quad (10)$$

but not the concentration of photons n_ν .

B. Thermodynamics

Nonequilibrium thermodynamics of chemical reactions crucially relies on the *local equilibrium* assumption. This means that all degrees of freedom other than those involved in reactions (i.e., concentrations) are considered to be at equilibrium.^{55,56} As a result, (i) the temperature T is well defined and it is set by the solvent, which acts as a thermal reservoir, and (ii) the chemical potentials of all chemical species are evaluated using their equilibrium expression in solution (expressed as energy minus temperature multiplying entropy) but evaluated at the nonequilibrium values of the concentrations.⁵⁶

By focusing first on the chemical species, we recall that, for an ideal dilute and homogeneous solution, the total internal energy U_{ch} and entropy S_{ch} per unit volume read⁴⁹

$$\begin{aligned} U_{\text{ch}} &= u_E^\circ [E] + u_{E^*}^\circ [E^*] \\ &= u_E^\circ [E] + (N_A \hbar \omega_\nu + u_E^\circ) [E^*], \end{aligned} \quad (11)$$

$$S_{\text{ch}} = (s_E + R) [E] + (s_{E^*} + R) [E^*], \quad (12)$$

where

$$s_E := s_E^\circ - R \ln [E], \quad s_{E^*} := s_{E^*}^\circ - R \ln [E^*] \quad (13)$$

denote the molar entropy of formation carried by the chemical species in the solution. In addition, R denotes the gas constant, while N_A denotes the Avogadro's number. Note that we assumed that the energy difference between the ground and the excited state is exactly the energy carried by a photon of frequency ν , i.e., $u_{E^*}^\circ - u_E^\circ = N_A \hbar \omega_\nu$. In Eq. (13), the s° terms take into account the possible degeneracy of the states, usually denoted g (i.e., $s^\circ = R \ln g$). The total free energy per unit volume of the chemical species consequently reads

$$F_{\text{ch}} = U_{\text{ch}} - T S_{\text{ch}}. \quad (14)$$

From this expression, the chemical potentials ensue,

$$\mu_E = \partial_{[E]} F_{\text{ch}} = u_E^\circ - T s_E \quad (15a)$$

$$= \mu_E^\circ + RT \ln [E], \quad (15b)$$

$$\mu_{E^*} = \partial_{[E^*]} F_{\text{ch}} = u_{E^*}^\circ - T s_{E^*} \quad (15c)$$

$$= \mu_{E^*}^\circ + RT \ln [E^*], \quad (15d)$$

where the standard chemical potentials are given by

$$\mu_E^\circ = u_E^\circ - T s_E^\circ, \quad \mu_{E^*}^\circ = u_{E^*}^\circ - T s_{E^*}^\circ. \quad (16)$$

We now extend this local equilibrium reasoning to the radiation. We thus introduce the equilibrium expressions for the thermodynamic potentials of radiation but evaluate them at arbitrary photons' concentration. The energy per unit volume and frequency carried by photons thus reads

$$U_{\text{ph}} = N_A \hbar \omega_\nu n_\nu, \quad (17)$$

and the entropy per unit volume and frequency (obtained using the expression derived for equilibrium Bose gases⁵⁶) reads

$$S_{\text{ph}} = R[(f_{\nu} + n_{\nu}) \ln(f_{\nu} + n_{\nu}) - n_{\nu} \ln n_{\nu} - f_{\nu} \ln f_{\nu}]. \quad (18)$$

Note that this local equilibrium assumption is equivalent to neglecting eigenstate coherences from the von Neumann entropy of the radiation. As further discussed in Sec. IV, it implies that the radiation may be out-of-equilibrium with respect to the thermal reservoir but incoherent. For large photon numbers—the limit of interest for this paper—the corresponding molar quantities are obtained as derivatives with respect to n_{ν} ,

$$u_{\nu} = \partial_{n_{\nu}} U_{\text{ph}} = N_{\text{A}} \hbar \omega_{\nu}, \quad (19)$$

$$s_{\nu} = \partial_{n_{\nu}} S_{\text{ph}} = R \ln\{(f_{\nu} + n_{\nu}) / n_{\nu}\}. \quad (20)$$

The energy of a photon is an intrinsic property that only depends on its frequency, while for chemical species, the same quantity must be specified with respect to some reference condition (the symbol \circ is used to indicate standard temperature and pressure).

In analogy with Eq. (15), we introduce the chemical potential of a mole of photons with frequency ν with respect to the temperature T of the solvent,

$$\mu_{\nu} = u_{\nu} - T s_{\nu} = N_{\text{A}} \hbar \omega_{\nu} - RT \ln \frac{f_{\nu} + n_{\nu}}{n_{\nu}}. \quad (21)$$

It measures the free energy carried by the photons with respect to the molecules in solution.

For *thermal radiation* at temperature T_r , we get that

$$\mu_{\nu}^{\text{th}} = u_{\nu} \left(1 - \frac{T}{T_r}\right), \quad (22)$$

since in that case,^{56,65}

$$n_{\nu}^{\text{th}} = \frac{f_{\nu}(n_{\nu})}{N_{\text{A}}}, \quad (23)$$

where

$$f_{\nu} = \frac{2\omega_{\nu}^2}{\pi c^3} \quad \text{and} \quad \langle n_{\nu} \rangle = \frac{1}{e^{N_{\text{A}} \hbar \omega_{\nu} / RT_r} - 1} \quad (24)$$

denote the density of photon states associated with the radiation and the Bose–Einstein distribution, respectively. When the temperature of radiation and solvent coincide, $T_r = T$, the chemical potential of the photons vanishes, in line with the fact that photons are not conserved in physical systems in contact with a thermal reservoir.⁴¹ Instead, in the limit of $T_r \rightarrow \infty$, the chemical potential of thermal photons coincides with the energy they carry. We note that expression (22) for chemical potential of thermal photons is consistent with previous formulations.^{42–44,65,66} Different from these, our formulation explicitly leverages the condition of local-equilibrium, in analogy with the approach of nonequilibrium thermodynamics of chemical reactions.^{49,52}

C. Local detailed balance

We now make use of local detailed balance^{67–69} to ensure the thermodynamic consistency of our kinetic modeling. This concept was first proposed in Ref. 70 to ensure that the stochastic dynamics of an open system correctly describes the effect of multiple reservoirs. For elementary chemical reactions, it stipulates that the

log-ratio of forward and backward rate constants is proportional to the difference of standard-state chemical potentials between products and reagents.^{46,49} For closed systems, this relation is essentially the law of mass action. Its extension to open systems^{71–73} ensures that microscopic reversibility holds at equilibrium.^{74–77}

For the quenching reaction (6), local detailed balance implies that

$$\frac{k_{+q}}{k_{-q}} = \exp\left\{\frac{\mu_{\text{E}^*}^{\circ} - \mu_{\text{E}}^{\circ}}{RT}\right\}, \quad (25)$$

and using (7) and (15), it enables us to identify the quenching affinity as

$$A_{\text{q}} \equiv \mu_{\text{E}^*} - \mu_{\text{E}} = RT \ln \frac{J_{+q}}{J_{-q}}. \quad (26)$$

This last relation binds the affinity to the reaction fluxes and implies that chemical forces and reaction currents are always aligned. We emphasize that Eqs. (25) and (26) hold for elementary reactions but not necessarily for coarse-grained processes, as we will see (see also Refs. 59 and 60).

For the photophysical reaction, the local detailed balance is essentially the Einstein relations, which were explicitly derived to ensure thermodynamic consistency in kinetic models,⁵⁷

$$\frac{k_{\text{a}}}{k_{\text{se}}} = \exp\left\{\frac{s_{\text{E}^*}^{\circ} - s_{\text{E}}^{\circ}}{R}\right\}, \quad (27a)$$

$$\frac{k_{\text{e}}}{k_{\text{se}}} = f_{\nu}. \quad (27b)$$

Substituting them into Eq. (4), we can express the forward and backward currents as

$$J_{+v} = k_{\text{a}} n_{\nu} [\text{E}] = e^{(s_{\text{E}^*}^{\circ} - s_{\text{E}}^{\circ}) / R} k_{\text{se}} n_{\nu} [\text{E}], \quad (28a)$$

$$J_{-v} = (k_{\text{e}} + k_{\text{se}} n_{\nu}) [\text{E}^*] = k_{\text{se}} (f_{\nu} + n_{\nu}) [\text{E}^*], \quad (28b)$$

and using Eqs. (15) and (20), we find

$$A_{\text{v}} \equiv \mu_{\text{E}} + \mu_{\nu} - \mu_{\text{E}^*} = RT \ln \frac{J_{+v}}{J_{-v}}, \quad (29)$$

where A_{v} is the affinity of the photophysical reaction.

From a thermodynamic standpoint, *equilibrium* is reached when all affinities vanish,

$$A_{\text{q}}^{\text{eq}} = \mu_{\text{E}^*}^{\text{eq}} - \mu_{\text{E}}^{\text{eq}} = 0, \quad (30a)$$

$$A_{\text{v}}^{\text{eq}} = \mu_{\text{E}}^{\text{eq}} + \mu_{\nu}^{\text{eq}} - \mu_{\text{E}^*}^{\text{eq}} = 0. \quad (30b)$$

Therefore, the chemical potential of the photons relative to the solution must vanish, $\mu_{\nu}^{\text{eq}} = 0$. This obviously implies that the temperature of the radiation and the temperature of the thermal reservoir must be equal [see Eq. (22)]. Finally, through Eqs. (26) and (29), Eq. (30) implies that the equilibrium currents vanish as well,

$$J_{\text{v}}^{\text{eq}} = 0 \quad \text{and} \quad J_{\text{q}}^{\text{eq}} = 0. \quad (31)$$

D. Molecules and radiation in a closed box

Consider the situation where the radiation and the molecules in solution at temperature T are enclosed in a box with walls perfectly reflecting at frequency ν , as depicted in Fig. 2. The system is thus closed with respect to both matter and photons but can still exchange heat with the solvent via the non-radiative quenching reaction. Its state is defined by the concentrations of photons and chemical species n_ν , $[E]$, and $[E^*]$, which obey the nonlinear dynamics (9). The condition of steady state implies equilibrium between the photophysical and quenching reactions [see Eq. (31)]. This also implies that

$$\frac{[E^*]^{\text{eq}}}{[E]^{\text{eq}}} = \frac{k_{-q}}{k_{+q}} = \frac{k_a n_\nu^{\text{eq}}}{k_e + k_{se} n_\nu^{\text{eq}}} \quad (32)$$

and more explicitly that

$$[E]^{\text{eq}} = E_0 - [E^*]^{\text{eq}} = \frac{k_{+q} E_0}{k_{+q} + k_{-q}} = \frac{(k_e + k_{se} n_\nu^{\text{eq}}) E_0}{k_e + k_{se} n_\nu^{\text{eq}} + k_a n_\nu^{\text{eq}}}, \quad (33)$$

where the total concentration of chemicals E_0 is set by the initial condition.

Turning to energetic considerations, the first law of thermodynamics expresses the fact that the change in internal energy of the system per unit time—evaluated using (11) and (17) with (9)—is due to the heat flow \dot{Q} entering the system,

$$d_t(U_{\text{ch}} + U_{\text{ph}}) = -N_A \hbar \omega_\nu J_q = \dot{Q}. \quad (34)$$

The second law states that the entropy production—defined as the entropy change in the system and in the reservoir (viz., the solvent)—must always be greater or equal than zero. This is verified as

$$\dot{\Sigma} = d_t(S_{\text{ch}} + S_{\text{ph}}) - \frac{\dot{Q}}{T} = J_\nu \frac{A_\nu}{T} + J_q \frac{A_q}{T} \geq 0, \quad (35)$$

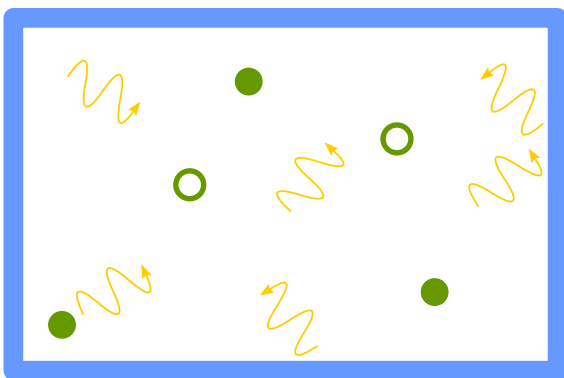


FIG. 2. Closed Box. Radiation interacting with species E and E^* via the scheme of Fig. 1 in an ideal dilute solution at temperature T inside a closed box with perfectly reflecting walls. Photons cannot leave the system, but their number is not conserved due to the coupling between light-induced and quenching reactions.

where we used (12) and (18) with (9) and the affinities (26) and (29). In particular, the inequality readily follows from Eq. (26).

We now show that this system relaxes to equilibrium and its relaxation is well characterized by the nonequilibrium total free energy

$$F = F_{\text{ch}} + F_{\text{ph}}, \quad (36)$$

with

$$F_{\text{ch}} = U_{\text{ch}} - TS_{\text{ch}} = (\mu_A - R)[E] + (\mu_{E^*} - R)[E^*], \quad (37a)$$

$$F_{\text{ph}} = U_{\text{ph}} - TS_{\text{ph}} = \mu_\nu n_\nu - RT f_\nu \ln(f_\nu + n_\nu), \quad (37b)$$

where the temperature T of the solution has been used as reference in the expression of both F_{ch} and F_{ph} . Indeed, from Eqs. (34) and (35), this total free energy can only decrease,

$$d_t F = d_t(F_{\text{ch}} + F_{\text{ph}}) = -T \dot{\Sigma} \leq 0. \quad (38)$$

In addition, the total free energy is lower bounded by its equilibrium value,

$$\begin{aligned} & (F_{\text{ch}} + F_{\text{ph}}) - (F_{\text{ch}}^{\text{eq}} + F_{\text{ph}}^{\text{eq}}) \\ &= [E^*] RT \ln \frac{[E^*]}{[E^*]_{\text{eq}}} - RT([E^*] - [E^*]_{\text{eq}}) \\ &+ [E] RT \ln \frac{[E]}{[E]_{\text{eq}}} - RT([E] - [E]_{\text{eq}}) \\ &+ f_\nu RT \ln \frac{f_\nu + n_\nu^{\text{eq}}}{f_\nu + n_\nu} + n_\nu RT \ln \frac{n_\nu(f_\nu + n_\nu^{\text{eq}})}{n_\nu^{\text{eq}}(f_\nu + n_\nu)} \geq 0. \end{aligned} \quad (39)$$

The first equality follows from

$$([E^*]_{\text{eq}} - [E^*]) \mu_{E^*}^{\text{eq}} = ([E]_{\text{eq}} - [E]) \mu_E^{\text{eq}} + (n_\nu^{\text{eq}} - n_\nu) \mu_\nu^{\text{eq}}, \quad (40)$$

which holds at any time by virtue of the equilibrium condition (30) and the conservation law (10). The inequality is set by the properties of logarithms (log inequality). Overall, Eqs. (38) and (39) show that the total free energy acts as a Lyapunov function: for any initial condition, the radiation will eventually thermalize with the solution by minimizing $F_{\text{ch}} + F_{\text{ph}}$.

We note that this treatment can be straightforwardly extended to a multi-frequency situation by including more chemical species or more photophysical and quenching transitions in the model.

E. Molecules and radiation in a transparent box

We now turn to the situation where an external mechanism keeps the concentration of photons n_ν constant inside the system, as illustrated in Fig. 3. We will refer to such a mechanism as *radiostat*, in analogy with *thermostats* and *chemostats*. As n_ν is not anymore a dynamical variable, only Eq. (9a) defines the dynamics, which is now linear and can be easily solved analytically. Equation (9b) only defines the rate at which photons must enter the system to keep their concentration constant. The steady-state condition will now

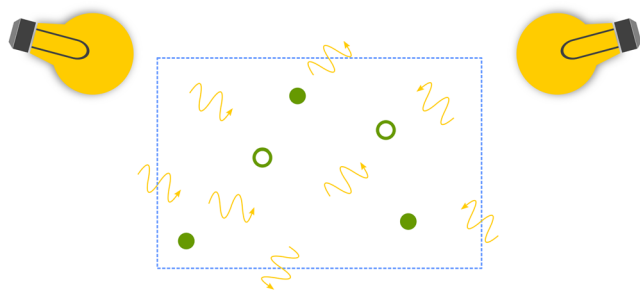


FIG. 3. Transparent box. Radiation interacting with species E and E* via the scheme of Fig. 1 in an ideal dilute solution at temperature T inside a box with perfectly transparent walls. The system is closed to exchanges of molecules but is open to photons, which can freely enter and exit the system. Photons' concentration is controlled by a radiostat.

generically produce nonvanishing currents $\bar{J}_v = \bar{J}_q$ corresponding to a nonequilibrium steady state,

$$[\bar{E}] = E_0 - [\bar{E}^*] = \frac{(k_{+q} + k_e + k_{se}n_v)E_0}{k_{+q} + k_{-q} + k_e + k_{se}n_v + k_a n_v}. \quad (41)$$

From now on, overbars will denote quantities at steady state. We also note that the ratio of out-of-equilibrium steady-state concentrations are governed by the relative magnitude of the kinetic rate constants rather than the relative thermodynamic stability of the molecules—as it is at equilibrium,

$$\frac{[\bar{E}^*]}{[\bar{E}]} = \frac{k_{-q} + k_a n_v}{k_{+q} + k_e + k_{se}n_v}. \quad (42)$$

As a special case, we consider the limit of very bright light: $n_v \rightarrow \infty$, which, for thermal radiation, corresponds to $T_r \rightarrow \infty$. In this case, the steady-state ratio (42) is governed by the ratio between the Einstein coefficients for absorption and stimulated emission and therefore by the ratio of the degeneracy between the ground and excited states [see Eq. (27a)].

We now proceed with thermodynamics. As the radiation inside the box is held constant by the radiostat, the only contribution to the change in the system's internal energy per unit time comes from the molecules in solution and can be decomposed—using (11) with (9), (4), and (7)—as

$$d_t U_{ch} = \dot{U}_v + \dot{Q}, \quad (43)$$

where

$$\dot{Q} = -N_A \hbar \omega_\nu J_q \quad (44)$$

is the rate of heat absorbed by the molecules and

$$\dot{U}_v = J_v u_\nu \quad (45)$$

is the rate of energy absorbed by the molecules from the photons. The entropy production rate now reads

$$\dot{\Sigma} = d_t S_{ch} - \frac{\dot{Q}}{T} - \dot{S}_v = J_v \frac{A_v}{T} + J_q \frac{A_q}{T} \geq 0. \quad (46)$$

In contrast to Eq. (35), $\dot{\Sigma}$ now accounts for the rate of entropy exchanged with the radiostat, $\dot{S}_v = J_v s_v$. To obtain the second equality ensuring the non-negativity of the entropy production, we used (12) with (9) and the affinities (26) and (29).

By combining Eqs. (43) and (46), we get the free-energy balance

$$T\dot{\Sigma} = \dot{F}_v - d_t F_{ch} \geq 0, \quad (47)$$

where

$$\dot{F}_v = J_v \mu_\nu \quad (48)$$

represents the rate of free energy absorbed by the molecules in solution from the radiation. This result shows that the variation of free energy in the system (e.g., the free energy stored) is upper bounded by the free energy absorbed from the photons.

We consider now the special case where the radiostat is a black body emitting thermal radiation at temperature T_r (e.g., a lamp). From Eqs. (20) and (23) and as expected from a thermal object, $\dot{U}_v = T_r \dot{S}_v$. As a consequence, the term describing radiation absorption in the energy balance (\dot{U}_v) enters the entropy balance as energy over the radiation temperature ($\dot{S}_v = \dot{U}_v / T_r$). In other words, the radiostat acts as a thermal reservoir—at temperature T_r —that exchanges heat with the systems through photons. The free energy carried by photons is given by the chemical potential μ_ν^{th} , Eq. (22). Hence, the rate of radiation free energy entering the system [Eq. (48)] becomes

$$\dot{F}_v = J_v \mu_\nu^{\text{th}} = \dot{U}_v \left(1 - \frac{T}{T_r}\right). \quad (49)$$

When $T_r = T$, we have $\dot{F}_v = 0$ and the system relaxes to equilibrium by minimizing F_{ch} , as in the closed case [see Eq. (47)]. However, whenever there is a mismatch between the temperature of the radiation reservoir and the temperature of the thermal reservoir, the system will be driven out-of-equilibrium and eventually reach a nonequilibrium steady state. At steady state, if $T_r > T$, the heat flows from the radiation to the solvent and vice versa if $T_r < T$. If one starts from an equilibrium solution at temperature T and turns on the radiation temperature $T_r > T$, the radiation heat flow can be used as a resource to drive and sustain accumulation of free energy in the system. Such energy storage phenomena is reminiscent of what happens in photosynthetic systems, as sunlight can be regarded as black body radiation at a temperature of about 5800 K.

When the temperature of the black body radiation becomes very large, $T_r \rightarrow \infty$ (more realistically when $N_A \hbar \omega_\nu \ll RT_r$), spontaneous emission becomes negligible compared to absorption and stimulated emission and the entropy of the radiation vanishes: $s_v \rightarrow 0$. As a consequence, the free energy absorbed is exclusively made of energy,

$$\dot{F}_v = J_v \mu_\nu \rightarrow J_v u_\nu = \dot{U}_v, \quad (50)$$

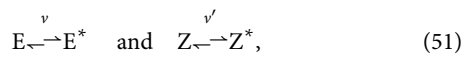
and the radiation can be regarded as the work source.

III. BASIC PHOTOCHEMICAL MECHANISMS

So far, we neglected the possibility that additional excited-state reactive events convert the species E^* into a different one, Z . The net effect, $E \xrightarrow{\text{light}} Z$, is, in fact, the crux of photochemistry. Often, the details of such excited-state dynamics are not explicitly specified in the kinetic descriptions.^{61,62} However, as we will see in this section, distinct mechanisms lead to different thermodynamics. We will consider two of the most common models for photochemical unimolecular reactions, namely, the adiabatic mechanism (Sec. III A) and the diabatic one (Sec. III B).^{4,78} To connect our results to experimental observations, we derive the coarse-grained description of both mechanisms in Sec. III C. Such a description does not require a detailed knowledge of excited state dynamics.

A. Adiabatic mechanism

Consider the photoisomerization scheme in Fig. 4. It comprises two elementary photophysical reactions,



and two thermally induced non-radiative reactions,

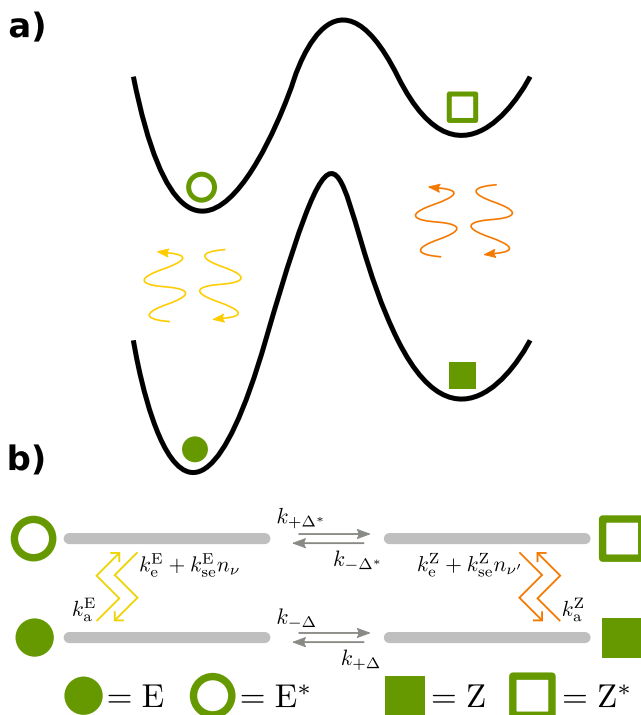
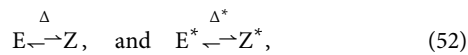


FIG. 4. Adiabatic mechanism. (a) Pictorial representation of the ground and excited state potential energy surfaces along the reaction coordinate interconverting species E and Z . (b) Schematic representation (also known as the Jablonski diagram) of the mechanism. This scheme can represent Förster cycles describing excited-state proton transfer reactions, where, for instance, species Z is a zwitterionic tautomer⁷⁹ of E .^{4,60,81}

which connect species in the same electronic state. Note that (i) the two photophysical reactions are activated by photons of different frequencies, ν and ν' , and (ii) we disregard the quenching reactions associated with the photophysical reactions (51).

For the moment, we consider the concentrations of photons n_ν and $n_{\nu'}$ as controlled by two different radiostats (see discussion in Sec. II E). The system's dynamics is ruled by the net absorption current of photons γ_ν going from E to E^* and the one of photons $\gamma_{\nu'}$ going from Z to Z^* ,

$$J_\nu = k_a^E n_\nu [E] - (k_e^E + k_{se}^E n_\nu) [E^*], \quad (53a)$$

$$J_{\nu'} = k_a^Z n_{\nu'} [Z] - (k_e^Z + k_{se}^Z n_{\nu'}) [Z^*], \quad (53b)$$

and by the reaction currents from E^* to Z^* and from Z to E ,

$$J_{\Delta^*} = k_{+\Delta^*} [E^*] - k_{-\Delta^*} [Z^*], \quad (54a)$$

$$J_\Delta = k_{+\Delta} [Z] - k_{-\Delta} [E]. \quad (54b)$$

It is easy to realize that, at the steady state, the currents must satisfy the following relations:

$$\bar{J} \equiv \bar{J}_\nu = \bar{J}_{\Delta^*} = -\bar{J}_{\nu'} = \bar{J}_\Delta, \quad (55)$$

with the special case $\bar{J} = 0$ describing equilibrium.

The total internal energy and entropy of the system are given by the expressions in Eqs. (11) and (12), respectively, but summed over all chemical species,

$$U_{\text{ch}} = \sum_{\substack{X=E, \\ E^*, Z, Z^*}} u_x^\circ [X], \quad S_{\text{ch}} = \sum_{\substack{X=E, \\ E^*, Z, Z^*}} (s_X + R) [X]. \quad (56)$$

The total free energy is thus given by $F_{\text{ch}} = U_{\text{ch}} - TS_{\text{ch}}$. Based on the results from Sec. II, Eq. (46), the entropy production can be expressed as the sum of currents times affinities,

$$\dot{\Sigma} = \sum_{\substack{\rho=\nu, \nu', \\ \Delta, \Delta^*}} J_\rho \frac{A_\rho}{T} = R \sum_{\substack{\rho=\nu, \nu', \\ \Delta, \Delta^*}} J_\rho \ln \frac{J_{+\rho}}{J_{-\rho}} \geq 0, \quad (57)$$

where the sum runs over all the reactions [Eqs. (53) and (54)]. As before [Eq. (47)], by combining Eqs. (57) and (56), we get the free-energy balance as

$$T\dot{\Sigma} = -d_t F_{\text{ch}} + \dot{F}_\nu + \dot{F}_{\nu'}, \quad (58)$$

where the free-energy absorbed has contributions from both the light sources: $\dot{F}_\nu = J_\nu \mu_\nu$, and $\dot{F}_{\nu'} = J_{\nu'} \mu_{\nu'}$.

We now focus on steady states. Using Eq. (55), we can rewrite the entropy production (58) as

$$T\dot{\Sigma} = \bar{J}(\mu_\nu - \mu_{\nu'}). \quad (59)$$

This result demonstrates that the adiabatic mechanism at steady state always converts photons with high chemical potential into pho-

tons with a lower one and thus the sign of the current \bar{J} is determined by the relative magnitude of μ_ν and $\mu_{\nu'}$. Equation (59) shows that when the photons have the same chemical potential, the system is detailed balanced, i.e., the unique steady state of the dynamics is an equilibrium state, where all currents in the system vanish according to condition (31). When combined with Eqs. (55) and (57), Eq. (59) yields the following relation:

$$\frac{k_a^E n_\nu k_{+\Delta^*} (k_e^Z + k_{se}^Z n_{\nu'}) k_{-\Delta}}{k_a^Z n_{\nu'} k_{-\Delta^*} (k_e^E + k_{se}^E n_\nu) k_{+\Delta}} = \exp\left\{\frac{\mu_\nu - \mu_{\nu'}}{RT}\right\}, \quad (60)$$

which is a consequence of the local detailed balance property discussed in Sec. II C. Importantly, this relation can be seen as a generalization of the concept known as the *Föster cycle*.^{4,80} Indeed, by using Eqs. (21) and (29), we can further recast Eq. (60) as

$$\ln \frac{k_{+\Delta^*}}{k_{-\Delta^*}} - \ln \frac{k_{+\Delta}}{k_{-\Delta}} = \frac{u_\nu - u_{\nu'}}{RT} + \frac{s_{Z^*}^{\circ} - s_{E^*}^{\circ}}{R} - \frac{s_{Z}^{\circ} - s_{E}^{\circ}}{R}. \quad (61)$$

Föster's relation is recovered when assuming that $s_{Z^*}^{\circ} - s_{E^*}^{\circ} = s_{Z}^{\circ} - s_{E}^{\circ}$,^{4,80,82} i.e., the excess of entropy of one molecule over the other is conserved between the two electronic states. It allows us to determine the equilibrium distribution of the excited state ($k_{+\Delta^*}$, $k_{-\Delta^*} = [Z^*]_{\text{eq}}$, $[E^*]_{\text{eq}}$) from the equilibrium distribution of the ground state ($k_{+\Delta}$, $k_{-\Delta} = [Z]_{\text{eq}}$, $[E]_{\text{eq}}$) (see discussion in Sec. IV for further remarks).

We conclude the discussion of the adiabatic mechanism by considering the case where photons γ_ν and $\gamma_{\nu'}$ comes from two different thermal sources with temperatures T_r and $T_{r'}$, respectively. In such a case, by virtue of the expression (22) for the photons chemical potentials, the steady state entropy production in Eq. (59) takes the following form:

$$\bar{\Sigma} = \bar{J} \left[u_\nu \left(\frac{1}{T} - \frac{1}{T_\nu} \right) - u_{\nu'} \left(\frac{1}{T} - \frac{1}{T_{\nu'}} \right) \right]. \quad (62)$$

This illustrates that the system may reach equilibrium ($\bar{\Sigma} = 0$) even in the presence of a temperature difference between the two light sources ($T_r \neq T_{r'}$). For this to happen, it suffices that the ratio of photons energies (u_ν , $u_{\nu'}$) makes the term in square brackets vanish. We note that detailed-balanced systems with temperature gradients are generally possible in *tightly coupled* (also referred to as strongly coupled) systems, where a single steady-state current [here \bar{J} , see Eq. (55)] is responsible for the whole dissipation.⁸³ Back to Eq. (62), when the two temperatures are equal ($T_r = T_{r'}$)—which happens when all the photons come from the same thermal source—the entropy production further reduces to

$$\bar{\Sigma} = \bar{J} N_A (\hbar\omega_\nu - \hbar\omega_{\nu'}) \left(\frac{1}{T} - \frac{1}{T_r} \right). \quad (63)$$

This shows that a nonequilibrium steady state may arise only if $T \neq T_r$ and the two radiative transitions couple photons with different frequencies, i.e., $u_\nu \neq u_{\nu'}$. If either of these conditions is not met, the system remains detailed balanced.

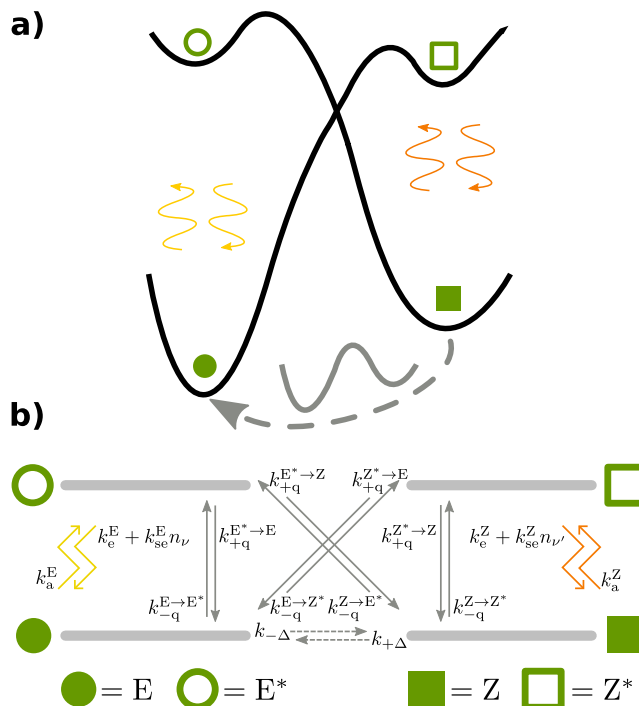
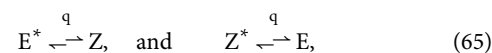
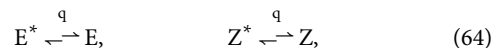


FIG. 5. Diabatic mechanism. (a) Pictorial representation of the ground and excited state potential energy surfaces along a reaction coordinate interconverting species E and Z through a conical intersection.^{28,29} The gray dashed arrow represents an alternative ground state pathway along a different reaction coordinate as in Ref. 84. (b) Schematic representation (also known as the Jablonski diagram⁴) of the mechanism. This scheme can represent typical E-Z photo-induced isomerizations⁹ of organic molecules.^{28,29}

B. Diabatic mechanism

We now consider the photoisomerization scheme in Fig. 5. As before, the excited states of E and Z are reached through elementary photophysical reactions [see Eqs. (51) and (53)]. However, in contrast to the adiabatic mechanism, the diabatic one includes a so-called *conical intersection*,^{28,29} i.e., the ground and excited potential energy surfaces cross each other, as in Fig. 5(a). Hence, multiple quenching thermal transitions connect excited and ground state species,



whose currents read

$$J_q^{E^* \rightarrow E} = k_{+q}^{E^* \rightarrow E} [E^*] - k_{-q}^{E \rightarrow E^*} [E], \quad (66a)$$

$$J_q^{Z^* \rightarrow Z} = k_{+q}^{Z^* \rightarrow Z} [Z^*] - k_{-q}^{Z \rightarrow Z^*} [Z], \quad (66b)$$

$$J_q^{E^* \rightarrow Z} = k_{+q}^{E^* \rightarrow Z} [E^*] - k_{-q}^{Z \rightarrow E^*} [Z], \quad (66c)$$

$$J_q^{Z^* \rightarrow E} = k_{+q}^{Z^* \rightarrow E} [Z^*] - k_{-q}^{E \rightarrow Z^*} [E]. \quad (66d)$$

On top of these, another non-radiative thermal reaction (represented by the dashed line in scheme 5) converts Z into E. This conversion happens along a reaction coordinate other than the photochemical one, and its current reads

$$J_{\Delta} = k_{+\Delta}[Z] - k_{-\Delta}[E]. \quad (67)$$

Note that because of the local detailed balance property for thermal reactions [Eq. (26)], the kinetic constants appearing in Eqs. (66) and (67) are not all independent from each other and must satisfy

$$\frac{k_{-q}^{E \rightarrow E^*} k_{+q}^{E^* \rightarrow Z} k_{+\Delta}}{k_{+q}^{E^* \rightarrow E} k_{-q}^{Z \rightarrow E^*} k_{-\Delta}} = \frac{k_{-q}^{Z \rightarrow Z^*} k_{+q}^{Z^* \rightarrow E} k_{-\Delta}}{k_{+q}^{Z^* \rightarrow Z} k_{-q}^{E \rightarrow Z^*} k_{+\Delta}} = 1. \quad (68)$$

If such relations were violated, then nonphysical cyclic currents would originate, thus preventing the reactions from reaching detailed balance in closed systems.

The entropy production rate of the diabatic mechanism can be formally written as in Eq. (58). However, in contrast to the adiabatic mechanism, the steady-state currents now must satisfy

$$\bar{J}_v = \bar{J}_q^{E^* \rightarrow E} + \bar{J}_q^{E^* \rightarrow Z}, \quad (69a)$$

$$\bar{J}_{v'} = \bar{J}_q^{Z^* \rightarrow Z} + \bar{J}_q^{Z^* \rightarrow E}, \quad (69b)$$

$$\bar{J}_{\Delta} = \bar{J}_q^{E^* \rightarrow Z} - \bar{J}_q^{Z^* \rightarrow E} = \bar{J}_q^{E^* \rightarrow Z} - \bar{J}_q^{Z^* \rightarrow E}. \quad (69c)$$

In particular, the two currents \bar{J}_v and $-\bar{J}_{v'}$ are not constrained to be equal. As a result, nonequilibrium steady states may arise also when the photons have the same chemical potential, giving rise to a steady-state entropy production of the form

$$T\dot{\bar{\Sigma}} = \dot{F}_v + \dot{F}_{v'} = \bar{J}_v \mu_v + \bar{J}_{v'} \mu_{v'}. \quad (70)$$

As we did for the adiabatic case (Sec. III A), we conclude this discussion by considering the case of photons γ_v and $\gamma_{v'}$ coming from two different thermal sources—with temperatures T_r and $T_{r'}$, respectively. In such a case, by virtue of expression (22) for the chemical potentials of the photons, the steady-state entropy production (70) takes the following form:

$$\dot{\bar{\Sigma}} = \bar{J}_v u_v \left(\frac{1}{T} - \frac{1}{T_r} \right) - \bar{J}_{v'} u_{v'} \left(\frac{1}{T} - \frac{1}{T_{r'}} \right). \quad (71)$$

As the diabatic mechanism is not tightly coupled—note two distinct current–force contribution at steady state, $\dot{\bar{\Sigma}}$ never vanishes if the radiation temperatures are different ($T_r \neq T_{r'}$). When these two temperatures are equal ($T_r = T_{r'}$) or when all the photons come from the same thermal source with temperature T_r , then the entropy production further reduces to

$$\dot{\bar{\Sigma}} = (\bar{J}_v N_A \hbar \omega_v + \bar{J}_{v'} N_A \hbar \omega_{v'}) \left(1 - \frac{T}{T_r} \right). \quad (72)$$

This relation shows that a photochemical process following the diabatic mechanism is detailed balanced (i.e., relaxes to equilibrium) only when all interacting photons have null chemical potential, namely, they are in thermal equilibrium with the solution at temperature T (see Sec. II C).

C. Coarse-grained description

To make contact with common experimental characterizations of photochemical mechanisms,⁵⁸ one should not rely on the detailed knowledge of the excited states dynamics.⁴ Hence, we now derive effective reactions for the photochemical mechanisms that only involve the species in the ground state, E and Z. This is achieved using a thermodynamically consistent coarse-graining procedure specifically developed for chemical reactions networks^{59,60} and based on a standard steady-state treatment of fast processes.⁸⁵ The underlying assumption is that the species in the excited state are sufficiently reactive for their concentrations to be considered small and always in a steady state, i.e.,

$$d_t[E^*] = 0, \quad (73a)$$

$$d_t[Z^*] = 0. \quad (73b)$$

This timescale-separation approximation is well justified as the lifetime of an excited intermediate is typically of 10^{-8} s in the absence of phosphorescent states (which are not taken into account in schemes 4 and 5).^{4,63} Imposing Eq. (73) implies that the concentrations of the excited species ($[E^*]$ and $[Z^*]$) become functions of the concentrations of the ground state species ($[E]$ and $[Z]$), which thus remain as the only dynamical variables in the model.⁸⁵ Therefore, photochemical processes are described as effective reactions involving ground state species only and whose kinetics is given by combinations of the kinetic constants of the elementary reactions. Standard experiments access precisely these effective kinetic parameters,⁵⁸ and we will show that their thermodynamic properties will depend on the underlying elementary models.

1. Coarse-grained adiabatic mechanism

In the adiabatic mechanism depicted in Fig. 4, conditions (73) allow the coarse-graining of the photochemical pathway $E \xrightarrow{h\nu} E^* \xrightarrow{-} Z^* \xrightarrow{-} Z$ into an effective reaction of the kind $E \xrightarrow{-} Z$, whose current reads

$$\begin{aligned} \Psi &= k_{+\Delta^*}[E^*] - k_{-\Delta^*}[Z^*] \\ &= \Phi^{E \rightarrow Z} k_a^E n_v [E] - \Phi^{Z \rightarrow E} k_a^Z n_{v'} [Z], \end{aligned} \quad (74)$$

where

$$\Phi^{E \rightarrow Z} = \frac{k_{+\Delta^*} (k_c^Z + k_{sc}^Z n_{v'})}{k_{-\Delta^*} (k_c^E + k_{se}^E n_v) + (k_{+\Delta^*} + k_c^E + k_{se}^E n_v) (k_c^Z + k_{sc}^Z n_{v'})}, \quad (75a)$$

$$\Phi^{Z \rightarrow E} = \frac{k_{-\Delta^*} (k_c^E + k_{se}^E n_v)}{k_{-\Delta^*} (k_c^E + k_{se}^E n_v) + (k_{+\Delta^*} + k_c^E + k_{se}^E n_v) (k_c^Z + k_{sc}^Z n_{v'})}. \quad (75b)$$

Together with the current of the non-radiative thermal reaction (54b), the current (74) fully determines the coarse-grained dynamics. The constants Φ in Eq. (75) express the probability that once a photon is absorbed, the photoisomerization happens. Importantly, these constants coincide with the “quantum yields” measured in experimental photochemistry, as discussed in Appendix A.

The single effective reaction obeys local detailed balance, as can be seen using Eqs. (60) and (75),

$$\frac{k_a^E n_\nu \Phi^{E \rightarrow Z}}{k_a^Z n_{\nu'} \Phi^{Z \rightarrow E}} = \exp\left\{\frac{\mu_\nu + \mu_E^\circ - \mu_{\nu'} - \mu_Z^\circ}{RT}\right\}. \quad (76)$$

This is due to the fact the adiabatic mechanism is tightly coupled, as it was shown for simple thermal reaction schemes.⁵⁹ As a consequence, the coarse-grained photochemical affinity can be written as

$$\mathcal{A} \equiv \mu_E + \mu_\nu - \mu_Z - \mu_{\nu'} = RT \ln\left\{\frac{k_a^E n_\nu \Phi^{E \rightarrow Z}[E]}{k_a^Z n_{\nu'} \Phi^{Z \rightarrow E}[Z]}\right\}. \quad (77)$$

Hence, also at the coarse-grained level, we correctly predict that the system reaches equilibrium [i.e., where all currents in the system vanish according to condition (31)] when the photons have the same chemical potential. When this is not the case, the system reaches a nonequilibrium steady state with current

$$\tilde{J} = \tilde{\Psi} = \tilde{J}_\Delta, \quad (78)$$

and combining Eqs. (54b), (74), (78), and (76), the concentration distribution is given by

$$\frac{[Z]}{[E]} = \frac{k_a^E n_\nu \Phi^{E \rightarrow Z} + k_{-\Delta}}{k_a^Z n_{\nu'} \Phi^{Z \rightarrow E} + k_{+\Delta}} = \mathcal{K} \exp\left\{\frac{\mu_E^\circ - \mu_Z^\circ}{RT}\right\}, \quad (79)$$

with

$$\mathcal{K} = \frac{\exp\left\{\frac{\mu_\nu - \mu_{\nu'}}{RT}\right\} + \frac{k_{+\Delta}}{k_a^Z n_{\nu'} \Phi^{Z \rightarrow E}}}{1 + \frac{k_{+\Delta}}{k_a^Z n_{\nu'} \Phi^{Z \rightarrow E}}}. \quad (80)$$

As in the case of Eq. (42), the ratio between steady-state concentrations in Eq. (79) is governed by the relative magnitude of the coarse-grained rate constants. In particular, it can be expressed as the product of the equilibrium ratio $[[Z]_{\text{eq}}/[E]_{\text{eq}}] = \exp\{(\mu_E^\circ - \mu_Z^\circ)/RT\}$ and a kinetic asymmetry factor (\mathcal{K}), as also found for molecular machine driven by Adenosine triphosphate (ATP) hydrolysis.⁸⁶ As expected from the discussion in Sec. III A, when photons ν and ν' have the same chemical potential $\mu_\nu - \mu_{\nu'} = 0$, the kinetic asymmetry factor is equal to one, and the ratio between steady-state concentrations coincides with the equilibrium ratio. Whether the steady-state ratio is larger or smaller than the equilibrium one is exclusively determined by μ_ν being larger or smaller than $\mu_{\nu'}$, respectively. As a consequence, consistent with Eq. (59), the sign of the steady-state current \tilde{J} in the adiabatic mechanism is only determined by the relative magnitude of μ_ν and $\mu_{\nu'}$ and not by the relative stability of species E and Z.

When all the photons come from the same thermal source at temperature T_r , the kinetic factor reads

$$\mathcal{K} = \frac{\exp\left\{\frac{\mu_\nu - \mu_{\nu'}}{RT}\left(1 - \frac{T}{T_r}\right)\right\} + \frac{k_{+\Delta}}{k_a^Z n_{\nu'} \Phi^{Z \rightarrow E}}}{1 + \frac{k_{+\Delta}}{k_a^Z n_{\nu'} \Phi^{Z \rightarrow E}}}. \quad (81)$$

In analogy with previous analysis of light-driven nonequilibrium chemical systems,⁸⁷ in the limit where photochemical steps are much faster than thermally induced ones ($k_{+\Delta} \ll k_a^Z n_{\nu'} \Phi^{Z \rightarrow E}$), we find

that the steady-state concentration distribution is controlled by the Carnot efficiency factor $1 - \frac{T}{T_r}$,

$$\frac{[Z]}{[E]} \approx \exp\left\{\frac{\mu_E^\circ - \mu_Z^\circ}{RT}\right\} \exp\left\{\frac{\mu_\nu - \mu_{\nu'}}{RT}\left(1 - \frac{T}{T_r}\right)\right\}. \quad (82)$$

When $T_r = T$, the system stays in thermal equilibrium, while in the limit of $T_r \rightarrow \infty$, the concentration distribution depends on the energy difference between the photons ν and ν' (i.e., on the energy differences between the species in the ground and in the excited state).⁸⁷ Instead, when the thermally induced processes are much faster than the light-induced ones (e.g., $k_{+\Delta} \gg k_a^Z n_{\nu'} \Phi^{Z \rightarrow E}$ and $k_{-\Delta} \gg k_a^E n_\nu \Phi^{E \rightarrow Z}$), the concentration distribution at steady state approximates the equilibrium one, $\mathcal{K} \simeq 1$.

Turning to thermodynamics, the coarse-grained entropy production for the coarse-grained adiabatic mechanism reads

$$\dot{\Sigma}^{\text{cg}} = \Psi \frac{\mathcal{A}}{T} + J_\Delta \frac{A_\Delta}{T}. \quad (83)$$

Provided that the timescale separation hypothesis leading to condition (73) holds, Eq. (83) coincides with Eq. (57), and the full entropy production can be quantified from the coarse-grained description: $\dot{\Sigma}^{\text{cg}} = \dot{\Sigma}$.⁶⁰ At steady state, where Eq. (78) applies, we find

$$T \dot{\Sigma}^{\text{cg}} = T \dot{\Sigma} = \tilde{J}(\mu_\nu - \mu_{\nu'}), \quad (84)$$

which exactly coincides with Eq. (59) regardless of how good is the timescale separation hypothesis.⁵⁹

2. Coarse-grained diabatic mechanism

In the diabatic mechanism depicted in Fig. 5, conditions (73) allow the coarse-graining of the four photochemical pathways $E \xleftrightarrow{\nu} E^* \xrightarrow{\nu'} Z$, $E \xrightarrow{\nu} E^* \xrightarrow{\nu'} E$, $Z \xleftrightarrow{\nu'} Z^* \xrightarrow{\nu} E$, and $Z \xleftrightarrow{\nu'} Z^* \xrightarrow{\nu} Z$ into four effective reactions of the kind $E \xleftrightarrow{\nu} Z$, $E \xleftrightarrow{\nu} E$, $Z \xleftrightarrow{\nu'} E$, and $Z \xleftrightarrow{\nu'} Z$, respectively. The corresponding coarse-grained currents read

$$\Psi_\nu^{E \rightarrow Z} = \Phi^{E^* \rightarrow Z} k_a^E n_\nu [E] - \Gamma_\nu^{E \rightarrow Z}, \quad (85a)$$

$$\Psi_\nu^{E \rightarrow E} = \Phi^{E^* \rightarrow E} k_a^E n_\nu [E] - \Gamma_\nu^{E \rightarrow E}, \quad (85b)$$

$$\Psi_{\nu'}^{Z \rightarrow E} = \Phi^{Z^* \rightarrow E} k_a^Z n_{\nu'} [Z] - \Gamma_{\nu'}^{Z \rightarrow E}, \quad (85c)$$

$$\Psi_{\nu'}^{Z \rightarrow Z} = \Phi^{Z^* \rightarrow Z} k_a^Z n_{\nu'} [Z] - \Gamma_{\nu'}^{Z \rightarrow Z}. \quad (85d)$$

All these four effective reactions coarse-grain pathways in which either E or Z absorb a photon and then thermally quench toward either the same or the isomeric ground state. As before, the Φ terms express the quantum yields of the various processes, while the Γ terms are proportional to the (very small) backward-quenching thermal rates and are therefore experimentally negligible. Expressions of both Φ s and Γ s are reported in Appendix B. The overall net rate of E to Z conversion due to photochemical processes is given by the difference

$$\Psi_\nu^{E \rightarrow Z} - \Psi_{\nu'}^{Z \rightarrow E} \approx \Phi^{E^* \rightarrow Z} k_a^E n_\nu [E] - \Phi^{Z^* \rightarrow E} k_a^Z n_{\nu'} [Z], \quad (86)$$

while the net absorption currents can be expressed as

$$J_v = \Psi_v^{E \rightarrow Z} + \Psi_v^{E \rightarrow E} \approx (\Phi^{E^* \rightarrow Z} + \Phi^{E^* \rightarrow E}) k_a^E n_v [E], \quad (87a)$$

$$J_{v'} = \Psi_{v'}^{Z \rightarrow E} + \Psi_{v'}^{Z \rightarrow Z} \approx (\Phi^{Z^* \rightarrow E} + \Phi^{Z^* \rightarrow Z}) k_a^Z n_{v'} [Z]. \quad (87b)$$

At steady state, the following equality holds:

$$\bar{J}_\Delta = \bar{\Psi}_v^{E \rightarrow Z} - \bar{\Psi}_{v'}^{Z \rightarrow E}, \quad (88)$$

and the concentration distribution is approximated by

$$\frac{[Z]}{[E]} \approx \frac{k_a^E n_v \Phi^{E^* \rightarrow Z} + k_{-\Delta}}{k_a^Z n_{v'} \Phi^{Z^* \rightarrow E} + k_{+\Delta}}. \quad (89)$$

We do not report the full analytical expression, which can be obtained from Eqs. (67), (85), and (88).

In contrast to the adiabatic mechanism, coarse-grained reactions (85) for the diabatic mechanism do not satisfy any local detailed balance property. This is due to the fact that the diabatic mechanism is not tightly coupled in analogy to what was shown for reaction schemes modeling active membrane transport.⁵⁹ As a consequence, in the diabatic mechanism, the steady-state distribution in Eq. (89) cannot be rewritten as the product of the equilibrium distribution times a kinetic asymmetry factor independent of the relative stability of E and Z [cf. Eq. (79)]. In fact, the steady-state distribution in Eq. (89) depends on many details of the systems (states stability included) and it is not solely controlled by the relative magnitude of the photons' chemical potentials. Moreover, thermodynamic affinities can be solely defined in terms of chemical potentials, since they do not equate the log ratio of forward and backward currents [cf. Eq. (77)],

$$\begin{aligned} \mathcal{A}_v^{E \rightarrow Z} &= \mu_E + \mu_v - \mu_Z, & \mathcal{A}_v^{E \rightarrow E} &= \mu_v, \\ \mathcal{A}_{v'}^{Z \rightarrow E} &= \mu_Z + \mu_{v'} - \mu_E, & \mathcal{A}_{v'}^{Z \rightarrow Z} &= \mu_{v'}. \end{aligned} \quad (90)$$

In analogy with Eq. (83), the entropy production at the coarse-grained level reads^{59,60}

$$\begin{aligned} \dot{\Sigma}^{\text{cg}} &= J_\Delta \frac{A_\Delta}{T} + \Psi_v^{E \rightarrow Z} \frac{\mathcal{A}_v^{E \rightarrow Z}}{T} + \Psi_{v'}^{Z \rightarrow E} \frac{\mathcal{A}_{v'}^{Z \rightarrow E}}{T} \\ &+ \Psi_v^{E \rightarrow E} \frac{\mathcal{A}_v^{E \rightarrow E}}{T} + \Psi_{v'}^{Z \rightarrow Z} \frac{\mathcal{A}_{v'}^{Z \rightarrow Z}}{T}. \end{aligned} \quad (91)$$

At the steady state, using Eqs. (87) and (88), we find

$$\begin{aligned} T \dot{\Sigma}^{\text{cg}} &= (\bar{\Psi}_v^{E \rightarrow Z} + \bar{\Psi}_v^{E \rightarrow E}) \mu_v + (\bar{\Psi}_{v'}^{Z \rightarrow E} + \bar{\Psi}_{v'}^{Z \rightarrow Z}) \mu_{v'} \\ &= \bar{J}_v \mu_v + \bar{J}_{v'} \mu_{v'}. \end{aligned} \quad (92)$$

As for the adiabatic mechanism, the steady-state coarse-grained entropy production [Eq. (92)] coincides with the fine-grained one [Eq. (70)],⁵⁹ while in transient, regimes the effectiveness of the coarse-grained estimation depends on the validity of the timescale separation approximation [Eq. (73)].⁶⁰

IV. CONCLUSIONS AND DISCUSSION

In this paper, we introduced a modern formulation of nonequilibrium thermodynamics for systems made of chemicals interacting with incoherent radiation. Our description leverages the advances made for purely thermal chemical-reactions systems in isothermal ideal dilute solutions.^{47,49}

In Sec. II, we focused on a simple photophysical mechanism in which photons may excite a chemical species. By using the property of local detailed balance, viz., Einstein relations, we established a rigorous and thermodynamically consistent description of such excitation dynamics. The main outcome of our description follows as we show that the nonequilibrium free energy of chemicals plus radiation F [Eq. (36)] acts as a Lyapunov function for a closed system relaxing to thermal equilibrium (Sec. II D): F never increases and reaches its minimum at thermal equilibrium. Mathematically, this proves that the equilibrium state is globally stable.

While we focused on chemical systems, the approach developed in Sec. II B is general and can be used to study the thermodynamics of any kind of light-reacting matter, provided that the local equilibrium assumption holds. This means that the sole mechanism generating dissipation are the reacting events mediating changes in the molar concentration of photons and molecules. In the absence of reacting events, the molecules are at equilibrium in solution and the photons are at (possibly another) equilibrium in the volume occupied by the solution. Our treatment is also consistent with previous results obtained in more specific contexts, including semiconductor physics.^{42,44}

The results derived in Sec. III on basic photochemical mechanisms describing light-induced conversion of a chemical species (E) into another one (Z) are compatible with both experimental observations and previous theoretical descriptions. The adiabatic mechanism analyzed in Sec. III A is usually employed to model excited-state proton-transfer reactions, where E, Z, and their excited counterparts differ by the location of an acidic proton.^{4,80,81} As first pointed out by Förster,⁸² a shift in the acid–base equilibrium distribution between E and Z under light irradiation from a single source cannot be reached if the two excitation frequencies ν and ν' are equal.^{4,80,81} This observation is predicted by Eq. (59) in the general case and by Eq. (63) when the light source is a thermal one. Indeed, in both cases, there is no steady-state entropy production when $\nu = \nu'$, meaning that the system will stay at the acid–base equilibrium when irradiated. In addition, a generalization of the well-known concept of Förster cycle^{80,82} is offered by Eq. (60), which allows us to link ratios between the kinetic constants of the thermally induced non-radiative steps with the differences in excitation frequencies and standard entropies of the chemical species involved.

The diabatic mechanism analyzed in Sec. III B models the most common mechanism for photochemical E–Z isomerizations, where E and Z are, for instance, two stable configurations of an alkene or an azobenzene derivative.^{4,28,29} A major biological example of this kind of processes is the primary event in vision.⁸ For an example drawn from synthetic chemistry, we refer to the minimalist two-step light-driven molecular motor proposed in Ref. 84, whose first experimental implementation is based on imine groups^{88,89} and for which the scheme in Fig. 5 is a realistic model. Contrary to the adiabatic case, we demonstrated that a photoisomerization following the diabatic mechanism can be driven out-of-equilibrium with only one

frequency of light. This is in line with common models describing synthetic light-driven molecular motors.^{61,62}

The typically very high reactivity of excited intermediates in fine-grained photochemical reaction schemes makes the coarse-grained level of description developed in Sec. III C the most natural framework to compare theory with experiments. Indeed, coarse-grained expressions (74) and (86) are usually employed to fit experimental data and measure quantum yields.⁵⁸ Crucially, despite the fact that the coarse-grained currents (74) and (86) read identically, only for the adiabatic mechanism, one can estimate its dissipative contribution—the affinity—using the local detailed balance property at the level of coarse-grained fluxes [see Eq. (77)]. Contrary, for photochemical processes following the diabatic mechanism (the most common mechanism for photo-isomerizations^{4,28}), the sole knowledge of the quantum yields appearing in the coarse-grained net current (86) is not enough to estimate the dissipation. Instead, one needs to measure also the net absorption currents in Eq. (87), which allow us to apply either Eqs. (91) or (92) for computing the dissipation.

V. PERSPECTIVES

Our results constitute a step forward toward a unified nonequilibrium thermodynamics for chemical systems powered by different free-energy sources, such as chemostats and light. They also open the way to performance studies on free-energy transduction and storage in finite-time and far from equilibrium, which were inaccessible before.

Building on the results developed in this paper, we plan next to generalize our thermodynamic theory to continuous absorption/emission spectra, as well as elementary bimolecular photo-physical (e.g., bimolecular quenchings) and photochemical (e.g., photodissociations) light-induced processes. Another limitation to overcome in order to model real experiments is to go beyond the assumption that the radiostats are able to fix the same molar concentration of photons in the whole system. While this is a good approximation for ideal dilute solutions irradiated uniformly, photochemical experiments are usually conducted with light coming from one side of the system. This may give rise to the nonlinear shielding effect.⁹⁰ We anticipate that such extensions will allow to study interesting nonequilibrium phenomena such as bistability and oscillations in photochemical systems,^{90–92} whose thermodynamic cost and constraints have never been compared with the cost of chemically driven oscillations.⁵⁴

ACKNOWLEDGMENTS

The authors are thankful to F. Avanzini and D. Accomasso for insightful discussions. E.P. and M.E. acknowledge funding from the European Research Council, project NanoThermo (ERC-2015-CoG Agreement No. 681456). R.R. was funded by *The Martin A. and Helen Chooljian* Membership in Biology at the Institute for Advanced Study.

APPENDIX A: ON THE QUANTUM YIELD

The standard textbook definition^{4,58} of quantum yield for the direct $E^* \rightarrow Z^* \rightarrow Z$ process reads

$$\Phi_0^{E \rightarrow Z} = \frac{k_+^*}{k_+^* + (k_e^E + k_{se}^E n_\nu)} \frac{(k_e^Z + k_{se}^Z n_\nu)}{k_-^* + (k_e^Z + k_{se}^Z n_\nu)}, \quad (A1)$$

but we can also consider an indirect process of the kind $E^* \rightarrow Z^* \rightarrow E^* \rightarrow Z^* \rightarrow Z$, for which the quantum yield reads

$$\Phi_1^{E \rightarrow Z} = \Phi_0^{E \rightarrow Z} \frac{k_-^*}{k_-^* + (k_e^Z + k_{se}^Z n_\nu)} \frac{k_+^*}{k_+^* + (k_e^E + k_{se}^E n_\nu)}. \quad (A2)$$

The global quantum yield of the photoisomerization is therefore the sum of the quantum yields of all the possible paths going from E^* to Z . By taking all of them into account, we correctly find

$$\begin{aligned} \Phi^{E \rightarrow Z} &= \sum_{i=0}^{\infty} \Phi_i^{E \rightarrow Z} \\ &= \Phi_0^{E \rightarrow Z} \sum_{i=0}^{\infty} \left(\frac{k_-^*}{k_-^* + (k_e^Z + k_{se}^Z n_\nu)} \frac{k_+^*}{k_+^* + (k_e^E + k_{se}^E n_\nu)} \right)^i. \end{aligned} \quad (A3)$$

APPENDIX B: COARSE-GRAINED Φ 's AND Γ 's TERMS FOR THE DIABATIC MECHANISM

The quantum yields in the coarse-grained version of the diabatic mechanism read

$$\Phi^{E^* \rightarrow Z} = \frac{k_{+q}^{E^* \rightarrow Z}}{k_{+q}^{E^* \rightarrow Z} + k_{+q}^{E^* \rightarrow E} + (k_e^E + k_{se}^E n_\nu)}, \quad (B1a)$$

$$\Phi^{E^* \rightarrow E} = \frac{k_{+q}^{E^* \rightarrow E}}{k_{+q}^{E^* \rightarrow Z} + k_{+q}^{E^* \rightarrow E} + (k_e^E + k_{se}^E n_\nu)}, \quad (B1b)$$

$$\Phi^{Z^* \rightarrow E} = \frac{k_{+q}^{Z^* \rightarrow E}}{k_{+q}^{Z^* \rightarrow E} + k_{+q}^{Z^* \rightarrow Z} + (k_e^Z + k_{se}^Z n_\nu)}, \quad (B1c)$$

$$\Phi^{Z^* \rightarrow Z} = \frac{k_{+q}^{Z^* \rightarrow Z}}{k_{+q}^{Z^* \rightarrow E} + k_{+q}^{Z^* \rightarrow Z} + (k_e^Z + k_{se}^Z n_\nu)}, \quad (B1d)$$

where k_q 's are the quenching thermal rates going from one excited species to its ground state isomer. Quantum yields here express the probabilities that given that either E or Z absorbed a photon, the photoisomerization happens or the excited state quenches back to the initial ground state. The Γ terms are proportional to the (very small) backward quenching thermal rates,

$$\Gamma_\nu^{E \rightarrow Z} = \frac{[Z]k_{-q}^{Z \rightarrow E^*} (k_e^E + k_{se}^E n_\nu + k_{+q}^{E^* \rightarrow E}) - [E]k_{-q}^{E \rightarrow E^*} k_{+q}^{E^* \rightarrow Z}}{k_{+q}^{E^* \rightarrow Z} + k_{+q}^{E^* \rightarrow E} + (k_e^E + k_{se}^E n_\nu)}, \quad (B2a)$$

$$\Gamma_\nu^{E \rightarrow E} = \frac{[E]k_{-q}^{E \rightarrow E^*} (k_e^E + k_{se}^E n_\nu + k_{+q}^{E^* \rightarrow Z}) - [Z]k_{-q}^{Z \rightarrow E^*} k_{+q}^{E^* \rightarrow E}}{k_{+q}^{E^* \rightarrow Z} + k_{+q}^{E^* \rightarrow E} + (k_e^E + k_{se}^E n_\nu)}, \quad (B2b)$$

$$\Gamma_\nu^{Z \rightarrow E} = \frac{[E]k_{-q}^{E \rightarrow Z^*} (k_e^Z + k_{se}^Z n_\nu + k_{+q}^{Z^* \rightarrow Z}) - [Z]k_{-q}^{Z \rightarrow Z^*} k_{+q}^{Z^* \rightarrow E}}{k_{+q}^{Z^* \rightarrow E} + k_{+q}^{Z^* \rightarrow Z} + (k_e^Z + k_{se}^Z n_\nu)}, \quad (B2c)$$

$$\Gamma_\nu^{Z \rightarrow Z} = \frac{[Z]k_{-q}^{Z \rightarrow Z^*} (k_e^Z + k_{se}^Z n_\nu + k_{+q}^{Z^* \rightarrow Z}) - [E]k_{-q}^{E \rightarrow Z^*} k_{+q}^{Z^* \rightarrow Z}}{k_{+q}^{Z^* \rightarrow E} + k_{+q}^{Z^* \rightarrow Z} + (k_e^Z + k_{se}^Z n_\nu)}, \quad (B2d)$$

and are therefore not experimentally accessible with standard techniques.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

- ¹M. Ghil and V. Lucarini, “The physics of climate variability and climate change,” *Rev. Mod. Phys.* **92**, 035002 (2020).
- ²M. Calvin, “The photosynthetic carbon cycle,” *J. Chem. Soc.* 1895–1915 (1956).
- ³R. E. Blankenship, *Molecular Mechanisms of Photosynthesis*, 2nd ed. (Wiley-Blackwell, 2014).
- ⁴V. Balzani, P. Ceroni, and A. Juris, *Photochemistry and Photophysics: Concepts, Research, Applications*, 1st ed. (Wiley-VCH Verlag, 2014).
- ⁵J. Barber, “Photosynthetic energy conversion: Natural and artificial,” *Chem. Soc. Rev.* **38**, 185–196 (2009).
- ⁶I. Navizet, Y.-J. Liu, N. Ferré, D. Roca-Sanjuán, and R. Lindh, “The chemistry of bioluminescence: An analysis of chemical functionalities,” *ChemPhysChem* **12**, 3064–3076 (2011).
- ⁷T. Yoshizawa and G. Wald, “Pre-lumirhodopsin and the bleaching of visual pigments,” *Nature* **197**, 1279–1286 (1963).
- ⁸D. Polli, P. Altoè, O. Weingart, K. M. Spillane, C. Manzoni, D. Brida, G. Tomasello, G. Orlandi, P. Kukura, R. A. Mathies, M. Garavelli, and G. Cerullo, “Conical intersection dynamics of the primary photoisomerization event in vision,” *Nature* **467**, 440–443 (2010).
- ⁹D. Dattler, G. Fuks, J. Heiser, E. Moulin, A. Perrot, X. Yao, and N. Giuseppone, “Design of collective motions from synthetic molecular switches, rotors, and motors,” *Chem. Rev.* **120**, 310–433 (2020).
- ¹⁰N. Koumura, R. W. J. Zijlstra, R. A. van Delden, N. Harada, and B. L. Feringa, “Light-driven monodirectional molecular rotor,” *Nature* **401**, 152–155 (1999).
- ¹¹G. Ragazzon, M. Baroncini, S. Silvi, M. Venturi, and A. Credi, “Light-powered autonomous and directional molecular motion of a dissipative self-assembling system,” *Nat. Nanotechnol.* **10**, 70–75 (2015).
- ¹²J. Berná, D. A. Leigh, M. Lubomska, S. M. Mendoza, E. M. Pérez, P. Rudolf, G. Teobaldi, and F. Zerbetto, “Macroscopic transport by synthetic molecular machines,” *Nat. Mater.* **4**, 704–710 (2005).
- ¹³R. Eelkema, M. M. Pollard, J. Vicario, N. Katsonis, B. S. Ramon, C. W. M. Bastiaansen, D. J. Broer, and B. L. Feringa, “Nanomotor rotates microscale objects,” *Nature* **440**, 163 (2006).
- ¹⁴Q. Li, G. Fuks, E. Moulin, M. Maaloum, M. Rawiso, I. Kulic, J. T. Foy, and N. Giuseppone, “Macroscopic contraction of a gel induced by the integrated motion of light-driven molecular motors,” *Nat. Nanotechnol.* **10**, 161–165 (2015).
- ¹⁵J. T. Foy, Q. Li, A. Goujon, J.-R. Colard-Itté, G. Fuks, E. Moulin, O. Schiffmann, D. Dattler, D. P. Funeriu, and N. Giuseppone, “Dual-light control of nanomachines that integrate motor and modulator subunits,” *Nat. Nanotechnol.* **12**, 540–545 (2017).
- ¹⁶V. Serreli, C.-F. Lee, E. R. Kay, and D. A. Leigh, “A molecular information ratchet,” *Nature* **445**, 523–527 (2007).
- ¹⁷G. Steinberg-Yfrach, P. A. Liddell, S.-C. Hung, A. L. Moore, D. Gust, and T. A. Moore, “Conversion of light energy to proton potential in liposomes by artificial photosynthetic reaction centres,” *Nature* **385**, 239–241 (1997).
- ¹⁸I. M. Bennett, H. M. V. Farfano, F. Bogani, A. Primak, P. A. Liddell, L. Otero, L. Sereno, J. J. Silber, A. L. Moore, T. A. Moore, and D. Gust, “Active transport of Ca²⁺ by an artificial photosynthetic membrane,” *Nature* **420**, 398–401 (2002).
- ¹⁹S. Bhosale, A. L. Sisson, P. Talukdar, A. Fürstenberg, N. Banerji, E. Vauthey, G. Bollot, J. Mareda, C. Röger, F. Würthner, N. Sakai, and S. Matile, “Photoproduction of proton gradients with π -stacked fluorophore scaffolds in lipid bilayers,” *Science* **313**, 84–86 (2006).
- ²⁰T. von Grothuss, “Auszug aus vier abhandlungen physikalisch-chemischen inhalts,” *Ann. Phys.* **61**, 50–74 (1819).
- ²¹J. W. Draper, “On a change produced by exposure to the beams of the sun, in the properties of an elementary substance,” *Philos. Mag.* **23**, 388 (1843).
- ²²C. G. Lemoine, “Etudes quantitatives sur l’action chimique de la lumière pour la décomposition mutuelle de l’acide oxalique et du chlorure ferrique,” *Ann. Chim. Phys.* **VI**, 433–540 (1895).
- ²³G. Ciamician, “Sur les actions chimiques de la lumière,” *Bull. Soc. Chim. Fr.* **IV**, i–xxvii (1908).
- ²⁴K. K. Rohatgi-Mukherjee, *Fundamentals of Photochemistry* (New Age International, 1978).
- ²⁵A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, 1st ed. (Dover Publications, Inc., 1996).
- ²⁶E. J. Bowen, *The Chemical Aspects of Light* (Clarendon Press, 1946).
- ²⁷I. Fleming, *Molecular Orbitals and Organic Chemical Reactions* (John Wiley & Sons, Ltd., 2010).
- ²⁸M. Boggio-Pasqua, *Computational Mechanistic Photochemistry: The Central Role of Conical Intersections*, Habilitation à Diriger des Recherches (Université Toulouse III, 2015).
- ²⁹M. Persico and G. Granucci, *Photochemistry: A Modern Theoretical Perspective*, 1st ed. (Springer International Publishing, 2018).
- ³⁰M. A. Robb, M. Garavelli, M. Olivucci, and F. Bernardi, “A computational strategy for organic photochemistry,” *Rev. Comput. Chem.* **15**, 87–146 (2000).
- ³¹G. Granucci, M. Persico, and A. Toniolo, “Direct semiclassical simulation of photochemical processes with semiempirical wave functions,” *J. Chem. Phys.* **114**, 10608–10615 (2001).
- ³²G. Kirchhoff, “Ueber das verhältniss zwischen dem emissionsvermögen und dem absorptionsvermögen der körper für wärme und licht,” *Ann. Phys.* **185**, 275–301 (1860).
- ³³W. Wien, “Ueber die energievertheilung im emissionsspectrum eines schwarzen körpers,” *Ann. Phys.* **294**, 662–669 (1896).
- ³⁴Lord Rayleigh, “Remarks upon the law of complete radiation,” *Philos. Mag.* **49**, 539–540 (1896).
- ³⁵M. Planck, “Ueber das gesetz der energieverteilung im normalspectrum,” *Ann. Phys.* **309**, 553–563 (1901).
- ³⁶A. Einstein, “Ueber einen die erzeugung und verwandlung des liches betreffenden heuristischen gesichtspunkt,” *Ann. Phys.* **322**, 132–148 (1905).
- ³⁷R. T. Ross, “Thermodynamic limitations on the conversion of radiant energy into work,” *J. Chem. Phys.* **45**, 1–7 (1966).
- ³⁸R. T. Ross and M. Calvin, “Thermodynamics of light emission and free-energy storage in photosynthesis,” *Biophys. J.* **7**, 595–614 (1967).
- ³⁹R. T. Ross, “Some thermodynamics of photochemical systems,” *J. Chem. Phys.* **46**, 4590–4593 (1967).
- ⁴⁰P. Würfel, “The chemical potential of radiation,” *J. Phys. C: Solid State Phys.* **15**, 3967–3985 (1982).
- ⁴¹F. Herrmann and P. Würfel, “Light with nonzero chemical potential,” *Am. J. Phys.* **73**, 717–721 (2005).
- ⁴²P. Würfel and U. Würfel, *Physics of Solar Cells: From Basic Principles to Advanced Concepts*, 3rd ed. (Wiley VCH, 2016).
- ⁴³H. Ries and A. J. McEvoy, “Chemical potential and temperature of light,” *J. Photochem. Photobiol., A* **59**, 11–18 (1991).
- ⁴⁴P. Gräber and G. Milazzo, *Bioenergetics*, 1st ed. (Birkhäuser Basel, 1997).
- ⁴⁵P. Gaspard, “Fluctuation theorem for nonequilibrium reactions,” *J. Chem. Phys.* **120**, 8898–8905 (2004).
- ⁴⁶H. Qian and D. A. Beard, “Thermodynamics of stoichiometric biochemical networks in living systems far from equilibrium,” *Biophys. Chem.* **114**, 213–220 (2005).
- ⁴⁷M. Polettini and M. Esposito, “Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws,” *J. Chem. Phys.* **141**, 024117 (2014).
- ⁴⁸H. Ge and H. Qian, “Nonequilibrium thermodynamic formalism of nonlinear chemical reaction systems with Waage–Guldberg’s law of mass action,” *Chem. Phys.* **472**, 241–248 (2016).
- ⁴⁹R. Rao and M. Esposito, “Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics,” *Phys. Rev. X* **6**, 041064 (2016).
- ⁵⁰R. Rao and M. Esposito, “Conservation laws and work fluctuation relations in chemical reaction networks,” *J. Chem. Phys.* **149**, 245101 (2018).

- ⁵¹E. Penocchio, R. Rao, and M. Esposito, "Thermodynamic efficiency in dissipative chemistry," *Nat. Commun.* **10**, 3865 (2019).
- ⁵²F. Avanzini, E. Penocchio, G. Falasco, and M. Esposito, "Nonequilibrium thermodynamics of non-ideal chemical reaction networks," *J. Chem. Phys.* **154**, 094114 (2021).
- ⁵³G. Falasco, R. Rao, and M. Esposito, "Information thermodynamics of Turing patterns," *Phys. Rev. Lett.* **121**, 108301 (2018).
- ⁵⁴F. Avanzini, G. Falasco, and M. Esposito, "Thermodynamics of chemical waves," *J. Chem. Phys.* **151**, 234103 (2019).
- ⁵⁵I. Prigogine, "Le domaine de validité de la thermodynamique des phénomènes irréversibles," *Physica* **15**, 272–284 (1949).
- ⁵⁶D. Kondepudi and I. Prigogine, *Modern Thermodynamics: From Heat Engines to Dissipative Structures*, 2nd ed. (Wiley, 2014).
- ⁵⁷A. Einstein, "Strahlungs-emission und absorption nach der quantentheorie," *Dtsch. Phys. Ges.* **18**, 318–323 (1916).
- ⁵⁸M. Montalti, A. Credi, L. Prodi, and T. Gandolfi, *Handbook of Photochemistry*, 3rd ed. (CRC Press, 2006).
- ⁵⁹A. Wachtel, R. Rao, and M. Esposito, "Thermodynamically consistent coarse graining of biocatalysts beyond Michaelis–Menten," *New J. Phys.* **20**, 042002 (2018).
- ⁶⁰F. Avanzini, G. Falasco, and M. Esposito, "Thermodynamics of non-elementary chemical reaction networks," *New J. Phys.* **22**, 093040 (2020).
- ⁶¹E. M. Geertsema, S. J. van der Molen, M. Martens, and B. L. Feringa, "Optimizing rotary processes in synthetic molecular motors," *Proc. Natl. Acad. Sci. U. S. A.* **106**, 16919–16924 (2009).
- ⁶²A. Sabatino, E. Penocchio, G. Ragazzon, A. Credi, and D. Frezzato, "Individual-molecule perspective analysis of chemical reaction networks: The case of a light-driven supramolecular pump," *Angew. Chem., Int. Ed.* **58**, 14341–14348 (2019).
- ⁶³T. Förster, "Primary photophysical processes," *Pure Appl. Chem.* **34**, 225–234 (1973).
- ⁶⁴C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Atom-Photon Interactions: Basic Processes and Applications* (Wiley VCH, 1998).
- ⁶⁵G. Meszéna and H. V. Westerhoff, "Non-equilibrium thermodynamics of light absorption," *J. Phys. A: Math. Gen.* **32**, 301–311 (1999).
- ⁶⁶Z. Chen, S. Mo, and P. Hu, "Recent progress in thermodynamics of radiation—exergy of radiation, effective temperature of photon and entropy constant of photon," *Sci. China, Ser. E: Technol. Sci.* **51**, 1096 (2008).
- ⁶⁷S. Katz, J. L. Lebowitz, and H. Spohn, "Phase transitions in stationary nonequilibrium states of model lattice systems," *Phys. Rev. B* **28**, 1655–1658 (1983).
- ⁶⁸C. Maes, "Local detailed balance," *SciPost Phys. Lect. Notes* **32**, 1–17 (2021).
- ⁶⁹G. Falasco and M. Esposito, "Local detailed balance across scales: From diffusions to jump processes and beyond," *Phys. Rev. E* **103**, 042114 (2021).
- ⁷⁰J. L. Lebowitz and P. G. Bergmann, "New approach to nonequilibrium process," *Phys. Rev.* **99**, 578–587 (1955).
- ⁷¹T. L. Hill, *Free Energy Transduction in Biology* (Academic Press, New York, 1977).
- ⁷²J. Schnakenberg, "Network theory of microscopic and macroscopic behavior of master equation systems," *Rev. Mod. Phys.* **48**, 571–585 (1976).
- ⁷³C. Y. Mou, J. I. Luo, and G. Nocolis, "Stochastic thermodynamics of nonequilibrium steady states in chemical reaction systems," *J. Chem. Phys.* **84**, 7011–7017 (1986).
- ⁷⁴R. Wegscheider, "Über simultane gleichgewichte und die beziehungen zwischen thermodynamik und reaktionskinetik homogener systeme," *Z. Phys. Chem.* **39U**, 257 (1902).
- ⁷⁵D. G. Blackmond, "'If pigs could fly' chemistry: A tutorial on the principle of microscopic reversibility," *Angew. Chem., Int. Ed.* **48**, 2648–2654 (2009).
- ⁷⁶R. D. Astumian, "Microscopic reversibility as the organizing principle of molecular machines," *Nat. Nanotechnol.* **7**, 684–688 (2012).
- ⁷⁷R. D. Astumian, "Trajectory and cycle-based thermodynamics and kinetics of molecular machines: The importance of microscopic reversibility," *Acc. Chem. Res.* **51**, 2653–2661 (2018).
- ⁷⁸T. Förster, "Diabatic and adiabatic processes in photochemistry," *Pure Appl. Chem.* **24**, 443–450 (1970).
- ⁷⁹P. Muller, "Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)," *Pure Appl. Chem.* **66**, 1077–1184 (1994).
- ⁸⁰Z. R. Grabowski and W. Rubaszewska, "Generalised Förster cycle. Thermodynamic and extrathermodynamic relationships between proton transfer, electron transfer and electronic excitation," *J. Chem. Soc., Faraday Trans. 1* **73**, 11–28 (1977).
- ⁸¹L. G. Arnaut and S. J. Formosinho, "Excited-state proton transfer reactions I. Fundamentals and intermolecular reactions," *J. Photochem. Photobiol., A* **75**, 1–20 (1993).
- ⁸²T. Förster, "Die ph-abhängigkeit der fluoreszenz von naphthalinderivaten," *Z. Elektrochem. Angew. Phys. Chem.* **54**, 531–535 (1950).
- ⁸³M. Esposito, D. Lindenberg, and C. Van den Broeck, "Universality of efficiency at maximum power," *Phys. Rev. Lett.* **102**, 130602 (2009).
- ⁸⁴J.-M. Lehn, "Conjecture: Imines as unidirectional photodriven molecular motors—motional and constitutional dynamic devices," *Chem. - Eur. J.* **12**, 5910–5915 (2006).
- ⁸⁵K. J. Laidler, *Chemical Kinetics* (HarperCollins Publishers, New York, 1987).
- ⁸⁶R. D. Astumian and M. Bier, "Mechanochemical coupling of the motion of molecular motors to ATP hydrolysis," *Biophys. J.* **70**, 637–653 (1996).
- ⁸⁷R. D. Astumian, "Optical vs. chemical driving for molecular machines," *Faraday Discuss.* **195**, 583–597 (2016).
- ⁸⁸L. Greb and J.-M. Lehn, "Light-driven molecular motors: Imines as four-step or two-step unidirectional rotors," *J. Am. Chem. Soc.* **136**, 13114–13117 (2014).
- ⁸⁹L. Greb, A. Eichhöfer, and J. M. Lehn, "Synthetic molecular motors: Thermal inversion and directional photoinduced C=N bond rotation of camphorquinone imines," *Angew. Chem., Int. Ed.* **54**, 14345–14348 (2015).
- ⁹⁰B. Borderie, D. Lavabre, J. C. Micheau, and J. P. Laplante, "Onlinear dynamics, multiple steady states, and oscillations in photochemistry," *J. Phys. Chem.* **96**, 2953–2961 (1992).
- ⁹¹A. Nitzan and J. Ross, "Oscillations, multiple steady states, and instabilities in illuminated systems," *J. Chem. Phys.* **59**, 241–250 (1973).
- ⁹²P. L. Gentili, B. Bartolomei, and J.-C. Micheau, "Light-driven artificial neuron models based on photoswitchable systems," *Dyes Pigm.* **187**, 109086 (2021).

The following article is reprinted from
[S. CORRA, M. T. BAKIĆ, J. GROPPI, M. BARONCINI, S. SILVI, E. PENOCCHIO,
M. ESPOSITO and A. CREDI, *Nature Nanotechnology* (2022)]
under the conditions of the Springer Nature Limited license².

The page numbers placed in the outer margins provide a continuous pagination throughout the thesis.

² <https://www.nature.com/nature-portfolio/reprints-and-permissions/permissions-requests>



Kinetic and energetic insights into the dissipative non-equilibrium operation of an autonomous light-powered supramolecular pump

Stefano Corra ^{1,2,6}, Marina Tranfić Bakić ^{1,2,6}, Jessica Groppi ¹, Massimo Baroncini ^{1,3}, Serena Silvi^{1,4}, Emanuele Penocchio ⁵, Massimiliano Esposito⁵ and Alberto Credi^{1,2}✉

Natural and artificial autonomous molecular machines operate by constantly dissipating energy coming from an external source to maintain a non-equilibrium state. Quantitative thermodynamic characterization of these dissipative states is highly challenging as they exist only as long as energy is provided. Here we report on the detailed physicochemical characterization of the dissipative operation of a supramolecular pump. The pump transduces light energy into chemical energy by bringing self-assembly reactions to non-equilibrium steady states. The composition of the system under light irradiation was followed in real time by ¹H NMR for four different irradiation intensities. The experimental composition and photon flow were then fed into a theoretical model describing the non-equilibrium dissipation and the energy storage at the steady state. We quantitatively probed the relationship between the light energy input and the deviation of the dissipative state from thermodynamic equilibrium in this artificial system. Our results provide a testing ground for newly developed theoretical models for photoactivated artificial molecular machines operating away from thermodynamic equilibrium.

Natural and artificial molecular machines, such as molecular motors and pumps, can be rationally described as networks of reactions where energy-harvesting chemical^{1–10} or photochemical^{11–20} reactions are coupled with large-amplitude intramolecular motions or self-assembly processes that can be driven towards a non-equilibrium state (Fig. 1a)^{21–26}.

The determination of the speciation in the out-of-equilibrium state is highly desirable, as several thermodynamic and kinetic parameters of the machine operation could be extracted. However, the characterization of dissipative states is extremely challenging as it requires any analysis to be performed during operation, that is in the presence of the fuel-to-waste conversion^{27–29}.

We recently reported on a series of light-powered molecular motors based on a pseudorotaxane-complex architecture. The operational principle relies on the reversible photoisomerization of the azobenzene moiety of the axle which enables a flashing energy ratchet mechanism (Fig. 1b). As the system realizes linear-directed transport of one molecular component relative to another, as a result of the Brownian ratchet mechanism, we refer to it as a supramolecular pump. Under continuous irradiation, the crown ether macrocycle undergoes autonomous and directional cycling between the assembled and disassembled states^{21,30,31}. In the setup described here, our pump operates ‘freewheel’ and, unlike other systems^{8–10}, it cannot build up a concentration gradient. Nonetheless, its dissipative operation as a molecular motor was proved by comparing the thermodynamic and kinetic data for the two—*E* and *Z*—isomers, as well as by ¹H NMR while irradiating the sample solution inside the NMR probehead^{30–32}. Specifically, this technique allows for the real-time monitoring of the system composition at the dissipative steady state. Building upon these results, we focused here on the

physicochemical characterization of the operation cycle and dissipative steady state of a previously reported supramolecular pump. The cycling rate, quantum yield, energy storage and power conversion efficiency were derived from the experimental concentrations at different incident photon flows. Moreover, we probed a relationship between the light energy input and the deviation of the dissipative state from the equilibrium composition in this artificial system.

Structure and operation of the supramolecular pump

The supramolecular pump is composed of a dibenzo[24]crown-8 (DB24C8) ether macrocycle (**1**) and a molecular axle (**2**⁺), which comprises a secondary ammonium recognition site for **1** flanked by a photoswitchable azobenzene gate and a cyclopentyl ‘pseudostopper’ (Fig. 1c). Briefly, in acetonitrile the macrocycle encircles the secondary ammonium station of the axle to form a stable hydrogen-bonded rotaxane-type complex for both the *E* and *Z* isomers of **2**⁺. The *E* complex is slightly more stable than the *Z* complex ($K_a^E > K_a^Z$) and its formation occurs almost exclusively by slippage of the ring over the *E*-azobenzene moiety (Table 1). The non-photoactive pseudostopper was selected to ensure that the formation of the *Z* complex occurred almost exclusively by slippage of the ring over the cyclopentyl moiety^{30,31}.

The *E* → *Z* isomerization of the axle has, therefore, two key consequences: (1) the destabilization of the threaded complex and (2) the increase of the activation energy for the slippage of the ring over the azobenzene extremity^{30,31,33}. As a result, light-induced unidirectional transport of the rings over the pump occurs according to an energy ratchet mechanism (Fig. 1b). As both *E* and *Z* isomers of azobenzene are photoreactive and exhibit highly overlapped absorption spectra, photons of the same wavelength can trigger both *E* → *Z* and

¹CLAN-Center for Light Activated Nanostructures, Istituto ISOF-CNR, Bologna, Italy. ²Dipartimento di Chimica Industriale ‘Toso Montanari’, Università di Bologna, Bologna, Italy. ³Dipartimento di Scienze e Tecnologie Agro-alimentari, Università di Bologna, Bologna, Italy. ⁴Dipartimento di Chimica ‘G. Ciamician’, Università di Bologna, Bologna, Italy. ⁵Department of Physics and Materials Science, University of Luxembourg, Luxembourg City, Luxembourg. ⁶These authors contributed equally: Stefano Corra, Marina Tranfić Bakić. ✉e-mail: alberto.credi@unibo.it

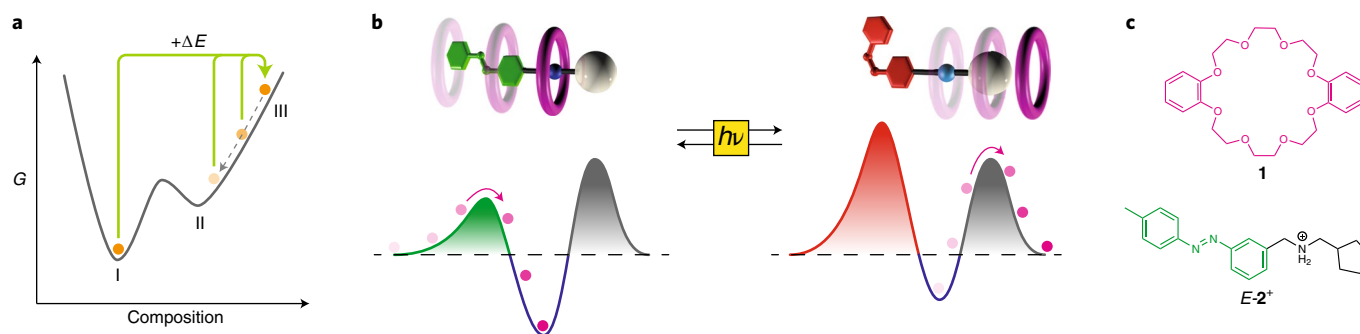


Fig. 1 | Simplified energy diagrams of operation of the pump and molecular structures of the components. **a**, Idealized landscape of the total Gibbs free energy (G) of a chemical system, showing the global thermodynamic minimum that corresponds to the thermal equilibrium (I), a local equilibrium state (II) and a dissipative non-equilibrium state (III). State III exists only if energy is continuously supplied to the system (solid green arrows) to prevent its relaxation to state II (dashed grey arrow). **b**, Schematic representation of the relative unidirectional translation of molecular ring and axle components triggered by light, that is, a supramolecular pump powered by light. The simplified potential energy curves (free energy versus ring–axle distance) corresponding to the two structures illustrate the energy ratchet mechanism that rectifies the Brownian motion of the ring. **c**, Molecular structure of the components of the supramolecular pump of the present study: the DB24C8 macrocycle **1** and the non-symmetric molecular axle $E-2^+$.

$Z \rightarrow E$ isomerization. The system is, thus, able to repeat the cycle autonomously under constant illumination, reaching and maintaining a stationary non-equilibrium state dissipating light energy.

The operation of the molecular pump can be described with the reaction network depicted in Fig. 2. Two thermal processes (self-assembly, numbered 1 and 3, following mass-action kinetics) and two photochemical processes (E/Z isomerization, numbered 2 and 4, following nonlinearized photokinetics) are involved in the cycle. Only the two thermal processes must satisfy microscopic reversibility, as photochemical processes result from reaction mechanisms involving excited states of the molecules^{34–37}. Since the photoisomerization quantum yields and molar absorption coefficients of the free and complexed axle are almost identical (Supplementary Table 2), they exhibit essentially the same E/Z composition at the photostationary state (PSS), that is $K_{hv}^u = K_{hv}^c$. Conversely, the equilibrium constants for the self-assembly processes are different, specifically $K_a^E > K_a^Z$ (ref. 30,31). As a consequence, under irradiation, detailed balance in this reaction network cannot be fulfilled and the cycle is travelled clockwise with a non-zero net rate^{31,38}. The ability of the system to cycle preferentially clockwise with respect to anticlockwise—that is, its kinetic asymmetry—is quantified by the ratcheting constant (K_r), defined as the ratio between the clockwise and anticlockwise rates^{39,40}. For our system under irradiation, K_r is well approximated by the ratio $K_a^E/K_a^Z = 2$ meaning that, on average, the system loops anticlockwise once every two cycles in the clockwise direction (Supplementary Section 3.3).

Study of the light-dependent dissipative regime

A typical feature of dissipative systems is that a fraction of the energy input is employed to shift concentrations away from equilibrium values and is, thus, stored as free energy into the system^{3,19,41,42}. In fact, in a closed cycle of reactions when detailed balance is not fulfilled, the cycling net rate must be equal for all steps and non-zero ($v_{cy} \equiv v_1 = v_2 = -v_3 = -v_4 \neq 0$)³⁴. To satisfy this condition, the concentration of the species under operation must adjust to a level which is different from that of equilibrium. This is exactly what is observed for the cycle of the supramolecular pump (Fig. 2a)³¹.

Since the rates of the individual photochemical reactions (v_2 and v_4) are dictated by the photon flow (Fig. 2b), variation of the latter will modify the overall cycling rate. Therefore, the self-assembly steps (1 and 3) must adapt their net rate of reaction to the new cycling rate, adjusting the concentration of reagents and products to a different level. As a consequence, for reactions (1) and (3) a larger

Table 1 | Thermodynamic and kinetic parameters^a in air-equilibrated CD_3CN at 298 K for the rotaxane-type complexes formed between **1 and either $E-2^+$ or $Z-2^+$**

	K_a (mol l ⁻¹) ^b	k_{in} (l mol ⁻¹ s ⁻¹) ^c	k_{out} (s ⁻¹) ^d
1 + $E-2^+$	230 ± 30	16 ± 3 ^e	0.07 ^f
1 + $Z-2^+$	115 ± 35 ^e	(3.1 ± 0.8) × 10 ⁻² ^h	(2.7 ± 0.5) × 10 ⁻⁴ ^h

^aFrom ref. 30. ^bAssociation constant. ^cThreading rate constant. ^dDethreading rate constant. ^eDetermined by stopped-flow ultraviolet–visible absorption method. ^fCalculated as k_{in}/K_a . ^gCalculated as k_{in}/k_{out} . ^hDetermined from time-dependent ¹H NMR concentration profiles.

deviation from the equilibrium concentrations is expected at a higher photon flow. In turn, this means that, also in a light-driven system, the amount of stored energy should be correlated to the amount of energy provided by the external reservoir^{37,41}.

To probe this hypothesis it is crucial to measure the speciation in real time upon operation of the pump. In situ irradiation in the NMR probehead, as demonstrated recently, is perfectly suited for this task^{31,32}. We employed an optical fibre to deliver nearly monochromatic ultraviolet light ($\lambda_{irr} = 365 \pm 5$ nm) to the solution inside the NMR tube (Supplementary Fig. 2). The monochromatic setup is advantageous because (1) the behaviour of the system can be appropriately described using spectroscopic and photochemical parameters independently measured for the pump components (Supplementary Table 2) and (2) the photon flow can be accurately determined by chemical actinometry. Both these points are fundamental to perform quantitative numerical simulations to support the pump operating cycle.

Typically, an equimolar mixture of $E-2^+$ and **1** (8.9 mM) was equilibrated in the dark and then irradiated with full-power 365 nm light (photon flow, 4.4×10^{-8} Einstein s⁻¹) until a constant Z/E composition was reached (35 min). During this time, the E -configured axle and complex were photoconverted into the corresponding Z species faster than the disassembly of the Z complex. Thus, the Z species were produced in relative concentrations that were not consistent with K_a^Z ; specifically, the Z complex was kinetically trapped in a concentration higher than its equilibrium value. In the first instance, after reaching the PSS, the irradiation was turned off and spectra were acquired over time. In the dark, a decrease in the concentration of the Z complex, accompanied by a corresponding increase in the concentration of Z axle, was observed,

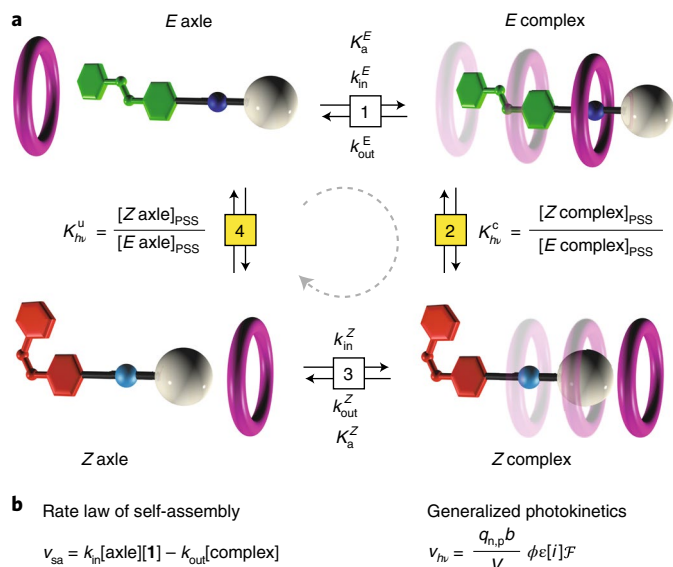


Fig. 2 | Operative reaction network of the supramolecular pump.

a, Thermal self-assembly equilibria (1 and 3) and photochemical isomerization reactions (2 and 4). The fictitious ‘equilibrium constants’ for the photochemical steps are defined as the ratio of the concentrations of the products and reactants at the PSS for the respective reactions. The dashed circular arrow indicates the net direction of travelling of the cycle. Conventionally all the parameters are positive for the reactions read from left to right and from top to bottom. **b**, Left: general rate law for the self-assembly reactions (1 and 3) following mass-action kinetics. Right: generalized photokinetics for photochemical reactions (2 and 4). $q_{n,p}$, photon flow; ϕ , quantum yield; ϵ , molar absorption coefficient; \mathcal{F} , photokinetic factor; i , photoactive species; b , optical path; V , volume of irradiated solution.

as **1** dethreaded to satisfy the equilibrium condition. After about 2 h (local) equilibrium was reached with a decrease in the *Z* complex concentration of about 10% (Fig. 3a). It should be noted that the *Z* species are kinetically trapped with respect to azobenzene isomerization because the thermodynamically stable isomer is *E*. For the sake of simplicity, from this point onward we shall refer to the kinetically trapped state of the system as an ‘equilibrium state’ referring to the local equilibrium of the *Z* complex self-assembly reaction. It has to be made clear, though, that the true equilibrium state of our system is the absence of *Z* species. It is also important to note that the thermal $Z \rightarrow E$ isomerization during this interval was negligible.

In the second instance, after reaching the PSS the irradiation was continued with the same photon flow (4.4×10^{-8} Einstein s^{-1}), or with a flow attenuated to 50%, 25% or 10% of the initial value (Supplementary Fig. 2). In all cases, new PSSs were reached within about 1 h in which the concentration of the *Z* complex was higher than the value reached in the dark (Fig. 3b–e). On the other hand, the concentration of the *Z* axle was lower than that in the dark (Supplementary Fig. 4). These results confirm that the system reached a dissipative non-equilibrium steady state in which the *Z* complex was kinetically accumulated at the expense of the *Z* axle³¹. The concentrations of the *Z* complex and *Z* axle progressively deviated more from their equilibrium values upon increasing the photon flow. Hence, we can qualitatively conclude that a larger energy input—more photons per unit time—drives the system farther away from equilibrium.

Numerical simulations, performed using experimentally determined (photo)chemical and photophysical parameters, provided time-dependent concentration profiles for all species. It is noteworthy

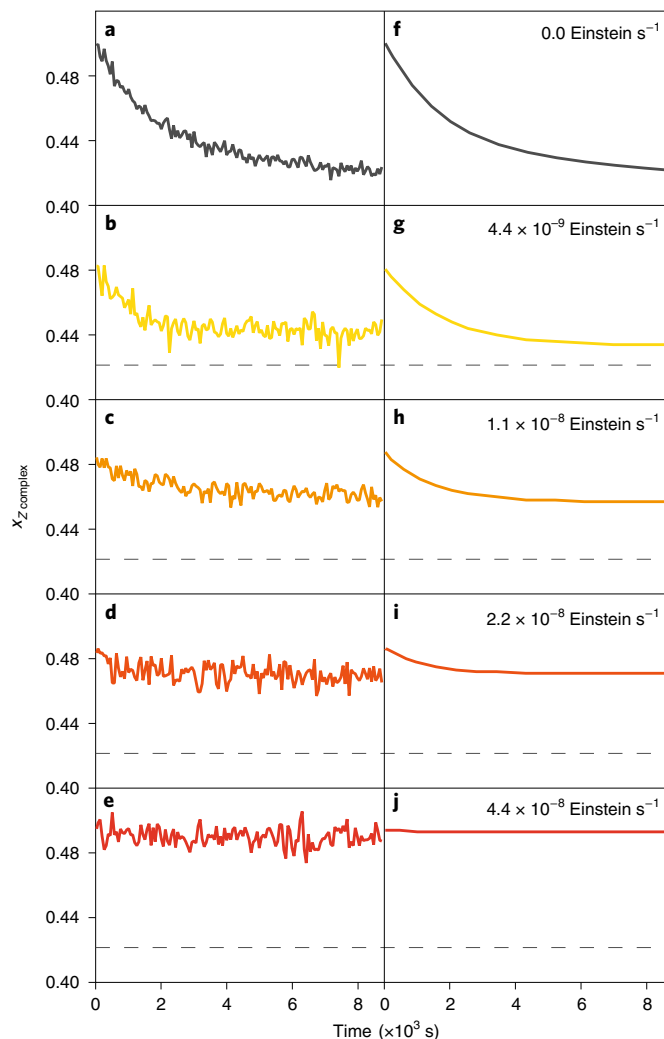


Fig. 3 | Time-dependent concentration profiles of complexed $Z\text{-}2^+$.

a–j, Experimental (left) and simulated (right) molar fraction profiles after photogeneration of the kinetically trapped *Z* complex in the dark (**a, f**) and upon constant irradiation at 365 nm with different light intensities (**b–e** and **g–j**). The photon flow employed to maintain the dissipative state is indicated on the right-hand side of each panel. The dashed lines mark the molar fraction of the *Z* complex at the local equilibrium in the dark corresponding to the plateau of the curve in **a**. Conditions: CD_3CN , 298 K, $[1] = [2^+]$ (initial concentrations for each experiment are reported in Supplementary Table 2).

that in the simulations the concentrations were assumed to be homogeneous in space. This assumption is not obvious a priori, as the intensity of the photon flux is dependent on the radius of the NMR tube. Nonetheless, the concentration profiles are in excellent agreement with the experimental data (Fig. 3f–j and Supplementary Fig. 4), confirming the fast diffusion of the molecules with respect to the cycling time of the pump. Moreover, the simulated data also confirm that the disassembly of the *Z* complex acts as the bottleneck for the reaction network. Thus, the *Z* complex is kinetically accumulated in different amounts depending on the intensity of the incident light.

Kinetic and energetic considerations

The possibility to directly measure or estimate by means of simulations the concentration of the species at the dissipative steady state provided us the unprecedented opportunity to gain quantitative

Table 2 | Kinetic and energetic parameters of the cycle of the 1/2⁺ ensemble for dissipative non-equilibrium operation under different photon flows^a

q_p (Einstein s ⁻¹) ^b	v_3 (μM s ⁻¹) ^c	Φ_{cy} [N_{hv} per cycle] ^d	$\Delta\mu_1$ (J mol ⁻¹) ^e	$\Delta\mu_3$ (J mol ⁻¹) ^e	$T\Delta_{cy,sa}$ (J mol ⁻¹) ^f	ΔG_{sa} (J l ⁻¹) ^g	η (%) ^h
0	-	-	0	0	0	0	-
4.4×10^{-9}	-0.25	0.11 [9]	-40	1,380	1,420	0.25	0.3
1.1×10^{-8}	-0.34	0.06 [16]	-64	1,776	1,840	0.48	0.2
2.2×10^{-8}	-0.38	0.04 [28]	-89	1,961	2,050	0.61	0.1
4.4×10^{-8}	-0.42	0.02 [53]	-149	1,902	2,051	0.63	0.08

^aConditions: CD₃CN, 298 K, [1] = [2⁺] (initial concentrations for each experiment are reported in Supplementary Table 2), λ_{irr} = 365 nm. All the properties refer to the whole system with homogeneous (well-mixed) concentrations. ^bPhoton flow at 365 nm, determined by chemical actinometry. ^cRate of reaction (3) at the stationary state calculated according to equation (1); see Fig. 2 for the sign convention. ^dQuantum yield of the cycle calculated using Supplementary equation (S2); numbers in brackets are the number of absorbed photons per cycle (1/ Φ_{cy}). ^eChemical potential of reaction according to Supplementary equation (S9), determined from simulated (reaction (1)) or experimental (reaction (3)) concentration values. ^fFree energy dissipated by self-assembly steps during a cycle of operation at the steady state. ^gFree energy density stored in the self-assembly steps at the non-equilibrium steady state. ^hEnergy transduction efficiency, calculated as the ratio between $T\Delta_{cy,sa}$ and the total amount of free energy absorbed during a cycle of operation at the steady state.

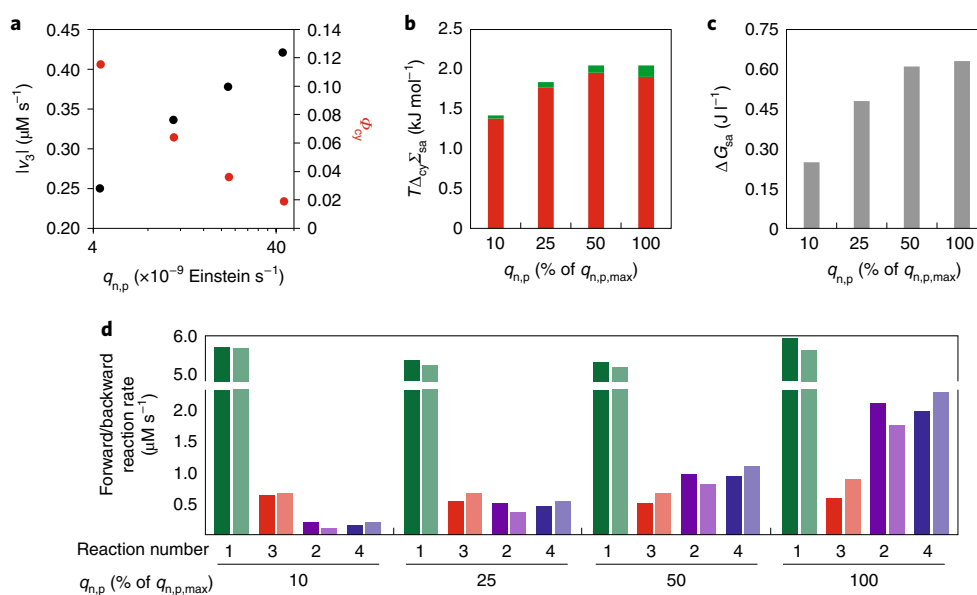


Fig. 4 | Dependence of the kinetic and thermodynamic parameters on the photon flow. a, Net cycling rate (black dots) and quantum yield of cycling (red dots) at 298 K. **b**, Free energy dissipated by the self-assembly steps during one cycle of operation at the steady state. The fractions of free energy dissipated by reactions (1) (green bars) and (3) (red bars) are indicated. **c**, Free energy density stored in the self-assembly steps. **d**, Simulated forward (dark bars) and backward (light bars) rates for reactions (1) (green), (2) (purple), (3) (red) and (4) (blue) at different photon flows. In the bar graphs the photon flow is expressed as a percentage of the maximum flow ($q_{n,p,max} = 4.4 \times 10^{-8}$ Einstein s⁻¹ = 100%).

information on the reaction network under autonomous cycling away from equilibrium.

The first quantity that can be extracted from our data is the rate of cycling under irradiation. As discussed above, in a closed cycle of reactions, at the steady state, all the reactions must proceed at the same net rate. Therefore, the rate of cycling at the steady state can be calculated on any reaction of the network (Fig. 2). The rate of reaction (3) (v_3) was calculated using a mixed-order kinetic equation (equation (1)) from the experimental concentrations at the steady state and the rate constants (Table 2).

$$v_3 = k_{in}^Z [Z \text{ axle}] [1] - k_{out}^Z [Z \text{ complex}] \equiv -v_{cy} \quad (1)$$

The rate v_3 is negative for all photon flows, indicating that the Z complex is undergoing dethreading and that the cycle shown in Fig. 2 is being travelled clockwise (dashed circular arrow), consistent with previous considerations on the kinetic asymmetry. As anticipated, the cycling rate (v_{cy}) becomes larger upon increasing

the photon flow (Table 2 and Fig. 4a). Finally, the K_r of the system, calculated from the simulated clockwise and anticlockwise reaction rates, is close to the estimated value of 2 (Supplementary Table 3).

On the contrary, the quantum yield of the cycle (Φ_{cy} , Table 2), that is the number of molecules that have completed a full cycle upon the absorption of one photon, decreases from 11% at the lowest photon flow to 2% at the highest photon flow (Fig. 4a). This means that, on average at the lowest light intensity, a given ring/axle pair needs to absorb nine photons to complete a cycle, while this number becomes more than five times larger at the highest intensity (Table 2).

The chemical potential difference ($\Delta\mu$) of each process at the steady state can be directly calculated from the experimental or simulated concentrations (Supplementary Section 5.1) and can be interpreted as a measure of how far the reaction is from equilibrium⁴³. The steady state $\Delta\mu$ values of the self-assembly of the E (reaction (1), $\Delta\mu_1$) and the Z (reaction (3), $\Delta\mu_3$) complexes both increase in magnitude upon increasing the photon flow (Table 2).

This means that both reactions are progressively displaced farther from their equilibrium state; however, they exhibit opposite trends. Specifically, $\Delta\mu_3$ becomes more positive upon increasing the photon flow, showing that larger light intensities drive reaction (3) more towards the product. Such an observation is consistent with the kinetic accumulation of the *Z* complex and with the increasingly negative values of v_3 . Conversely, $\Delta\mu_1$ becomes more negative at higher light intensities, indicating that reaction (1) is progressively shifted back towards the reagents. Therefore, at the steady state the *Z* complex and *E* axle are accumulated, whereas the *E* complex and *Z* axle are depleted due to purely kinetic effects⁴⁴.

From a thermodynamic point of view, our system can be thought of as a thermal machine powered by the gradient between the temperature *T* of the solvent (the cold reservoir) and the temperature T_{hv} of the radiation (the hot reservoir), which can be correlated to the experimental photon flow using Planck's law (Supplementary Section 5.2). As a consequence, its efficiency is limited by Carnot's theorem^{35,36}. The gradient of temperature powering the machine also relates to the chemical potential gradient generated by the radiation (Supplementary Table 7). Larger photon flows correspond to a higher temperature of the hot reservoir, thus producing a larger free energy gradient to drive the machine farther away from equilibrium^{35–37,45}. In the present case, the gradient is about one order of magnitude higher than that delivered by the hydrolysis of adenosine triphosphate (ATP) in physiological conditions and comparable to that generated by, for example, the decarboxylation of the fluorenylmethoxycarbonyl group employed to power other synthetic molecular motors¹².

Each time a photon is absorbed by the system, its free energy can be either dissipated or converted into chemical free energy by changing the concentration distribution^{37,41,43,46}. This can be expressed with the following instantaneous free energy balance:

$$\dot{W}_{hv} = d_t G + T \dot{\Sigma} \quad (2)$$

where \dot{W}_{hv} , $d_t G$ and $T \dot{\Sigma}$ are the free energy densities absorbed, stored and dissipated, respectively, by the whole system per unit time. By integrating equation (2) it is possible to compute the amount of free energy dissipated and stored in any time interval (Supplementary Section 5.3).

When the system reaches the steady state, the power absorbed from the radiation is no longer stored in the system ($d_t G = 0$), but is entirely dissipated to sustain the non-equilibrium steady state⁴⁶. Thus, over one cycle, $\Delta_{cy} W_{hv} = T \Delta_{cy} \Sigma$. This free energy dissipation can be split into two contributions: that of the self-assembly and that of the isomerization processes, namely $T \Delta_{cy} \Sigma = T \Delta_{cy} \Sigma_{sa} + T \Delta_{cy} \Sigma_{iso}$ (ref. 47). The self-assembly dissipation ($T \Delta_{cy} \Sigma_{sa}$) is the portion of the absorbed free energy responsible for keeping the self-assembly steps out of equilibrium and for sustaining the unidirectional motion of the rings with respect to the axles. This value can also be calculated from $\Delta\mu_1$ and $\Delta\mu_3$ (Supplementary Section 5.3.1). Dissipation increases with the photon flow (Table 2), consistently with the self-assembly steps being progressively shifted away from their equilibrium state. The fraction of free energy input dissipated by the self-assembly reactions is the efficiency of free energy transduction from the photo-isomerization processes towards the self-assembly ones ($\eta = T \Delta_{cy} \Sigma_{sa} / \Delta_{cy} W_{hv}$)⁴⁷. The efficiency is relatively small at all operating regimes, meaning that most of the free energy harvested from the radiation is dissipated by the photoisomerization steps.

The free energy density stored in the self-assembly steps (ΔG_{sa}) coincides with the difference between the Gibbs free energy densities in the steady and local equilibrium states ($\Delta G_{sa} = G_{light} - G_{dark}$, see Supplementary Section 5.3.2). Thus, it can be computed from the experimentally determined steady state concentrations. It is noteworthy that the free energy storage increases with the photon flow until it reaches a plateau around 0.6 J l^{-1} .

Interestingly, two distinct trends can be identified in the dependence of the determined quantities on the light intensity. Upon increasing the photon flow, ΔG_{sa} , $T \Delta_{cy} \Sigma_{sa}$ and v_{cy} increase, reaching a plateau between 1 and $2 \times 10^{-8} \text{ Einstein s}^{-1}$, while η and Φ_{cy} decrease (Fig. 4a–c). This observation is in qualitative agreement with the results of stochastic simulations performed on a closely related molecular pump, which predicted that the network can process photons efficiently only as long as the photoreactions are the rate-limiting steps⁴⁸. To support this claim, we compared the forward and backward rates of each reaction in the network (Fig. 4d). At the steady state, the forward and backward reaction rates (Fig. 4d) can assume different values for the four processes, provided that their relative difference is the same for all reactions. In fact, for photon flows up to $1.1 \times 10^{-8} \text{ Einstein s}^{-1}$ (that is, 25% of $q_{n,p,max}$), the photochemical reactions are slower than the thermal ones and determine the overall cycling rate in either direction. Conversely, at higher flows reaction (3) becomes rate limiting (red bars in Fig. 4d), while photoreactions proceed faster dissipating a larger fraction of light energy, thus reducing the energy transduction efficiency.

The observed decrease of the quantum yield of the cycle (Φ_{cy}) (Table 2) is also coherent with this picture. Notably, already at the lowest flow employed, completing one cycle requires nine photons (Table 2), a number significantly larger than the theoretical minimum value of two. This means that, in the investigated light intensity range, a given axle (complex) undergoes several $E \rightleftharpoons Z$ photoisomerization events before threading (dethreading). Increasing the rate of the photoreactions increases the number of unfruitful $E \rightleftharpoons Z$ isomerization events. As indicated by earlier simulations⁴⁸, only at low photon rates can the pump operate close to the maximum quantum efficiency.

Conclusions

A combination of experimental data and numerical simulations was used to characterize from a kinetic and thermodynamic point of view the dissipative photostationary states of a supramolecular pump for four intensities of incident light. Our measurements quantitatively probed the relationship between the deviation from thermodynamic equilibrium and the photon flow and provide an unprecedented insight in the non-equilibrium behaviour of (photo)chemical reaction networks. Although in the current design the stored energy cannot be converted into work, our analysis allows the maximum amount of free energy which can potentially be converted into work to be quantified. We believe that this approach is applicable to investigate any kind of light-fuelled non-equilibrium chemical ensemble, providing a testing ground for recently developed theoretical models²⁷. We envision that our results will stimulate research on new, more sophisticated light-driven artificial molecular machines and materials capable of operating away from thermodynamic equilibrium.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41565-022-01151-y>.

Received: 23 December 2021; Accepted: 12 May 2022;

Published online: 27 June 2022

References

- Schliwa, M. & Woehlke, G. Molecular motors. *Nature* **422**, 759–765 (2003).
- Goodsell, D. S. *The Machinery of Life* (Copernicus, 2009).
- Solis Muñana, P. et al. Substrate-induced self-assembly of cooperative catalysts. *Angew. Chem. Int. Ed.* **57**, 16469–16474 (2018).

4. Yang, S. et al. Chemical fueling enables molecular complexification of self-replicators. *Angew. Chem. Int. Ed.* **60**, 11344–11349 (2021).
5. Boekhoven, J. et al. Dissipative self-assembly of a molecular gelator by using a chemical fuel. *Angew. Chem. Int. Ed.* **49**, 4825–4828 (2010).
6. Wilson, M. R. et al. An autonomous chemically fueled small-molecule motor. *Nature* **534**, 235–240 (2016).
7. Borsley, S., Kreidt, E., Leigh, D. A. & Robert, B. M. W. Autonomous fuelled directional rotation about a covalent single bond. *Nature* **604**, 80–85 (2022).
8. Amano, S., Fielden, S. D. P. & Leigh, D. A. A catalysis-driven artificial molecular pump. *Nature* **594**, 529–534 (2021).
9. Cheng, C. et al. An artificial molecular pump. *Nat. Nanotechnol.* **10**, 547–553 (2015).
10. Qiu, Y. et al. A precise polyrotaxane synthesizer. *Science* **368**, 1247–1253 (2020).
11. Astumian, R. D. Optical vs. chemical driving for molecular machines. *Faraday Discuss.* **195**, 583–597 (2016).
12. Kathan, M. & Hecht, S. Photoswitchable molecules as key ingredients to drive systems away from the global thermodynamic minimum. *Chem. Soc. Rev.* **46**, 5536–5550 (2017).
13. Baroncini, M., Silvi, S. & Credi, A. Photo- and redox-driven artificial molecular motors. *Chem. Rev.* **120**, 200–268 (2020).
14. Feringa, B. L. The art of building small: from molecular switches to motors (Nobel Lecture). *Angew. Chem. Int. Ed.* **56**, 11060–11078 (2017).
15. Ragazzon, G., Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Light-powered autonomous and directional molecular motion of a dissipative self-assembling system. *Nat. Nanotechnol.* **10**, 70–75 (2015).
16. Ceroni, P., Credi, A. & Venturi, M. Light to investigate (read) and operate (write) molecular devices and machines. *Chem. Soc. Rev.* **43**, 4068–4083 (2014).
17. Foy, J. T. et al. Dual-light control of nanomachines that integrate motor and modulator subunits. *Nat. Nanotechnol.* **12**, 540–545 (2017).
18. Ikegami, T., Kageyama, Y., Obara, K. & Takeda, S. Dissipative and autonomous square-wave self-oscillation of a macroscopic hybrid self-assembly under continuous light irradiation. *Angew. Chem. Int. Ed.* **55**, 8239–8243 (2016).
19. Herder, M. & Lehn, J.-M. The photodynamic covalent bond: sensitized alkoxyamines as a tool to shift reaction networks out-of-equilibrium using light energy. *J. Am. Chem. Soc.* **140**, 7647–7657 (2018).
20. Greb, L., Eichh fer, A. & Lehn, J.-M. Synthetic molecular motors: thermal N inversion and directional photoinduced C=N bond rotation of camphorquinone imines. *Angew. Chem. Int. Ed.* **54**, 14345–14348 (2015).
21. Cheng, C., McGonigal, P. R., Stoddart, J. F. & Astumian, R. D. Design and synthesis of nonequilibrium systems. *ACS Nano* **9**, 8672–8688 (2015).
22. Merindol, R. & Walther, A. Materials learning from life: concepts for active, adaptive and autonomous molecular systems. *Chem. Soc. Rev.* **46**, 5588–5619 (2017).
23. Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotechnol.* **13**, 882–889 (2018).
24. van Rossum, S. A. P., Tena-Solsona, M., van Esch, J. H., Eelkema, R. & Boekhoven, J. Dissipative out-of-equilibrium assembly of man-made supramolecular materials. *Chem. Soc. Rev.* **46**, 5519–5535 (2017).
25. Das, K., Gabrielli, L. & Prins, L. J. Chemically fueled self-assembly in biology and chemistry. *Angew. Chem. Int. Ed.* **60**, 20120–20143 (2021).
26. Mattia, E. & Otto, S. Supramolecular systems chemistry. *Nat. Nanotechnol.* **10**, 111–119 (2015).
27. Sorrenti, A., Leira-Iglesias, J., Sato, A. & Hermans, T. M. Non-equilibrium steady states in supramolecular polymerization. *Nat. Commun.* **8**, 15899 (2017).
28. Geertsema, E. M., van der Molen, S. J., Martens, M. & Feringa, B. L. Optimizing rotary processes in synthetic molecular motors. *Proc. Natl Acad. Sci. USA* **106**, 16919–16924 (2009).
29. Wilcken, R., Gerwien, A., Huber, L. A., Dube, H. & Riedle, E. Quantitative in-situ NMR illumination for excitation and kinetic analysis of molecular motor intermediates. *ChemPhotoChem* **6**, e202100232 (2022).
30. Corra, S. et al. Artificial supramolecular pumps powered by light. *Chem. Eur. J.* **27**, 11076–11083 (2021).
31. Canton, M. et al. Second-generation light-fueled supramolecular pump. *J. Am. Chem. Soc.* **143**, 10890–10894 (2021).
32. Nitschke, P., Lokesh, N. & Gschwind, R. M. Combination of illumination and high resolution NMR spectroscopy: key features and practical aspects, photochemical applications, and new concepts. *Prog. Nucl. Magn. Reson. Spectrosc.* **114**, 86–134 (2019).
33. Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Reversible photoswitching of rotaxane character and interplay of thermodynamic stability and kinetic lability in a self-assembling ring–axle molecular system. *Chem. Eur. J.* **16**, 11580–11587 (2010).
34. Blackmond, D. G. “If pigs could fly” chemistry: a tutorial on the principle of microscopic reversibility. *Angew. Chem. Int. Ed.* **48**, 2648–2654 (2009).
35. Ross, R. T. Some thermodynamics of photochemical systems. *J. Chem. Phys.* **45**, 1–7 (1966).
36. Porter, G. Transfer and storage of chemical and radiation potential. *J. Chem. Soc., Faraday Trans. 2* **79**, 473–482 (1983).
37. Penocchio, E., Rao, R. & Esposito, M. Nonequilibrium thermodynamics of light-induced reactions. *J. Chem. Phys.* **155**, 114101 (2021).
38. Onsager, L. Reciprocal relations in irreversible processes. I. *Phys. Rev.* **37**, 405–426 (1931).
39. Astumian, R. D. & Bier, M. Mechanochemical coupling of the motion of molecular motors to ATP hydrolysis. *Biophys. J.* **70**, 637–653 (1996).
40. Astumian, R. D. Stochastic conformational pumping: a mechanism for free-energy transduction by molecules. *Annu. Rev. Biophys.* **40**, 289–313 (2011).
41. Kondepudi, D. & Prigogine, I. *Modern Thermodynamics: From Heat Engines to Dissipative Structures* (Wiley, 1998).
42. Rao, R. & Esposito, M. Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. *Phys. Rev. X* **6**, 041064 (2016).
43. Demirel, Y. & Gerbaud, V. in *Nonequilibrium Thermodynamics* 4th edn, Ch. 8 (Elsevier, 2019).
44. Lehn, J.-M. Perspectives in chemistry—aspects of adaptive chemistry and materials. *Angew. Chem. Int. Ed.* **54**, 3276–3289 (2015).
45. Ries, H. & McEvoy, A. J. Chemical potential and temperature of light. *J. Photochem. Photobiol. A* **59**, 11–18 (1991).
46. Penocchio, E., Rao, R. & Esposito, M. Thermodynamic efficiency of dissipative chemistry. *Nat. Commun.* **10**, 3865 (2019).
47. Amano, S. et al. Insights from an information thermodynamics analysis of a synthetic molecular motor. *Nat. Chem.* <https://doi.org/10.1038/s41557-022-00899-z> (2022).
48. Sabatino, A., Penocchio, E., Ragazzon, G., Credi, A. & Frezzato, D. Individual-molecule perspective analysis of chemical reaction networks: the case of a light-driven supramolecular pump. *Angew. Chem. Int. Ed.* **58**, 14341–14348 (2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Methods

Dissipative operation of the supramolecular pump. Typically, an equimolar solution of **1** and *E*-**2**⁺ in air-equilibrated CD₃CN (8.9 × 10⁻³ M, 0.6 ml) was allowed to reach thermodynamic equilibrium in the dark at 298 K. Subsequently, a non-equilibrium mixture of the corresponding free and complexed *Z*-**2**⁺ was obtained by photoisomerization of the azobenzene unit. Photoisomerization was performed in situ using a light-emitting diode Illuminator (1.5 W; λ_{max} = 369 nm; full-width at half-maximum, 15.56 nm) equipped with a bandpass filter centred at 365 ± 5 nm. A quartz optical fibre (core 1,000 μm, 5 m) was used to channel light to the solution. The terminal end of the optical fibre (the quartz core) was exposed and submerged into the solution within the NMR tube. Upon reaching a stable *Z/E* composition (PSS), irradiation was interrupted (dark, relaxation to local equilibrium) or regulated (light on, dissipative operation) using the appropriate neutral density filter and the time-dependent concentration changes of the species were followed using ¹H NMR.

Dynamic simulations. The linear time-variant (LTV) system of differential equations describing the pumping cycle was implemented considering three subsequent time-invariant steps (LTI): (a) equilibration in the dark, (b) photoisomerization and (c) operation regime. For each LTI step the reaction rates for all processes and species concentrations were computed separately at any time point. The final concentrations were taken as initial conditions for the subsequent LTI step.

Data availability

All data needed to evaluate the conclusions are present in the main text and/or the Supplementary Information. Additional data related to this paper may be requested

from the corresponding author on reasonable request. Source data are provided with this paper.

Acknowledgements

Financial support from the EU (H2020 ERC AdG 692981 and ERC-2015-CoG n. 681456) and the Ministero dell'Università e della Ricerca (PRIN 20173L7W8K and 201732PY3X, FARE R16S9XXKX3) is gratefully acknowledged.

Author contributions

S.C. and J.G. synthesized the compounds. M.T.B., S.C., S.S. and M.B. designed and performed the kinetic experiments and analysed the data. S.C. carried out numerical simulations. E.P. and M.E. performed the thermodynamic analysis. A.C. and S.C. wrote the manuscript with input from all authors. All authors discussed the results and commented on the manuscript. A.C. conceived the project and directed the research.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41565-022-01151-y>.

Correspondence and requests for materials should be addressed to Alberto Credi.

Peer review information *Nature Nanotechnology* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Supplementary information

Kinetic and energetic insights into the dissipative non-equilibrium operation of an autonomous light-powered supramolecular pump

In the format provided by the authors and unedited

Kinetic and energetic insights into the dissipative non-equilibrium operation of an autonomous light-powered supramolecular pump

Stefano Corra,^{‡,1,2} Marina Tranfić Bakić,^{‡,1,2} Jessica Groppi,¹ Massimo Baroncini,^{1,4} Serena Silvi,^{1,3} Emanuele Penocchio⁵, Massimiliano Esposito⁵, Alberto Credi^{1,2,*}

¹ CLAN-Center for Light Activated Nanostructures, Istituto ISOF-CNR, via Gobetti 101, 40129 Bologna, Italy

² Dipartimento di Chimica Industriale “Toso Montanari”, Università di Bologna, viale del Risorgimento 4, 40136 Bologna, Italy

³ Dipartimento di Chimica “G. Ciamician”, Università di Bologna, via Selmi 2, 4126 Bologna, Italy

⁴ Dipartimento di Scienze e Tecnologie Agro-alimentari, Università di Bologna, viale Fanin 44, 40127 Bologna, Italy

⁵ Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg City, G. D. Luxembourg

[‡] these authors contributed equally to the work

SUPPLEMENTARY INFORMATION

1. Materials and methods	2
2. Synthesis and characterization of the supramolecular pump	3
3. Operation of the Supramolecular Pump	6
4. Fitting and Numerical Simulations	10
5. Thermodynamic Analysis	14
7. References	23

1. Materials and methods

General materials. All reagents and chemicals were purchased from Sigma-Aldrich or VWR international and used as received unless otherwise stated. Flash column chromatography was performed using Sigma Aldrich Silica 40 (230-400 mesh size or 40-63 μm) as the stationary phase. Thin layer chromatography was performed on TLC Silica gel 60 F254 coated aluminum plates from Merck.

NMR Spectroscopy. NMR spectra were recorded on an Agilent DD2 spectrometer operating at 500 MHz. Chemical shifts are quoted in parts per million (ppm) relative to tetramethylsilane using the residual solvent peak as a reference standard and all coupling constants (J) are expressed in Hertz (Hz).

NMR Photochemistry Photochemical reactions were performed in air-equilibrated CD_3CN solutions at 298 K inside NMR tubes in the spectrometer probehead, using a Prizmatix UHP-T-365-SR LED Illuminator (1.5 W, $\lambda_{\text{max}} = 369$ nm, FWHM, 15.56 nm) equipped with an FCA-SMA adaptor for optical fiber. The desired irradiation wavelength of 365 nm was selected using the appropriate hard coated OD 4.0 bandpass filter. Neutral density AR coated filters were used to regulate intensity of incident light when required. Quartz optical fiber (core 1000 μm , 5 m) equipped with a SMA connector on one end was purchased from Thorlabs. The other end of the optical fiber was scraped to remove the protective coatings, exposing the quartz core, and submerged into the solution within the NMR tube to be irradiated. The emission spectra of the light source (optical fiber end with protective coatings removed) were measured with an AVANTES Star Line AvaSpec-ULS2048CL-EVO-RS spectrometer. The emission intensity was selected using the appropriate neutral density filter with known optical density, purchased from Edmund Optics.

Least square fitting and numerical simulations. Fitting of the experimental data to the appropriate kinetic model was performed using Berkeley Madonna 10. Numerical simulations were performed with MATLAB R2018b implementing the kinetic model including all the kinetic and photokinetic equations as a system of differential equations that was solved numerically with ode15s solver.

2. Synthesis and characterization of the supramolecular pump

2.1 Synthesis of the axle

The axle was synthesized according to a previously published procedure.¹ Analytical data were in line with those previously reported.

2.2. ¹H NMR characterization of axles and complexes

¹H NMR spectra of *E* and *Z* axles and complexes were consistent with those previously reported in ref. 1 and summarized in Fig. S1.

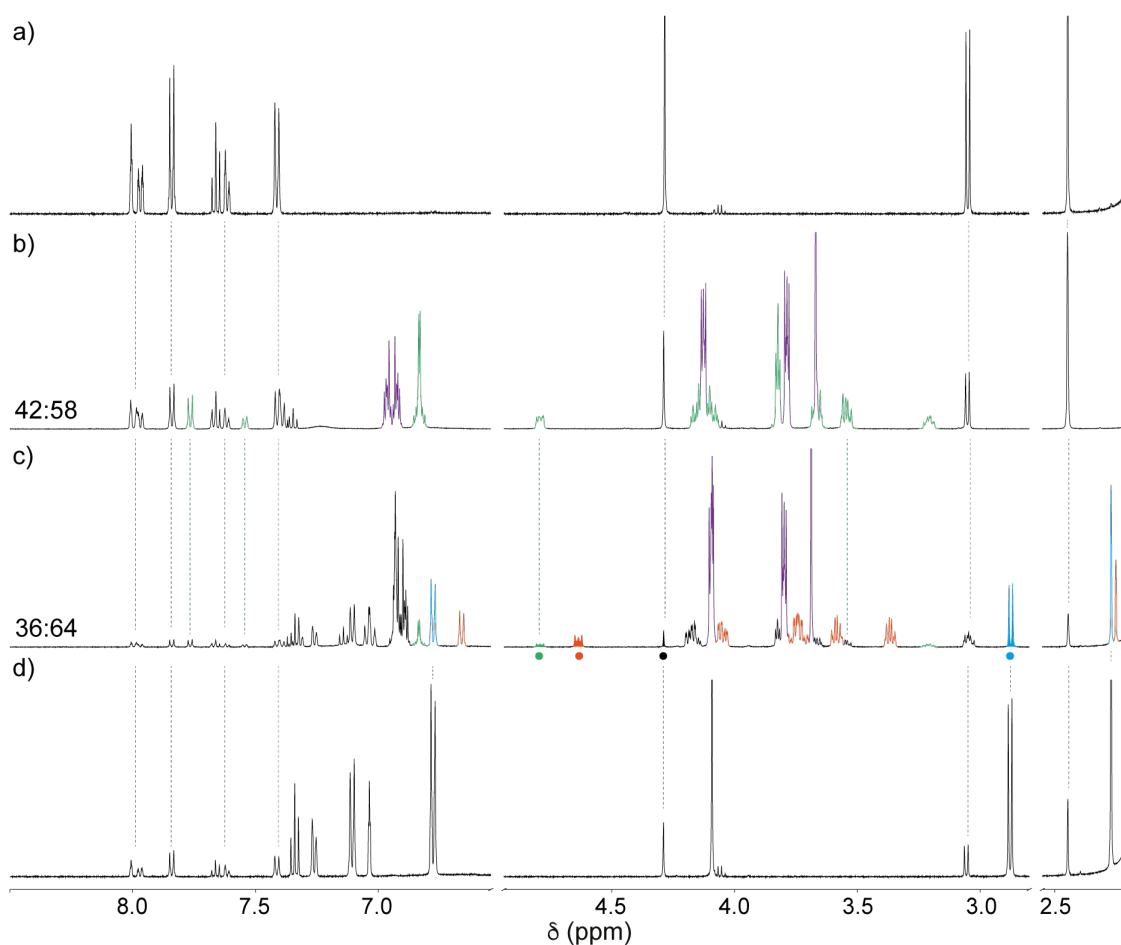


Figure S1. Typical ¹H NMR spectra (500 MHz, CD₃CN, 298 K) of a) *E*-2⁺. b) A 1:1 equilibrated mixture of *E*-2⁺ and 1, the signals univocally associated with the complex are highlighted in green, the signals of free 1 are highlighted in purple. c) The same sample after exhaustive irradiation ($\lambda_{irr} = 365$ nm, 30 min) and equilibration for 3 hours in dark. Signals of free *Z*-1⁺ are highlighted in cyan, signals univocally attributed to the *Z*-complex are highlighted in orange. d) same sample as (a) after exhaustive irradiation ($\lambda_{irr} = 365$ nm, 30 min). The dots underneath the filled signals in (c) mark the signals univocally assigned to the species followed for the kinetic measurements: green: 1·*E*-2⁺; orange: 1·*Z*-2⁺; cyan: *Z*-2⁺; black: *E*-2⁺. Ratios in (b) and (c) indicate the complex associated fraction (complex:axle) obtained by integration of univocally identified, non-overlapping, signals for complex and free axle.

2.3 NMR Photochemistry

About 6 cm of the coating were removed from the terminal end of a quartz optical fiber (~5 m). The exposed quartz core was sanded in order to diffuse light into solution. In a typical NMR experiment, the exposed core of the fiber was immersed in 0.6 mL of an equimolar solution ($\sim 9 \times 10^{-3}$ M) of *E-2*⁺ and **1** inside an NMR tube. The other extremity of the fiber was connected to a LED apparatus equipped with two filter holders. The wavelength of emission was 365 ± 5 nm, selected using a hard coated bandpass interference filter. The number of incident photons was selected using an additional neutral density (ND) filter with known optical density.

2.3.1 Optical Fiber Emission Spectra

The polychromatic emission spectra of the light source at different light intensities obtained with the appropriate neutral density filter consist of broad bands with a maximum at 369 nm (Fig. S2a). The 365 ± 5 nm interference filter significantly narrowed the emission spectrum (Fig. S2b). In both setups, the integral of the emission bands, which is proportional to the output power, changes linearly with the transmittance of the ND filter (insets in Fig. S2). The emission spectra reported in Fig. S2b correspond to the “monochromatic” irradiation wavelength used in all the photochemical experiments.

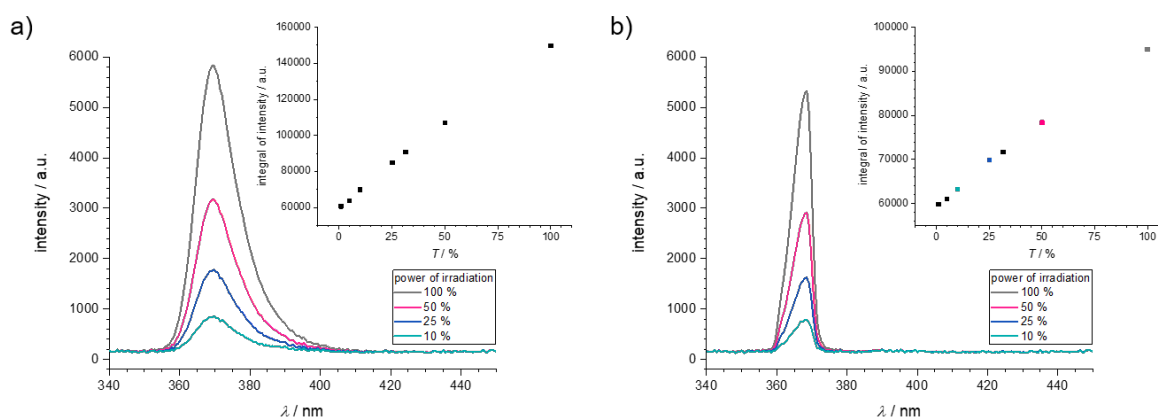


Figure S2. Emission spectra of the light source a) polychromatic setup b) monochromatic setup. In the inset the plot of integrated intensity on the percentage of transmitted light achieved with the appropriate ND filter. Light intensity values corresponding to those used for pump operation are reported in the same color as the corresponding spectrum in the inset.

2.3.2 NMR Chemical Actinometry

In order to estimate the photon flow of the optical fiber chemical actinometry was performed using the same experimental setup as for the molecular pump irradiation experiments using azobenzene as the actinometer. A solution of azobenzene in CD₃OD (1.19×10^{-2} M, 0.6 mL) was irradiated at 365 nm inside the NMR tube and *E*→*Z* isomerization was followed by ¹H NMR (Fig. S3). Photostationary state (PSS) was reached in about 50 min. The photon flow ($q_{n,p}$) and the ϵ_Z were

determined by least-square fitting of the concentration profiles of *E*- and *Z*-azobenzene to the photokinetic equation set (see below eq. S8 and S9), considering the thermal *Z*→*E* isomerization. Molar absorption coefficient for *E*-azobenzene² and quantum yields³ in methanol were taken from literature data, while the obtained molar absorption coefficient for *Z*-azobenzene at 365 nm is in line with the reported data³. A photon flow ($q_{n,p,max}$) of 4.4×10^{-8} Einstein s⁻¹ was obtained, corresponding to an output power of about 14 mW.

Table S1. Photophysical data of azobenzene in methanol used for actinometry.

ϵ (M ⁻¹ cm ⁻¹)		ϕ^{E-Z} [a]	ϕ^{Z-E} [a]	k_{Δ} (s ⁻¹)
<i>E</i> -azobenzene	<i>Z</i> -azobenzene			
375 ^[b]	92 ^[c]	0.12	0.34	1.0×10^{-6}

^[a]Ref. 3. ^[b]Ref. 2. ^[c]Fitted parameter.

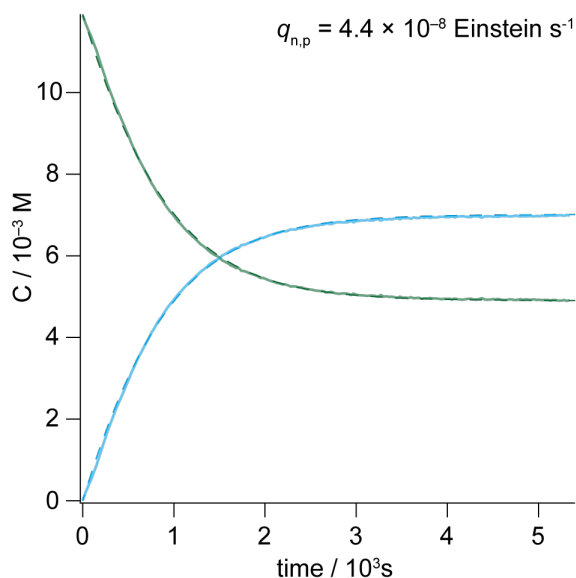


Figure S3. ¹H NMR kinetic of photoconversion of azobenzene (1.19×10^{-2} M, CD₃OD, 298 K). *E*-azobenzene: green line, *Z*-azobenzene: blue line. Solid lines represent the experimental concentration profile, dashed lines are least square fitting to the photokinetic equations set.

Since the radiant power increases linearly with the transmittance (Fig. S2), the number of incident photons for the numerical simulations (see below) was calculated from the 100% intensity diminished by the optical density of the filter used according to the following equation.

$$q_{n,p} = q_{n,p,max} \times 10^{-O.D.} \quad (S1)$$

3. Operation of the Supramolecular Pump

Typically, an equimolar mixture of **1** and $E\text{-}2^+$ ($\sim 9 \times 10^{-3}$ M) was allowed to reach thermodynamic equilibrium in the dark, then a non-equilibrium mixture of Z axle and Z complex was obtained by photoisomerization of the azobenzene unit. Photoisomerization was performed *in situ* using the setup described in section 2.3.

Upon reaching a stable Z/E composition (PSS) irradiation was interrupted (Fig. S4, black trace) or reduced (Fig. S4, colored traces) and the time-dependent changes in concentration of the photoactive species were followed by ^1H NMR.

The initial concentrations of **1** and $E\text{-}2^+$ used in each experiment are reported in Table S2.

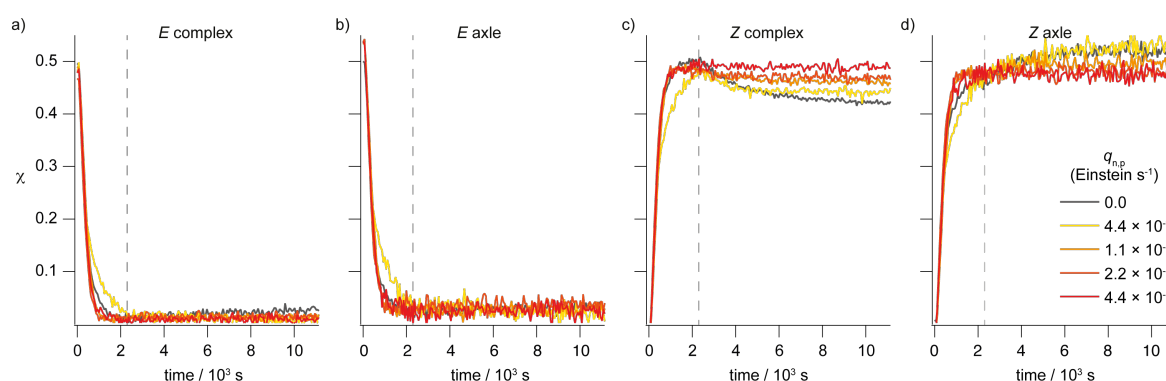
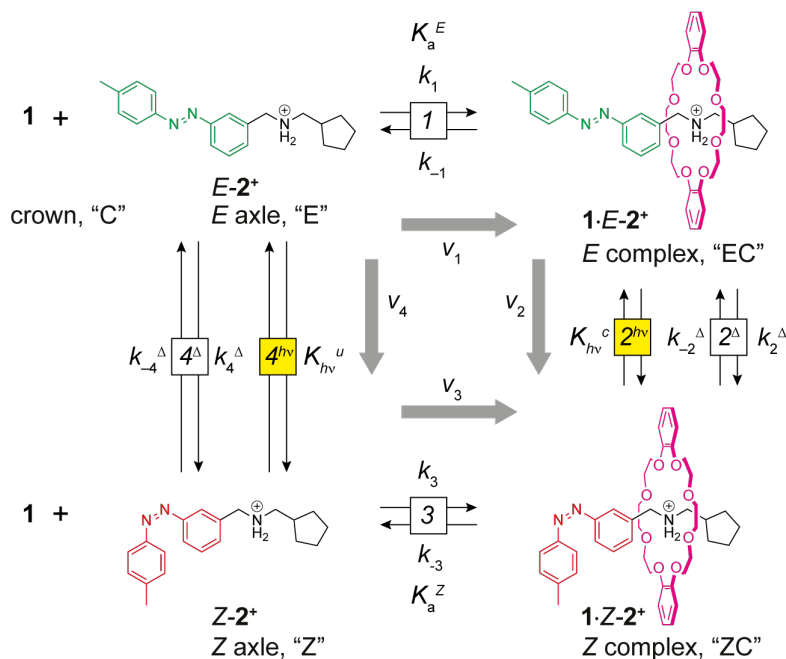


Figure S4. Time-dependent molar fraction profiles of E complex (a), E axle (b), Z complex (c), and Z axle (d) at different light intensities generated by light-induced isomerization of an equilibrated mixture of $E\text{-}2^+$ and **1**. The gray dashed line indicates the time at which PSS is reached. Data obtained from ^1H NMR data (500 MHz, CD_3CN , 298 K, initial concentrations for each experiment are reported in Table S2).

Table S2. Experimental initial concentrations of **1** and $E\text{-}2^+$.

$q_{n,p}$ (Einstein s^{-1})	0	4.4×10^{-9}	1.1×10^{-8}	2.2×10^{-8}	4.4×10^{-8}
(% of $q_{n,p,max}$)	0%	10%	25%	50%	100%
$[2^+]_0$ (M)	8.91×10^{-3}	7.43×10^{-3}	7.43×10^{-3}	7.43×10^{-3}	8.17×10^{-3}
$[1]_0$ (M)	8.98×10^{-3}	7.48×10^{-3}	7.48×10^{-3}	7.48×10^{-3}	8.23×10^{-3}

The complete closed reaction network describing the pumping cycle for compounds **1** and **2⁺** is reported in Scheme S1. This set of six reactions (four thermal and two photochemical) was implemented in appropriate kinetic models for fitting or simulating the experimental data (see section 4).



Scheme S1. Complete reaction network for compound **2⁺** and **1** used to fit and simulate the experimental behavior. Reactions are numbered clockwise starting from the top. "K" and "k" indicate equilibrium and rate constants respectively. " $h\nu$ " and " Δ " indicate the photochemical and thermal ($E \rightarrow Z$ and $Z \rightarrow E$) isomerization processes respectively. Conventionally, the rates of reaction are positive for the reaction read from left to right and from top to bottom (gray arrows).

3.1 Cycling Rate

Under light irradiation, at the steady state the rates of all reactions must be equal in magnitude and their value corresponds to the cycling rate. The rate of reaction 3 was determined using the appropriate rate law (equation 1 of the manuscript) and the experimental concentration of complex, axle, and macrocycle. The steady state rates reported in the manuscript were measured by averaging the rate values over the last 15 minutes.

3.2 Cycle Quantum Yield

The cycle quantum yield (Φ_{cy}) represents the number of cycles performed in a given interval of time divided by the corresponding number of photons absorbed in the same time interval (equation S2). The number of moles of photons absorbed by the system per unit time was determined from the photons emitted by the light source considering the absorbance of the solution at 365 nm (section 2.2). The number of moles of pumps that completed a cycle in a given time is equal to the rate of cycling multiplied by the time interval and by the volume of the irradiated solution. The rate of cycling at steady state was measured from the rate of reaction 3 (section 3.1).

$$\Phi_{cy} = \frac{N_{cycles}}{N_{ph,abs}} = \frac{v_{cycling} \times \Delta t \times V}{q_{n,p}(1-10^{-A}) \times \Delta t} \quad (S2)$$

3.3 Kinetic Asymmetry

With reference to Scheme S1, the kinetic asymmetry (ratcheting constant, K_r) of the cycle can be rigorously quantified as:⁴

$$K_r = \frac{r_1 r_2 r_{-3} r_{-4}}{r_{-1} r_{-2} r_3 r_4} \quad (S3)$$

where, r_i and r_{-i} are the rates of the i -th process forward and backward respectively. Since the reactions describe a closed cycle and considering that the rate for processes 2 and 4 is the sum of the photochemical and thermal reaction rates ($r_i = r_i^{hv} + r_i^{\Delta}$), K_r can also be written in terms of the rate constants of the processes reported in Scheme S1:

$$K_r = \frac{k_1(k_2^{hv} + k_2^{\Delta})k_{-3}(k_{-4}^{hv} + k_{-4}^{\Delta})}{k_{-1}(k_{-2}^{hv} + k_{-2}^{\Delta})k_3(k_4^{hv} + k_4^{\Delta})} = K_a^E \frac{F_{hv}\epsilon_{EC}\phi_{EC} + k_2^{\Delta}}{F_{hv}\epsilon_{ZC}\phi_{ZC} + k_{-2}^{\Delta}} (K_a^Z)^{-1} \frac{F_{hv}\epsilon_Z\phi_Z + k_{-4}^{\Delta}}{F_{hv}\epsilon_E\phi_E + k_4^{\Delta}}$$

Where $F_{hv} = \frac{q_{n,p} \cdot b (1-10^{-A})}{V \cdot A}$ and $A = b \cdot \sum_i \epsilon_i \cdot [i]$.

In the dark ($q_{n,p} = 0$) F_{hv} is null and $K_r = 1$ due to the microscopic reversibility of the thermal processes.⁵ Conversely, in case of irradiation ($q_{n,p} > 0$), the kinetic constants for thermal isomerization reactions (indicated with a “ Δ ” superscript) can be neglected with respect to the photochemical ones. As a result, the light-dependent parts of the equation simplify to the photostationary state composition for axle and complex. Moreover, for the present system is fair to assume that these are equally populated ($K_{hv}^u = K_{hv}^c$) and therefore:

$$K_r \approx K_a^E \frac{\epsilon_{EC}\phi_{EC}}{\epsilon_{ZC}\phi_{ZC}} (K_a^Z)^{-1} \frac{\epsilon_Z\phi_Z}{\epsilon_E\phi_E} = K_a^E K_{hv}^c (K_a^Z)^{-1} (K_{hv}^u)^{-1} \approx \frac{K_a^E}{K_a^Z} = 2$$

In Table S3 are reported the absolute rates of reaction for all processes obtained from numerical simulations (see section 4.2) and the ratcheting constants at different photon flows calculated using equation S3.

Table S3. Computed forward and backward reaction rates obtained from numerical simulations and calculated ratcheting constants.

	$q_{n,p}$ (Einstein s^{-1})	4.4×10^{-9}	1.1×10^{-8}	2.2×10^{-8}	4.4×10^{-8}
thermal reactions	r_1 (10^{-6} M s^{-1})	5.70	5.34	5.33	5.90
	r_{-1} (10^{-6} M s^{-1})	5.65	5.24	5.16	5.59
	r_3 (10^{-6} M s^{-1})	0.625	0.575	0.526	0.599
	r_{-3} (10^{-6} M s^{-1})	0.676	0.679	0.692	0.906
photoreactions	r_2 (10^{-6} M s^{-1})	0.213	0.497	0.979	2.09
	r_{-2} (10^{-6} M s^{-1})	0.147	0.378	0.797	1.77
	r_4 (10^{-6} M s^{-1})	0.192	0.470	0.952	1.99
	r_{-4} (10^{-6} M s^{-1})	0.234	0.566	1.11	2.29
	$K_r^{[a]}$	1.9	1.9	2.0	2.2

^[a]Calculated using equation S3 from computed reaction rates.

4. Fitting and Numerical Simulations

4.1 Fitting of the dethreading kinetic

Second order threading rate constant ($k_{in}^E = k_1$) and equilibrium constant (K_a^E) for the *E* isomer were determined independently by UV-Vis and ^1H NMR.¹ The rate constant of dethreading of compound $Z\text{-}2^+$ (k_{-3}) was determined by fitting of the time-dependent concentration profiles of *Z* axle and *Z* complex reaching the local according to a kinetic model including the thermal reactions of Scheme S1. The average value obtained from four independent “relaxation” experiments (black trace in figure S4) is $(2.3 \pm 0.4) \times 10^{-4} \text{ s}^{-1}$, which is in excellent agreement with the previously reported data (Table S4) and was used in all the following calculations.

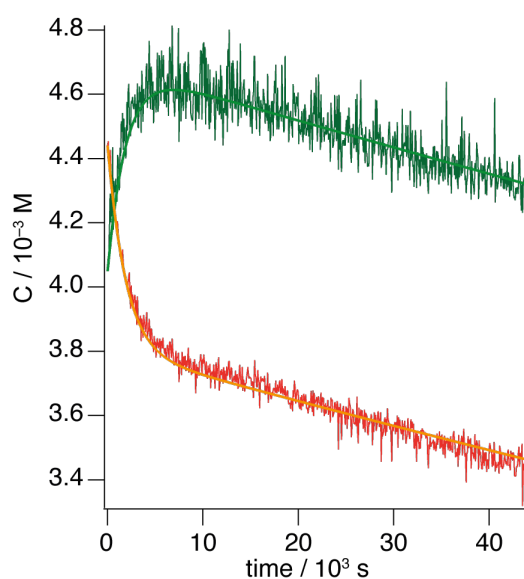


Figure S5. Typical kinetic of equilibration of a mixture of $Z\text{-}2^+$ and **1**, generated by fast light-induced isomerization of the corresponding equilibrated mixture of $E\text{-}2^+$ and **1**. Data obtained from ^1H NMR measurements (500 MHz, CD_3CN , 298 K). *Z* complex: red trace; *Z* axle: green trace. Solid lines represent the least square fitting according to the thermal reactions of scheme S1. Conditions: CD_3CN , $C = 8.9 \text{ mM}$, 298 K.

Table S4. Comprehensive table of the experimental thermodynamic, kinetic, photophysical, and photochemical parameters for axle $E-2^+$, $Z-2^+$, and for the corresponding [2]pseudorotaxane in air-equilibrated CH_3CN .^[a]

	K_a (M^{-1})	k_{in} ($\text{M}^{-1} \text{s}^{-1}$)	k_{out} (s^{-1})	k_{Δ} (s^{-1})	ϕ ^[b]	ϵ_{365} ($\text{M}^{-1} \text{cm}^{-1}$)
$E-2^+$	[c]	[c]	[c]	[c]	0.23 ($E \rightarrow Z$)	3900
$Z-2^+$	[c]	[c]	[c]	1.9×10^{-6}	0.58 ($Z \rightarrow E$)	< 100
$1 \cdot E-2^+$	230 ± 30	16 ^[d]	0.07 ^[e]	[c]	0.22 ($E \rightarrow Z$)	4800
$1 \cdot Z-2^+$	170 ± 30 115 ± 35 ^[f]	$3.1 \pm 0.8 \times 10^{-2}$ ^[g]	$2.7 \pm 0.5 \times 10^{-4}$ ^[g] $2.3 \pm 0.4 \times 10^{-4}$ ^[h]	4.8×10^{-6}	0.59 ($Z \rightarrow E$)	< 100

^[a]Data reported in ref. 1. *Italicized* values were newly determined in this work. ^[b]Determined at 365 nm. The corresponding isomerization process is given in brackets. ^[c]Parameter is not relevant for the compound. ^[d]Determined by stopped-flow UV-Vis absorption method. ^[e]Calculated as k_{in}/K_a . ^[f]Calculated as $k_{\text{in}}/k_{\text{out}}$. ^[g]Determined by non-linear regression of the time-dependent ^1H NMR concentration profiles. ^[h]Same as (g) but determined in *this work* by regression of the dethreading kinetic. This value was used as the basis for the k_{-3} in the numerical simulations (see table S4).

4.2 Numerical Simulations

The kinetic behavior of the reaction network was simulated taking into consideration all the thermal and photochemical reactions in scheme S1. The reaction rates for all processes were computed separately at any point in time solving the linear system of differential equations (S4 to S8) that defines the pumping cycle with a stiff ODE solver (ode15s).

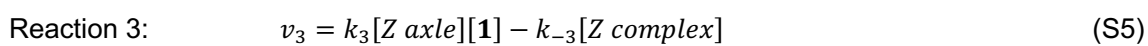
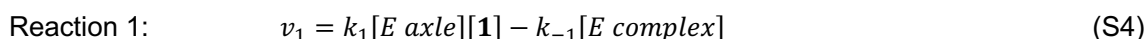
Concerning the implementation, the linear system describing the pumping cycle is time-variant (LTV), therefore simulations were carried out considering three time-invariant steps (LTI): a) equilibration, b) photoisomerization, and c) operation regime. The solutions from the previous LTI were taken as initial conditions for the subsequent one. The system is assumed to be in well-mixed conditions. Thus, diffusion of species is never considered rate limiting and all concentrations used in the rate laws are homogeneous in space. Details for the three steps are given below:

- First the macrocycle and E -axle association in the dark was simulated using the initial experimental concentrations of $E-2^+$ and **1** as initial conditions for the numerical solution. A photon flow of $0.0 \text{ Einstein s}^{-1}$ ("light off") was set in this step (photochemical rates are neglected). A close match between the simulated and experimental equilibrium composition in the dark was obtained.
- Starting from these equilibrium concentrations of E axle and E complex the isomerization was simulated using the photokinetic equations (eq. S8). In this step the photon flow was set to the value of $4.4 \times 10^{-8} \text{ Einstein s}^{-1}$ for about 35 min, consistently with the experimental

procedure. Pleasingly, the simulated PSS concentrations (>95% conversion to Z isomers) matched very closely the experimental ones (Fig. S6) validating this methodology of simulation and the actinometry results (section 2.3).

- c) Finally, from the simulated PSS concentrations the photon flow was set to the actual experimental value, and the system was allowed to reach the local equilibrium state or the kinetic steady state respectively in absence or presence of the light. Five possible photon flows ranging from 0.0 to 4.4×10^{-8} Einstein s^{-1} were used in the simulations in line with the experiments.

The rates of thermal reactions (1,3, 2^Δ , and 4^Δ) were computed with the corresponding rate equations (S4 to S7). The rate constants of reaction 3 were adjusted within the experimental error to achieve the best overlay with the experimental concentration profiles (table S4). The rate constants for the thermal $E \rightarrow Z$ isomerization reactions (k_2^Δ and k_4^Δ), included in Scheme 1 for completeness, were neglected in the simulations. As a result, reactions 2^Δ and 4^Δ appear as unidirectional processes that convert Z isomers to the corresponding E form (equations S6 and S7).



Where the Δ superscripts refer to the thermal $Z \rightarrow E$ isomerization process.

Rates of $E \rightarrow Z$ and $Z \rightarrow E$ photochemical isomerization for the free and complexed axles (reactions 2^{hv} and 4^{hv}) were calculated using the photokinetic rate law (S8). Experimental quantum yields (ϕ) and molar absorption coefficients (ε) determined at 365 nm, as well as the photon flow determined by chemical actinometry ($q_{n,p,max}$) were used. The absorbance of the mixture at the irradiation wavelength (A_{TOT}) was calculated according to Beer-Lambert's law at any point in time from the calculated mixture composition. Equation S8 is the general photokinetics equation⁶ for the photoisomerization of the i -th species ($i = E \text{ axle}, Z \text{ axle}, E \text{ complex}, Z \text{ complex}$).

$$v_i = \frac{q_{n,p} \cdot b}{V} \cdot \phi^i \cdot \varepsilon_i \cdot [i] \frac{1 - 10^{-A_{TOT}}}{A_{TOT}} \quad (\text{S8})$$

It must be noted that each photoactive species needs two photokinetic rate laws accounting for both $E \rightarrow Z$ and $Z \rightarrow E$ photochemical isomerization processes.

A total of four photokinetic equations plus the six rate laws are needed to appropriately simulate the pumping cycle (Scheme S1). In Fig. S6 (top row) the simulated concentration profiles of all the

photoactive species are displayed. In table S5 the parameters used in the dynamic simulations reported in Fig. 3 of the main text and Fig. S6 are reported.

Table S5. Experimental initial concentrations of **1** and $E\text{-}2^*$, rate constants, and photophysical parameters used for the numerical simulations at the different photon flows.^[a]

$q_{n,p}$ (Einstein s^{-1})	0	4.4×10^{-9}	1.1×10^{-8}	2.2×10^{-8}	4.4×10^{-8}
(% of $q_{n,p,max}$)	0%	10%	25%	50%	100%
$[2^+]_0$ (M) ^[b]	8.91×10^{-3}	7.43×10^{-3}	7.43×10^{-3}	7.43×10^{-3}	8.17×10^{-3}
$[1]_0$ (M) ^[b]	8.98×10^{-3}	7.48×10^{-3}	7.48×10^{-3}	7.48×10^{-3}	8.23×10^{-3}
k_3 ($M^{-1} s^{-1}$) ^[c]	3.65×10^{-2}	3.7×10^{-2}	3.7×10^{-2}	3.5×10^{-2}	3.65×10^{-2}
k_{-3} (s^{-1}) ^[d]	2.25×10^{-4}	2.1×10^{-4}	2.0×10^{-4}	2.0×10^{-4}	2.25×10^{-4}
ϵ_{EC}	4800	4800	4800	4800	4800
ϵ_E	3900	3900	3900	3900	3900
ϵ_{ZC} ^[e]	30	31	30	31	30
ϵ_Z ^[e]	40	40	40	40	40
$\phi^{E \rightarrow Z}$	0.23	0.23	0.23	0.23	0.23
$\phi^{Z \rightarrow E}$	0.58	0.58	0.58	0.58	0.58
$\phi^{EC \rightarrow ZC}$	0.22	0.22	0.22	0.22	0.22
$\phi^{ZC \rightarrow EC}$	0.59	0.59	0.59	0.59	0.59

^[a]The parameters shaded in orange were adapted within the confidence interval of the experimental value to achieve a better overlay with the data. ^[b]Experimental concentration. ^[c]Adapted in the interval $(3.1 \pm 0.8) \times 10^{-4} M^{-1} s^{-1}$. ^[d]Adapted in the interval $(2.3 \pm 0.4) \times 10^{-4} s^{-1}$. ^[e]Value calculated in order for the ratio $\frac{\phi^{EZ} \epsilon_E}{\phi^{ZE} \epsilon_Z}$ to match the experimental composition at the PSS; minor modifications were performed to achieve a better overlay with the experimental data.

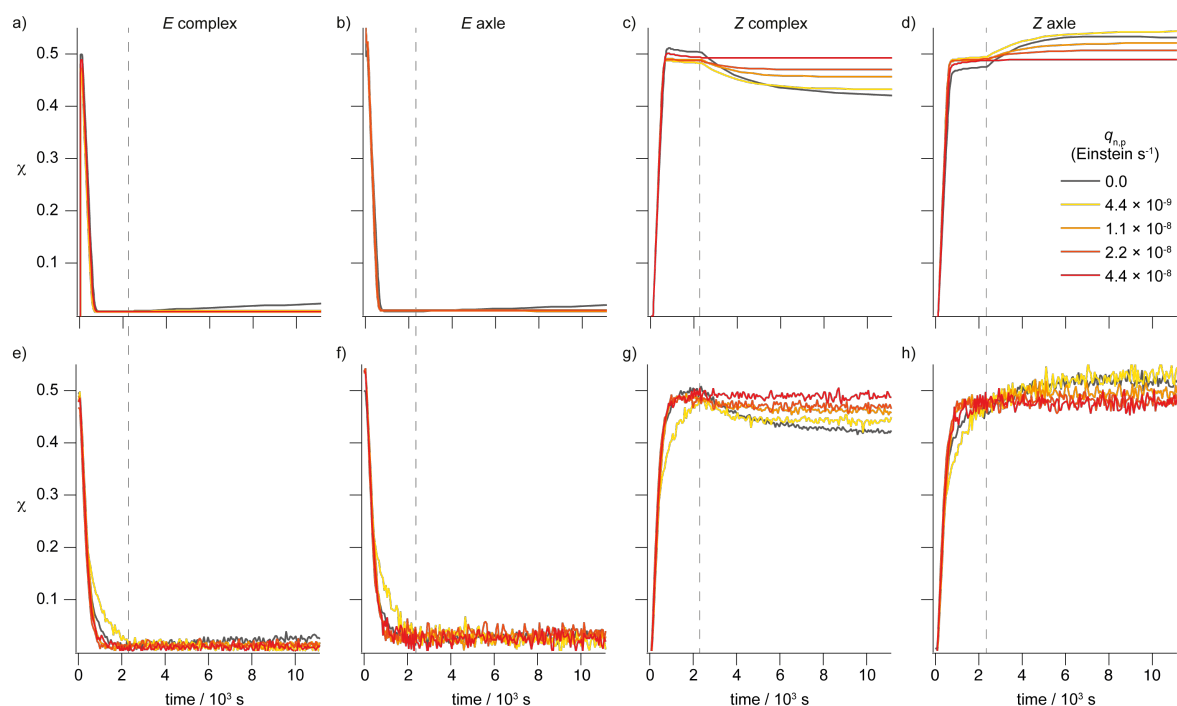


Figure S6. Simulated molar fraction profiles (top row) at different light intensities of all species generated by light-induced isomerization of an equilibrated mixture of *E-2** and **1**. On the bottom row the corresponding experimental profiles (Fig. S4) are reported for comparison. *E* complex (a, e), *E* axle (b, f), *Z* complex (c, g), and *Z* axle (d, h). The gray dashed line indicates the time at which PSS is reached. Data obtained from ¹H NMR data (500 MHz, CD₃CN, 298 K, initial concentrations for each experiment are reported in Table S4).

5. Thermodynamic Analysis

All the following thermodynamic considerations are made in the steady state regime, unless otherwise specified, that is, analogously to equilibrium condition, time-invariant. The steady state concentrations of all species extracted from time-dependent ¹H NMR experiments are reported in table S6.

Table S6. Experimental and simulated molar concentrations of **1**, *E-2**, *Z-2** and of the corresponding complexes at the local equilibrium (dark) or dissipative steady states (light on).

	$q_{n,p}$ (Einstein s ⁻¹)				
	Dark	4.4×10^{-9}	1.1×10^{-8}	2.2×10^{-8}	4.4×10^{-8}
[EC] (M) ^[a]	2.24×10^{-4}	8.07×10^{-5}	7.48×10^{-5}	7.34×10^{-5}	7.99×10^{-5}
[E] (M) ^[a]	2.00×10^{-4}	8.53×10^{-5}	8.32×10^{-5}	8.46×10^{-5}	9.01×10^{-5}
[ZC] (M) ^[b]	3.76×10^{-3}	3.28×10^{-3}	3.42×10^{-3}	3.49×10^{-3}	4.00×10^{-3}
[Z] (M) ^[b]	4.59×10^{-3}	3.96×10^{-3}	3.69×10^{-3}	3.57×10^{-3}	3.91×10^{-3}
[C] (M) ^[c]	4.95×10^{-3}	4.13×10^{-3}	3.94×10^{-3}	3.86×10^{-3}	4.13×10^{-3}

^[a]Determined from dynamics simulation. ^[b]Experimental molar concentrations calculated from the mole fractions (observable) averaged over the last 15 min and the initial concentration of axle. ^[c]Determined applying mass balance considerations.

5.1 Chemical Potential of Reaction

The non-standard chemical potential of reaction (also known as thermodynamic affinity when taken with a minus sign) was calculated from the concentrations maintained at the kinetic steady state away from equilibrium according to equation S9.⁷ At any point along the reaction coordinate the $\Delta\mu$ is given by the sum of the chemical potential of each species (μ_i) multiplied by the corresponding stoichiometric coefficient (negative for reagents). In particular, a positive value of $\Delta\mu$ corresponds to a steady state shifted towards the products ($\mu_P > \mu_R$), while the opposite holds for a negative $\Delta\mu$. For a generic reaction of the kind $A + B \rightarrow P$ the overall variation in chemical potential thus reads:

$$\Delta\mu = -\mu_A - \mu_B + \mu_P$$

Then:

$$\Delta\mu = -\mu_A^0 - RT \ln[A] - \mu_B^0 - RT \ln[B] + \mu_P^0 + RT \ln[P] = \underbrace{-\mu_A^0 - \mu_B^0 + \mu_P^0}_{\Delta\mu^0} + RT \ln \frac{[P]}{[A][B]}$$

$$\Delta\mu = RT \ln Q - RT \ln K \quad (\text{S9})$$

where K is the equilibrium constant of the reaction, Q is the reaction quotient. The sum of the non-standard chemical potential (thermodynamic affinities when taken with the opposite sign) along the cycle of scheme S1 is the free energy dissipated along one cycle of operation.

5.2 Thermodynamic of Radiation

Fig. S7a and b present a scheme of the irradiation setup used in this study, which served as the basis for the thermodynamic analysis. The first step of our thermodynamic analysis is to determine the concentration of photons in a volume element $dV_{cyl}(r)$ consisting of an infinitesimally thin cylindrical shell of thickness dr (inner radius r and outer radius $r + dr$) axial to the optical fiber and the NMR tube (highlighted in red Fig. S7c). The volume of the cylindrical shell is $dV_{cyl}(r) = \pi h((r + dr)^2 - r^2) = 2\pi h r dr$.

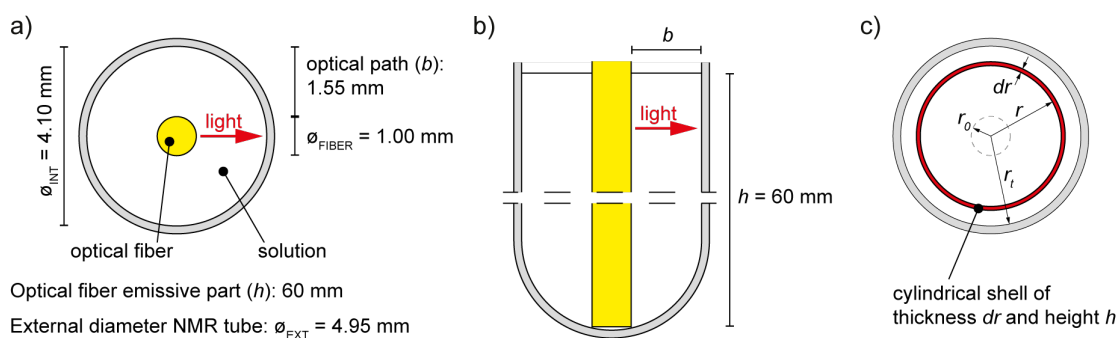


Figure S7. Scheme of the optical geometry and dimensions of the irradiation setup. a) Top view. b) Side view. c) Modelling of the infinitesimal volume travelled by light in a unit time used in the thermodynamic analysis ($r_0 = 0.5$ mm, $r_t = \varnothing_{INT}/2 = 2.05$ mm).

The moles of photons in this volume ($dN_{hv}(r)$) can be computed by multiplying the number of moles of photons impinging on the inner surface per unit time ($q_{n,p}(r)$) by the time that light takes to travel the distance dr ($t = dr/c$, where “ c ” is the speed of light):

$$dN_{hv}(r) = q_{n,p}(r) \cdot t = \frac{q_{n,p}(r)}{c} dr$$

The molar concentration of photons in the volume $dV_{cyl}(r)$ is then:

$$n_{hv}(r) = \frac{dN_{hv}(r)}{dV_{cyl}(r)} = \frac{q_{n,p}(r)}{2\pi hcr} = \frac{q_{n,p}}{2\pi hc} \cdot \frac{10^{-(r-r_0)\sum_i \varepsilon_i [\textit{i}]}}{r}$$

where the dependence of $q_{n,p}(r)$ on r has been expressed in terms of the Beer-Lambert's law and the total amount of photons emitted by the optical fiber ($q_{n,p}(r_0) \equiv q_{n,p}$). It is worth noting that the molar concentration of photons in the volume $dV_{cyl}(r)$ has, in principle, also a contribution n_{bb}^T (constant along r) from the black body emission of the solution due to its temperature T . However, such a contribution is negligible at any r when compared to that of the source at the wavelength and intensity considered.

To proceed with our analysis, we now define the effective temperature $T_{hv}(r)$ associated with the radiation in the volume $dV_{cyl}(r)$. Such a temperature is the one that a black body needs to have in order to generate a molar concentration of photons equal to $n_{hv}(r)$ in the interval selected by the interference filter (365 ± 5 nm), and can be computed by solving numerically the following equation:

$$n_{hv}(r) = \int_{360 \text{ nm}}^{370 \text{ nm}} n_{bb}(\lambda) d\lambda = \int_{360 \text{ nm}}^{370 \text{ nm}} \frac{8\pi}{N_A \lambda^4} \cdot \frac{1}{e^{hc/\lambda k_B T_{hv}} - 1} d\lambda$$

where N_A is the Avogadro's number, h the Planck's constant, and k_B the Boltzmann's constant. The above equation uniquely defines $T_{hv}(r)$ and is well justified within the monochromatic assumption. Indeed, as all the photons impinging on the system are considered at the same frequency, the exact shape of their distribution in frequency in the interval selected by the interference filter has no role in the treatment, and can therefore be assumed to be the black body one ($n_{bb}(\lambda)$) without altering the final results. This has the major advantage that the radiation in the volume $dV_{cyl}(r)$ can be considered as a heat source at temperature $T_{hv}(r)$. Since the solvent acts as another heat source at temperature T , we can properly think of the molecular motor as a

thermal engine working by virtue of the temperature gradient between the radiation and the solvent. As a consequence, we can immediately conclude that the maximum efficiency at which the motor can convert the energy absorbed from the radiation into work in the volume $dV_{cyl}(r)$ is limited by Carnot's theorem:

$$\eta_C(r) = 1 - \frac{T}{T_{hv}(r)}$$

We can have a first idea of the thermodynamics of the motor by computing the average temperature of the radiation and the average Carnot's efficiency (Table S7).

$$\langle T_{hv} \rangle = \frac{1}{b} \int_{r_0}^{r_t} T_{hv}(r) dr$$

$$\langle \eta_C \rangle = \frac{1}{b} \int_{r_0}^{r_t} \eta_C(r) dr$$

Note that, by virtue of the inequality holding between the harmonic mean and the arithmetic mean of a positive defined function, we have that:

$$\langle \eta_C \rangle = 1 - T \left\langle \frac{1}{T_{hv}} \right\rangle \leq 1 - \frac{T}{\langle T_{hv} \rangle}$$

and therefore the Carnot's efficiency computed by using the average temperature of the hot heat source ($\langle T_{hv} \rangle$) provides an upper bound to the maximum efficiency of the whole system considered as a thermal engine.

To have further quantitative insights on the thermodynamic characterization of the system, we introduce the chemical potential of the radiation as:⁸

$$\mu_{hv}(r) = N_A h\nu \cdot \eta_C(r)$$

In this context, the chemical potential of the radiation can be interpreted as the part of the molar energy coming from the light source which is actually available to the system in the volume $dV_{cyl}(r)$ (the free energy in the volume). It is also the amount of energy which can be used to take the system away from thermodynamic equilibrium by performing work on it. From the above expression, whenever the concentration of photons in the volume $dV_{cyl}(r)$ is such that $T_{hv}(r) = T$, with a molar concentration of photons equal to n_{bb}^T , corresponding to the black body emission of the solution at the temperature T , the chemical potential of the radiation is null. Therefore, the radiation cannot perform any work on the system in that volume element. In this condition which

corresponds to the absence of radiation from the source the only possible steady state for the system is the equilibrium one. In Table S7, the average chemical potential for the five experimental regimes is reported. The average chemical potential of the radiation in the system is way larger than the thermal energy (RT) and about one order of magnitude higher than the free energy released upon hydrolysis of ATP to ADP in physiological conditions. Moreover, by increasing the light intensity, the average chemical potential of the radiation also increases, consistently with the system being brought farther away from equilibrium at higher photon fluxes.

Table S7. Average temperature of radiation ($\langle T_{hv} \rangle$), Carnot's efficiency ($\langle \eta_C \rangle$), and chemical potential of radiation ($\langle \mu_{hv} \rangle$) at the investigated photon flows.^[a]

$q_{n,p}$ (Einstein s ⁻¹)	$\langle T_{hv} \rangle$ (K)	$\langle \eta_C \rangle$	$\langle \mu_{hv} \rangle$ (kJ mol ⁻¹) ^[a]
0	298	0	0 ^[b]
4.4×10^{-9}	1917	0.844	277
1.1×10^{-8}	2007	0.851	279
2.2×10^{-8}	2081	0.857	281
4.4×10^{-8}	2159	0.862	282

^[a]The energy of one mole of 365 nm photons was estimated as $N_A h \nu = 328$ kJ mol⁻¹. ^[b]The actual value of photon flow at which the force is null is not 0 Einstein s⁻¹, but rather a very small value of photon flow corresponding to a black body at 298 K emitting around 365 nm.

5.3 Energy Dissipation and Storage

We now couple the thermodynamic description with the system's kinetics. The number of moles of photons absorbed by the system per unit of time in the volume element $dV_{cyl}(r)$ is:

$$\begin{aligned} dq_{n,p}(r) &= q_{n,p}(r) - q_{n,p}(r + dr) = q_{n,p}(r)(1 - 10^{-dr \sum_i \varepsilon_i [i]}) \\ &= q_{n,p}(r) \sum_i \varepsilon_i [i] dr + \mathcal{O}(dr^2) \\ &\approx q_{n,p}(r) \sum_i \varepsilon_i [i] dr \end{aligned}$$

from which we can compute the free energy absorbed by the system per unit of time (\dot{W}) in the same volume element as:

$$d\dot{W}(r) = \mu_{hv}(r) dq_{n,p}(r) = \mu_{hv}(r) q_{n,p}(r) \sum_i \varepsilon_i [i] dr$$

However, since the quantum yield of photoisomerization is not unitary for the considered processes, the actual amount of absorbed free energy which can perform work on the system, thus modifying the species concentrations, is:

$$d\dot{W}_{hv}(r) = \mu_{hv}(r)q_{n,p}(r) \sum_i \phi_i \varepsilon_i [i] dr$$

Were the quantum yields (ϕ_i) are included to account for the amount of absorbed free energy which is unavoidably dissipated without leading to photoisomerization.

The integral of $d\dot{W}_{hv}(r)$ over the space occupied by the solution and divided by the volume yields the input power of the molecular motor which appears in equation (2) of the main text, that is the work performed on the system by the radiation:

$$\dot{W}_{hv} = \frac{1}{V} \sum_i \phi_i \varepsilon_i [i] q_{n,p} \int_{r_0}^{r_t} \mu_{hv}(r) 10^{-(r-r_0)\sum_i \varepsilon_i [i]} dr$$

At any time, the free energy absorbed per unit time and volume can either be dissipated or stored in the system as free energy. This is expressed by equation (2) of the main text:

$$\dot{W}_{hv} = d_t G + T\dot{\Sigma}$$

where the dissipation rate per unit of volume is always positive ($T\dot{\Sigma} \geq 0$) according to the second law of thermodynamics. Both the dissipation rate and the time derivative of the Gibbs free energy can be split into two contributions, one due to the isomerization steps ($T\dot{\Sigma}_{iso}$), and the other to the self-assembly steps ($T\dot{\Sigma}_{sa}$):

$$\begin{aligned} T\dot{\Sigma} &= T\dot{\Sigma}_{sa} + T\dot{\Sigma}_{iso} \\ d_t G &= d_t G_{sa} + d_t G_{iso} \end{aligned}$$

5.3.1 Energy dissipation

The part of the absorbed power which is dissipated by the self-assembly steps is calculated as:

$$T\dot{\Sigma}_{sa} = -(v_1 \Delta\mu_1 + v_3 \Delta\mu_3)$$

At the stationary state, since the Gibbs free energy is a state function, all the power absorbed by system is dissipated to sustain the non-equilibrium condition ($d_t G = 0$):

$$\dot{W}_{hv} = T\dot{\Sigma} = T\dot{\Sigma}_{sa} + T\dot{\Sigma}_{iso}$$

with $T\dot{\Sigma}_{iso} = \dot{W}_{hv} - T\dot{\Sigma}_{sa} \geq 0$. The above equation allows for the following thermodynamic interpretation. At the steady state, to maintain the free energy storage, that is keeping the self-assembly reactions away from equilibrium ($T\dot{\Sigma}_{sa} > 0$), part of the power absorbed by the photoisomerization steps (\dot{W}_{hv}) needs to be transferred to the self-assembly ones, allowing them to dissipate.

This means that the supramolecular pump operates as a (thermal) engine transducing free energy from the photons to the self-assembly processes through the photoisomerization processes. Therefore, free energy storage, and consequently unidirectional motion of the ring-axle pairs, are sustained by light energy absorption.

At the steady state $v_1 = -v_3 = v_{cy}$, thus $T\dot{\Sigma}_{sa} = -v_{cy}(\Delta\mu_1 - \Delta\mu_3)$. Upon integration of this equation over the time it takes for a mole of rings to complete one cycle ($\tau_{cy} = 1/v_{cy}V$) and multiplying by the volume, we find that the free energy dissipated by the motor per cycle of operation at the steady state reported in the main text is given by:

$$T\Delta_{cy}\Sigma_{sa} = -(\Delta\mu_1 - \Delta\mu_3)$$

In turns, this value also sets a limit to the amount of work which self-assembly steps can perform per mole of rings which complete one pumping cycle.

5.3.2 Energy storage

The free energy density stored in the self-assembly steps at steady state can be defined by virtue of the timescale separation between the relaxation times of the self-assembly steps and the isomerization ones. In fact, once the photon flow is stopped, the self-assembly steps reach their (local) equilibrium much faster than the isomerization reactions, which can equilibrate only via the very slow thermal relaxation steps. Therefore, the relative amounts of *E* and *Z* species are preserved. In these conditions, the free energy released by the relaxation to equilibrium of the self-assembly steps alone corresponds to the fraction of free energy they stored. In the following the suffix “dark” indicates that the quantity is referred to the local equilibrium, while the other quantities are considered at the dissipative steady state. The following mass balances hold true upon relaxation to the local equilibrium:

$$\begin{aligned} [EC] + [E] &= [EC]_{dark} + [E]_{dark} \\ [ZC] + [Z] &= [ZC]_{dark} + [Z]_{dark} \\ [EC] + [ZC] + [C] &= [EC]_{dark} + [ZC]_{dark} + [C]_{dark} \end{aligned}$$

Moreover, for the concentration at local equilibrium (in dark) the following relations hold true:

$$K_E = \frac{k_1}{k_{-1}} = \frac{[EC]_{dark}}{[E]_{dark}[C]_{dark}}$$

$$K_Z = \frac{k_3}{k_{-3}} = \frac{[ZC]_{dark}}{[Z]_{dark}[C]_{dark}}$$

With the above five equations, we can uniquely compute the (local) equilibrium concentrations ($[i]_{dark}$) by just knowing the steady state concentration in an experimental regime and the in/out kinetic rate constants of the self-assembly steps. The free energy stored in the self-assembly steps per unit volume at the steady state reached in any experimental regime as reported in the main text can then be computed as:

$$\Delta G_{sa} = G - G_{dark}$$

Since the Gibbs free energy of the system per unit of volume is $G = \sum_i [i](\mu_i - RT)$, then the stored energy density becomes:⁹

$$\Delta G_{sa} = \sum_i [i](\mu_i - RT) - [i]_{dark}(\mu_{i,dark} - RT) = RT \sum_i [i] \ln \frac{[i]}{[i]_{dark}} - [i] + [i]_{dark}$$

which was used to calculate the energy storage reported in the main text.

5.4 Energy Transduction Efficiency

The efficiency at which the motor, at steady state, converts the light energy into chemical energy available to the self-assembly reactions can be defined in two ways. The first, reported in the main text (η), consists in evaluating the fraction of work performed by the radiation (free energy) on the system which is successfully transferred to the self-assembly steps:

$$\eta = \frac{T\dot{\Sigma}_{sa}}{\dot{W}_{hv}}$$

This quantity is bounded between 0 and 1 and, being based on the free energy gradient generated by the radiation ($\langle \mu_{hv} \rangle$), can be directly compared between different motors regardless of the provided fuel. For example, the same quantity computed for a chemically driven rotary motor results to be 5 orders of magnitude lower.¹⁰ This can be explained by considering that the chemically driven rotary motor synthesized by the Leigh group works by transducing free energy only in form of information, while here $T\Delta_{cy}\Sigma_{sa}$ also have an energetic component due to the energy ratchet mechanism.

Additionally, in the spirit of Carnot's efficiency of thermal machines, we can evaluate the ratio between the free energy transduced to the self-assembly steps per unit time and unit volume (dissipation $T\dot{\Sigma}_{sa}$), and the heat absorbed from the radiation (the hot reservoir) per unit time and unit volume (\dot{Q}_{hv}), according to the following equation:

$$\eta_2 = \frac{T\dot{\Sigma}_{sa}}{\dot{Q}_{hv}} \leq \langle \eta_C \rangle$$

Where $\dot{Q}_{hv} = \frac{N_A hc}{\lambda v} q_{n,p} (1 - 10^{-b \sum_i \varepsilon_i [i]})$ is the total power absorbed by the sample in the unit time and volume. The fact that this efficiency is upper limited by the average Carnot efficiency (proof in section 5.4.1) makes η_2 an interesting quantity to understand how close the motor, seen as a thermal engine, works to the theoretical limit imposed by Carnot efficiency. The comparison is done in Table S8, which shows that at best this motor reaches 0.04% of Carnot efficiency and decreases with the photon flow. All the considerations in the main text aimed at rationalizing the trend of the energy transduction efficiency (η) hold for the Carnot-like efficiency η_2 .

Table S8. Ratio between the average Carnot-like and Carnot's efficiency ($\eta_2/\langle \eta_C \rangle$) at the investigated photon flows.

	$q_{n,p}$ (Einstein s ⁻¹)			
	4.4×10^{-9}	1.1×10^{-8}	2.2×10^{-8}	4.4×10^{-8}
$\eta_2/\langle \eta_C \rangle$ (%)	0.04	0.03	0.02	0.01

5.4.1 Proof of the bound $\eta_2 \leq \langle \eta_C \rangle$

Consider the following expression of the Second Law of thermodynamics as it applies for the system under study inside a volume element $dV_{cyl}(r)$ at the steady state:

$$T\dot{\Sigma}(r) = T\dot{\Sigma}_{iso}(r) + T\dot{\Sigma}_{sa} = \frac{d\dot{W}(r)}{dV_{cyl}(r)} - T\dot{\Sigma}_{sa} + T\dot{\Sigma}_{sa} = \dot{Q}_{hv}(r)\eta_C(r) - T\dot{\Sigma}_{sa} + T\dot{\Sigma}_{sa} \geq 0$$

where $T\dot{\Sigma}_{iso}(r)$ and $T\dot{\Sigma}_{sa}$ are both positive quantities and $\dot{Q}_{hv}(r) = N_A h v \frac{dq_{n,p}(r)}{dV_{cyl}(r)}$ is the heat per unit volume which is absorbed by the system in the volume element. Note that $T\dot{\Sigma}_{sa}$ does not depend on r due to the homogeneity in concentrations. From the above equation follows that:

$$0 \leq \frac{T\dot{\Sigma}_{sa}}{\underbrace{\dot{Q}_{hv}(r)\eta_C(r)}_{\eta}} = 1 - \frac{T\dot{\Sigma}_{iso}(r)}{\dot{Q}_{hv}(r)\eta_C(r)} \leq 1$$

and therefore $\eta_2(r) \leq \eta_c(r)$, with $\eta_2(r) = \frac{T\dot{\Sigma}_{sa}}{\dot{Q}_{hv}(r)}$.

By averaging over the entire system (see section 5.2) we have:

$$\langle \eta_2 \rangle = T\dot{\Sigma}_{sa} \left\langle \frac{1}{\dot{Q}_{hv}} \right\rangle \leq \langle \eta_c \rangle$$

And by virtue of the inequality holding between the harmonic and the arithmetic mean of a positive defined function, we have that:

$$\left\langle \frac{1}{\dot{Q}_{hv}} \right\rangle \geq \frac{1}{\langle \dot{Q}_{hv} \rangle} \equiv \frac{1}{\dot{Q}_{hv}}$$

and therefore $T\dot{\Sigma}_{sa} \left\langle \frac{1}{\dot{Q}_{hv}} \right\rangle \geq \frac{T\dot{\Sigma}_{sa}}{\langle \dot{Q}_{hv} \rangle} = \frac{T\dot{\Sigma}_{sa}}{\dot{Q}_{hv}} = \eta_2$, which proves the inequality $\eta_2 \leq \langle \eta_c \rangle$.

7. References

- Corra, S., Casimiro, L., Baroncini, M., Groppi, J., La Rosa, M., Tranfić Bakić, M., Silvi, S. & Credi, A. Artificial Supramolecular Pumps Powered by Light. *Chem. Eur. J.* **27**, 11076–11083 (2021).
- L. Vetráková, V. Ladányi, J. A. Anshori, P. Dvořák, J. Wirz, D. Heger The Absorption Spectrum of cis-Azobenzene. *Photochem. Photobiol. Sci.* **16**, 1749–1756 (2017).
- V. Ladányi, P. Dvořák, J. A. Anshori, L. Vetráková, J. Wirz, D. Heger Azobenzene Photoisomerization Quantum Yields in Methanol Redetermined. *Photochem. Photobiol. Sci.* **16**, 1757–1761 (2017).
- Astumian, R. D. & Bier, M. Mechanochemical Coupling of the Motion of Molecular Motors to ATP Hydrolysis. *Biophys. J.* **70**, 637-653 (1996).
- Onsager, L. Reciprocal Relations in Irreversible Processes. I. *Phys. Rev.* **37**, 405–426 (1931).
- Mausser, H. & Gauglitz, G. *Photokinetics: theoretical fundamentals and applications* (Elsevier, 1998).
- Demirel, Y. & Gerbaud, V. Chapter 8: Chemical Reactions. *Nonequilibrium Thermodynamics Transport and Rate Processes in Physical, Chemical and Biological Systems* (Elsevier, 2019).
- a) Ries, H. & McEvoy, A. J. Chemical Potential and Temperature of Light. *J. Photochem. Photobiol. A* **59**, 11–18 (1991); b) Gräber, P. & Milazzo, G. *Bioenergetics* (Birkhäuser, 1997); c) P. Würfel, P. & Würfel, U. *Physics of Solar Cells: From Basic Principles to Advanced Concepts* (Wiley-VCH, 2016); d) Penocchio, E., Rao, R. & Esposito, M. Nonequilibrium Thermodynamics of Light-Induced Reactions. *J. Chem. Phys.* **155**, 114101 (2021).
- The relation is obtained by considering that $\sum_i [i]_{dark} \mu_{i,dark} = \sum_i [i] \mu_{i,dark}$, which can be proven via chemical reaction network techniques by noticing that the vector with elements $\mu_{i,dark}$ is a left-null vector of the stoichiometric matrix when only self-assembly steps are considered.
- a) Wilson, M. R., Sola, J., Carlone, A., Goldup, S. M., Lebrasseur, N. & Leigh, D. A. An Autonomous Chemically Fueled Small-Molecule Motor. *Nature* **534**, 235-240 (2016); b) Amano, S., Esposito, M. Kreidt, E., Leigh, D. A., Penocchio, E. & Roberts, B. M. W. Insights from an Information Thermodynamics Analysis of a Synthetic Molecular Motor. *Nat. Chem.* (2022) doi: 10.1038/s41557-022-00899-z.

4

TOWARDS AN INFORMATION THERMODYNAMIC PERSPECTIVE

The following article is reprinted from
[E. PENOCCHIO, F. AVANZINI and M. ESPOSITO, *arXiv* 2204.02815. (2022), Ac-
cepted in *The Journal of Chemical Physics*]
under the conditions of the Creative Commons CC BY 4.0 license¹.

The page numbers placed in the outer margins provide a continuous pagi-
nation throughout the thesis.

¹ <https://creativecommons.org/licenses/by/4.0/>

Information Thermodynamics for Deterministic Chemical Reaction Networks

Emanuele Penocchio,^{*} Francesco Avanzini,[†] and Massimiliano Esposito[‡]

Complex Systems and Statistical Mechanics, Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg

(Dated: July 4, 2022)

Information thermodynamics relates the rate of change of mutual information between two interacting subsystems to their thermodynamics when the joined system is described by a bipartite stochastic dynamics satisfying local detailed balance. Here, we expand the scope of information thermodynamics to deterministic bipartite chemical reaction networks, namely, composed of two coupled subnetworks sharing species, but not reactions. We do so by introducing a meaningful notion of mutual information between different molecular features, that we express in terms of deterministic concentrations. This allows us to formulate separate second laws for each subnetwork, which account for their energy and information exchanges, in complete analogy with stochastic systems. We then use our framework to investigate the working mechanisms of a model of chemically-driven self-assembly and an experimental light-driven bimolecular motor. We show that both systems are constituted by two coupled subnetworks of chemical reactions. One subnetwork is maintained out of equilibrium by external reservoirs (chemostats or light sources) and powers the other via energy and information flows. In doing so, we clarify that the information flow is precisely the thermodynamic counterpart of an information ratchet mechanism only when no energy flow is involved.

I. INTRODUCTION

Engines commonly operate such that some components (e.g., pistons) directly interact with the power source to harvest energy, whereas some other components (e.g., wheels) produce the functionality the engine is designed for. Chemical engines [1], such as molecular motors, are no exception [2–4]. Indeed, they can be rationally described [5–9] and designed [1, 10–13] as chemical reaction networks (CRNs) where energy-harvesting chemical [14–20], photochemical [21–27] or electrochemical [27–29] processes are coupled with large-amplitude intramolecular motions or self-assembly reactions. These CRNs are often bipartite [30, 31], as the energy-harvesting processes act on some molecular properties (e.g., phosphorylation state, photo-isomerization state, oxidation/reduction state) while the processes realizing the engine’s functionality act on different properties (e.g., position in space, assembly state). When this is the case, the chemical species involved in the functioning of a chemical engine can be described with a double state (x, y) , and the whole network can be split into two coupled subnetworks characterizing the interconversion of x -states by one kind of processes (e.g., self assembly steps converting free states into assembled ones) possibly driven out-of-equilibrium by processes of another kind interconverting y -states (e.g., reaction with a chemical fuel acting as a substrate).

Assessing the thermodynamics of bipartite systems requires quantifying energy and information exchanges between the subnetworks by applying the tools of information thermodynamics [30–33]. Linear stochastic models of biochemical machines comprising unimolecular or pseudo-unimolecular reactions have been extensively studied from this perspective, revealing the fundamental role of information flows between different degrees of freedom in powering

such single molecule machines [34–39]. However, chemical engines may in general comprise nonlinear processes such as bimolecular reactions and are sometimes (almost always in case of synthetic ones) better described by deterministic dynamics expressed in terms of kinetic equations evolving experimentally measurable concentrations rather than probabilities. At present, a theoretical framework able to systematically address the information thermodynamics of nonlinear deterministic CRNs is missing.

Very recently, we applied information thermodynamics to an autonomous synthetic molecular motor working at the deterministic level [40]. Such analysis shows that information thermodynamics is in principle not limited to stochastic setups, but it strongly relies on the fact that the CRN describing the motor is only composed of unimolecular reactions. Indeed, when this is the case, kinetic equations can be mapped into a Markov jump process on a linear network of states by normalizing the concentrations by the total concentration, which is a conserved quantity in linear networks. Once this correspondence is in place, standard tools from information thermodynamics can be applied [31]. In particular, the mutual information [41], a central quantity in information thermodynamics, can be defined at the level of normalized concentrations [40].

In this paper, we go further by extending the framework of information thermodynamics to nonlinear deterministic CRNs, where multimolecular reactions cause the total concentration to vary in time. We start by introducing the setup of deterministic CRNs in Section II, where the crucial notion of bipartite networks is formally defined and used to identify the two subnetworks, which influence each other despite the fact that each reaction pertains unambiguously to only one subnetwork. In Section III, thermodynamic quantities are introduced for both chemically-driven and light-driven CRNs, and the second law for nonequilibrium regimes is formulated. The main result of this paper is obtained in Section IV, where the notion of mutual information for non-normalized concentration distributions is defined and used to formulate a second law for each subnetwork. Crucially, these second laws

^{*} emanuele.penocchio@uni.lu

[†] francesco.avanzini@uni.lu

[‡] massimiliano.esposito@uni.lu

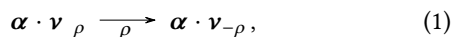
show that the two subnetworks exchange free-energy through information and energy flows. Our result is analogous to the one obtained in the framework of stochastic thermodynamics of bipartite systems [30, 31], but importantly information flow terms account here for the non-normalized concentration distributions. Furthermore, the resulting expressions of the subnetworks' entropy production also consider the variation of the subnetworks entropy due to the standard molar entropy carried by the chemical species and the contribution of non-bipartite species which may be present. To exemplify the use of our new framework, in Section V we apply it to two paradigmatic examples of out of equilibrium chemistry with nonlinear dynamics: a model for chemically-driven self-assembly [13, 42] and a light-driven bimolecular motor [24, 27, 43–45]. We briefly recap the basis of their functioning while focusing on the insights brought by the information thermodynamic analysis. Using numerical simulations, we also comment on their efficiency and we quantitatively investigate the correspondence between the concept of Brownian information ratchet [10, 46] and (thermodynamic) information flows.

Our approach can be applied to any CRN with a bipartite structure and coupled to any kind of reservoir. We restricted the presentation to the case of autonomous and homogeneous ideal dilute solutions, but extensions to nonautonomous [47], non-homogeneous [48, 49] and nonideal [50] CRNs are possible, as well as to cases where external species are continuously injected [51]. Based on the relevance of information thermodynamics as a tool to elucidate free-energy processing in the stochastic realm, we envision that our new framework will bring significant insights in the understanding of deterministic chemical processes, from synthetic molecular machines to complex biochemical networks.

Finally, we note that many forms of information processing are possible in CRNs [52, 53]: e.g., logic gates [54, 55], machine learning [56–58], sensing [59], copying and proof-reading [60–63], memories [64, 65]. At this stage, how the present framework may be suitable to study all of them is left for future inquiries.

II. SETUP

CRNs are treated here as ideal dilute solutions composed of chemical species, identified by the label $\alpha \in Z$, which undergo elementary [66] or coarse-grained [67–69] chemical reactions, identified by the index $\rho \in \mathcal{R}$:



with $\boldsymbol{\alpha} = (\dots, \alpha, \dots)^\top$ denoting the vector of chemical species, while $\boldsymbol{\nu}_{\rho}$ and $\boldsymbol{\nu}_{-\rho}$ denote the vectors of stoichiometric coefficients of reagents and products of reaction ρ . Throughout this paper, we will use single arrows like in Eq. (1) to denote reactions that are actually reversible: for every reaction $\rho \in \mathcal{R}$, the forward (resp. backward) reaction ρ (resp. $-\rho$) interconverts $\boldsymbol{\alpha} \cdot \boldsymbol{\nu}_{\rho}$ (resp. $\boldsymbol{\alpha} \cdot \boldsymbol{\nu}_{-\rho}$) into $\boldsymbol{\alpha} \cdot \boldsymbol{\nu}_{-\rho}$ (resp. $\boldsymbol{\alpha} \cdot \boldsymbol{\nu}_{\rho}$). This choice will simplify the hypergraph rep-

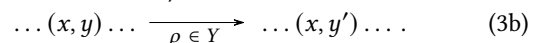
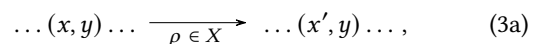
resentation of the two CRNs examined in Sec. V (see Figs. 2 and 3).

In deterministic CRNs, the abundance of the chemical species is specified by the concentration vector $\boldsymbol{z} = (\dots, [\alpha], \dots)^\top$, which follows the rate equation

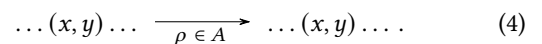
$$d_t \boldsymbol{z} = \mathbb{S} \boldsymbol{j}, \quad (2)$$

where we introduced the stoichiometric matrix \mathbb{S} and the current vector \boldsymbol{j} . Each ρ column \mathbb{S}_{ρ} of the stoichiometric matrix \mathbb{S} specifies the net variation of the number of molecules for each species undergoing the ρ reaction (1), $\mathbb{S}_{\rho} = \boldsymbol{\nu}_{-\rho} - \boldsymbol{\nu}_{\rho}$. The current vector $\boldsymbol{j} = (\dots, j^{\rho}, \dots)^\top$ specifies the net reaction current for every ρ reaction (1) as the difference between the forward and backward flux, i.e., $j^{\rho} = j^{\rho} - j^{-\rho}$. In *closed* CRNs, where the chemical reactions involve only the chemical species Z , the fluxes $j^{\pm\rho}$ depend only on the concentrations \boldsymbol{z} . For elementary reactions satisfying mass action kinetics [70–72], $j^{\pm\rho} = k_{\pm\rho} \boldsymbol{z}^{\boldsymbol{\nu}_{\pm\rho}}$, with $k_{\pm\rho}$ the kinetic constants (hereafter, for every vectors \boldsymbol{v} and \boldsymbol{w} , $\boldsymbol{v}^{\boldsymbol{w}} = \prod_i v_i^{w_i}$). In *open* CRNs, some chemical reactions exchange for instance matter and/or photons with external reservoirs. Thus, the corresponding fluxes $j^{\pm\rho}$ depend on the concentrations \boldsymbol{z} as well as on the coupling mechanisms with the reservoirs [51]. Note that, throughout the manuscript, we will omit for compactness of notation the functions' variables, e.g., $\boldsymbol{z} = \boldsymbol{z}(t)$ and $j^{\pm\rho} = j^{\pm\rho}(\boldsymbol{z})$ (if reaction ρ is not coupled to any reservoir).

We now define the formal conditions for the chemical species and reactions under which CRNs have a *bipartite* structure. An exemplification of the concept is provided in Figure 1. First, the chemical species must split as $Z = A \cup B$. The species b – the *bipartite* species – are univocally identified with a double *state* $(x, y) = \alpha \in b$. The species a – the *ancillary* species – are all the other species, when present. Second, the set of chemical reactions must split as $\mathcal{R} = A \cup X \cup Y$. The reactions X (resp. Y) interconvert the species b in such a way that only their x (resp. y) state changes:



In the above chemical reactions, we neglected the stoichiometric coefficients as well as the other species involved for illustrative reasons. The reactions A are those that do not interconvert the species b ($S_{\rho}^{\alpha} = 0 \forall \alpha \in b$ and $\forall \rho \in A$):



Notice that, when these conditions for the chemical species and reactions are satisfied, the species a can undergo every reaction $\rho \in \mathcal{R} = A \cup X \cup Y$ and no reactions can change both the state x and the state y of a b species. When these conditions are not satisfied, CRNs do not have a bipartite structure and their splitting into coupled subnetworks (discussed in the next paragraphs of this section) would not be possible. Finally, if the chemical reactions are coarse-grained, we further assume that the sets A , X and Y are independent sets

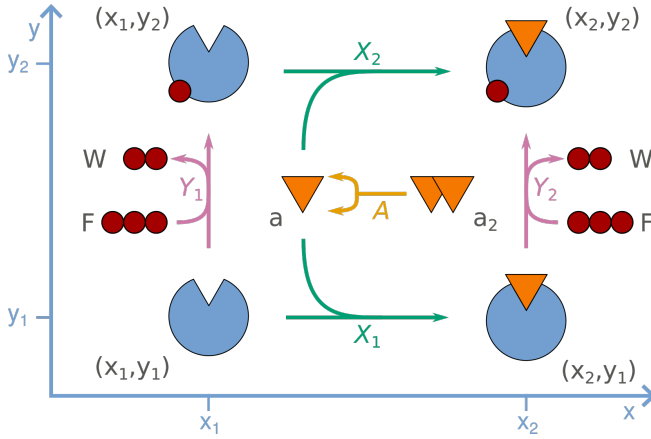


FIG. 1. **An example of bipartite network.** To illustrate our notion of bipartite structure in a CRN, we consider a generic molecule (blue “Pac-Man”) able to bind two substrates. One substrate (orange triangle, denoted a) is generated from a precursor (denoted a_2) via reaction A : $a_2 \rightarrow 2a$ (notice the use of hypergraph notation where the production of two molecules of a is represented by a bifurcating arrow). The substrate a can then bind the molecule via reactions X_1 or X_2 , which both change its x state from x_1 (free) to x_2 (bound). The other substrate (red sphere) is exchanged between the molecule and compounds F and W through reactions Y_1 and Y_2 , which both change the y state of the molecule from y_1 (free) to y_2 (bound). As a result, the molecule can be found in four double states (x, y) : (x_1, y_1) , (x_2, y_1) , (x_2, y_2) , (x_1, y_2) , which form the set of bipartite species (set b). All other species in the network (a_2 , a , F and W) can be treated either as ancillary species (set a) or, if their concentration is controlled by external reservoirs, as chemostatted species (i.e., controlled parameters). Crucially, as none of the transitions between the double states can change both x and y at the same time, the resulting CRN has a bipartite structure. To fix ideas, the Pac-Man shaped molecule can be thought of as an enzyme binding (state x) the monomeric form of a substrate and phosphorylated (state y) by adenosine triphosphate (F) to adenosine diphosphate (W) hydrolysis. A concrete example with a simpler yet similar bipartite structure is found in a recently reported single-molecule molecular motor, where the two states are the axial chirality (which can be in R state or S state) and the chemical state (which can be in diacid state or anhydride state) of the motor [73]. The latter results in a linear bipartite network to which the treatment in Ref. [40] straightforwardly applies. On the contrary, systems like the present network or the applications analyzed in Section V require a generalization of the previous framework due to the presence of nonlinear reactions and ancillary species.

of coarse-grained reactions. The reason for this will become clear in Subs. III B.

In bipartite CRNs, we can apply the same splittings of the chemical species and reactions to the stoichiometric matrix

$$\mathbb{S} = \begin{matrix} a \\ b \end{matrix} \begin{pmatrix} \mathbb{S}^a \\ \mathbb{S}^b \end{pmatrix}, \quad (5)$$

with

$$\mathbb{S}^a = \begin{matrix} A & X & Y \\ \mathbb{S}_A^a & \mathbb{S}_X^a & \mathbb{S}_Y^a \end{matrix}, \quad (6a)$$

$$\mathbb{S}^b = \begin{matrix} A & X & Y \\ 0 & \mathbb{S}_X^b & \mathbb{S}_Y^b \end{matrix}, \quad (6b)$$

to the current vector

$$\mathbf{j} = (j^A, j^X, j^Y), \quad (7)$$

and to the concentration vector

$$\mathbf{z} = (\mathbf{a}, \mathbf{b}). \quad (8)$$

The rate equation (2) thus becomes

$$d_t \mathbf{a} = \mathbb{S}^a \mathbf{j} = \mathbb{S}_A^a j^A \quad \mathbb{S}_X^a j^X \quad \mathbb{S}_Y^a j^Y, \quad (9a)$$

$$d_t \mathbf{b} = \mathbb{S}^b \mathbf{j} = \mathbb{S}_X^b j^X \quad \mathbb{S}_Y^b j^Y. \quad (9b)$$

To account for the total concentration of the bipartite species in the same state x or y , we define the following marginal concentrations

$$[x] = \sum_y [(x, y)], \quad (10a)$$

$$[y] = \sum_x [(x, y)], \quad (10b)$$

which are collected in the following vectors

$$\mathbf{x} = (\dots, [x], \dots)^\top, \quad (11a)$$

$$\mathbf{y} = (\dots, [y], \dots)^\top. \quad (11b)$$

Here, $[(x, y)]$ is the concentration of the specific bipartite species $\alpha \in b$ characterized by the double state (x, y) , i.e., $[(x, y)] = [\alpha]_{\alpha=(x,y)}$.

The bipartite structure of a CRN allows us to decompose it into two coupled *subnetworks* \mathcal{X} and \mathcal{Y} . The subnetwork \mathcal{X} represents the interconversion of the states $\{x\}$ of the bipartite species driven by the X reactions. Analogously, the subnetwork \mathcal{Y} represents the interconversion of the states $\{y\}$ driven by the Y reactions. Notice that this does not imply that the concentrations $[x]$ (resp. $[y]$) change only because of the reactions X (resp. Y) as we will see for the example in Subs. V A.

III. THERMODYNAMICS

Building on the setup of Sec. II, we introduce the thermodynamic theory of deterministic CRNs [50, 69, 74] and we specialize it to bipartite CRNs.

A. Chemical Potentials

The free energy contributions carried by the chemical species Z in ideal dilute solutions are given by the vector of chemical potentials [75]:

$$\boldsymbol{\mu}_Z = \boldsymbol{\mu}_Z^\circ - RT \ln \mathbf{z}, \quad (12)$$

where $\ln \mathbf{z} = (\dots, \ln[\alpha], \dots)^\top$ (hereafter, for every vector $\mathbf{v} = (\dots, v_i, \dots)^\top$, $\log \mathbf{v} = (\dots, \log v_i, \dots)^\top$), T is the temperature of the thermal bath, R is the gas constant, and $\boldsymbol{\mu}_Z^\circ$ is the vector of the standard chemical potentials, which in turn are given by the sum of a standard enthalpic \mathbf{h}_Z° and entropic \mathbf{s}_Z° contributions according to

$$\boldsymbol{\mu}_Z^\circ = \mathbf{h}_Z^\circ - T \mathbf{s}_Z^\circ. \quad (13)$$

From a thermodynamic standpoint, CRNs are said to be *open* when they exchange free energy with some external reservoirs besides the thermal bath at temperature T . For instance, *chemicals* can be exchanged with chemostats, i.e., particle reservoirs, and/or *photons* can be exchanged with radiation sources. Each reservoir is in turn characterized by a chemical potential.

For chemostats, their chemical potentials are given by the vector

$$\boldsymbol{\mu}_{\text{ch}} = \boldsymbol{\mu}_{\text{ch}}^\circ - RT \ln \mathbf{c}, \quad (14)$$

where $\boldsymbol{\mu}_{\text{ch}}^\circ$ and \mathbf{c} are the vectors of standard chemical potentials and of the concentrations of the exchanged chemicals, respectively.

For radiation sources, the chemical potential of photons γ_ν at frequency ν and concentration distribution n_ν reads [69]

$$\mu_\nu = u_\nu - RT \ln \left(\frac{f_\nu n_\nu}{n_\nu} \right), \quad (15)$$

where $u_\nu = N_A \hbar \omega_\nu$ is the energy carried by a mole of photons at frequency ν and $f_\nu = 2\omega_\nu^2 / N_A \pi c^3$ is the density of photon states at frequency ν (with N_A the Avogadro's number, \hbar the reduced Planck constant, c the speed of light and $\omega_\nu = 2\pi\nu$ the angular frequency). When the radiation source is thermal, photons are distributed according to the black body distribution at a certain temperature T_r :

$$n_\nu^{T_r} = \frac{f_\nu}{\exp(\hbar\omega_\nu/k_B T_r) - 1}, \quad (16)$$

with k_B the Boltzmann constant. As a consequence, their chemical potential becomes [76, 77]

$$\mu_\nu = u_\nu - RT \ln \left(\frac{f_\nu n_\nu^{T_r}}{n_\nu^{T_r}} \right) = u_\nu \left(1 - \frac{T}{T_r} \right). \quad (17)$$

Since a radiation source is in equilibrium with a CRN only when the corresponding chemical potential vanishes [69], i.e., $\mu_\nu = 0$, Eq. (17) implies that this happens for a thermal radiation source only when its temperature is the same as the one of the thermal bath, i.e., $T_r = T$.

In this work, we focus for simplicity on autonomous CRNs, namely, the chemical potentials of the reservoirs ($\boldsymbol{\mu}_{\text{ch}}$ in Eq. (14) for chemostats and μ_ν in Eq. (15) for radiation sources), and thus also the quantities \mathbf{c} and n_ν , are constant in time.

B. Thermodynamic Forces

Chemical reactions are driven by thermodynamic forces named affinities [78, 79],

$$\mathcal{A}_\rho = -\frac{\boldsymbol{\mu}_Z \cdot \mathbb{S}_\rho}{T} - F_\rho, \quad (18)$$

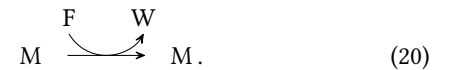
which have two contributions. The first contribution, i.e., $-\boldsymbol{\mu}_Z \cdot \mathbb{S}_\rho / T$, accounts for the variation of the free energy due to the interconversion of the Z species in solution via reaction ρ . The second, i.e., F_ρ , accounts for the *external force* due to the coupling with the reservoirs.

For elementary reactions, affinities (18) satisfy the local detailed balance condition,

$$\mathcal{A}_\rho = R \ln \frac{j^\rho}{j^{-\rho}}, \quad (19)$$

which implies that they always have the same sign of the corresponding reaction currents: $\mathcal{A}_\rho j^\rho \geq 0$. Coarse-grained reactions do not in general satisfy Eq. (19), and hence their affinities may not be aligned to the currents (unless they are tightly coupled [67]). However, every *independent* subset of coarse-grained reactions $C \subseteq \mathcal{R}$, namely, every subset whose underlying elementary reactions involve unique intermediate (coarse-grained) species which are not shared with other subsets, satisfies $\sum_{\rho \in C} \mathcal{A}_\rho j^\rho \geq 0$ [67, 68].

Example 1. Consider the following reaction (assumed as elementary for illustrative reasons) where the catalyst M interconverts the substrate (fuel) F into the product (waste) W which are exchanged with chemostats:



In this case, the affinity coincides with the external force,

$$\mathcal{A} = F = \frac{\mu_F - \mu_W}{T}, \quad (21)$$

while the reaction fluxes satisfy mass action kinetics,

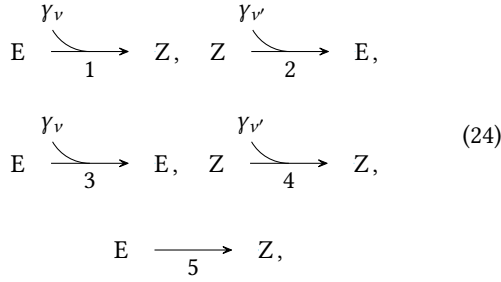
$$j = k [F][M], \quad j^- = k_- [W][M], \quad (22)$$

where $[F]$, $[W]$ and $[M]$ are the concentrations of the fuel, waste and catalyst, respectively. Since reaction (20) is elementary, it satisfies the local detailed balance condition (19), which in turn implies that the kinetic constants $\{k, k_-\}$ and the standard chemical potentials $\{\mu_F^\circ, \mu_W^\circ\}$ are not independent:

$$\frac{\mu_F^\circ - \mu_W^\circ}{T} = R \ln \frac{k}{k_-}, \quad (23)$$

where we used Eq. (14) to express the chemical potential of the chemostats.

Example 2. Consider the coarse-grained reactions



describing a typical photoisomerization process occurring via the so-called diabatic mechanism: upon the absorption of a photon γ_ν , a species in the E conformation is either converted into the Z conformation (reaction 1) or it decays back to the same conformation by dissipating all the absorbed energy (reaction 3); analogously, another photon $\gamma_{\nu'}$, with possibly a different frequency ν' , can trigger the conversion of Z into E (reaction 2) or be dissipated (reaction 4); finally, E and Z can also thermally interconvert (reaction 5). We assume for simplicity that only photons at the exact frequencies ν and ν' can trigger the corresponding transitions. In such a case, the affinities read

$$\begin{aligned}
 \mathcal{A}_1 &= \frac{\mu_E - \mu_Z}{T} - \frac{\mu_\nu}{T}, & \mathcal{A}_2 &= \frac{\mu_Z - \mu_E}{T} - \frac{\mu_{\nu'}}{T}, \\
 \mathcal{A}_3 &= \frac{\mu_\nu}{T}, & \mathcal{A}_4 &= \frac{\mu_{\nu'}}{T}, & \mathcal{A}_5 &= \frac{\mu_E - \mu_Z}{T},
 \end{aligned} \quad (25)$$

with $F_1 = F_3 = \mu_\nu/T$, $F_2 = F_4 = \mu_{\nu'}/T$, and $F_5 = 0$, while expressions for the reaction currents $\{j^1, j^2, j^3, j^4, j^5\}$ can be found in Ref. [69]. The reactions (24) are not tightly coupled and, consequently, affinities and currents are not in general aligned. However, as they constitute an independent subset of coarse-grained reactions, they satisfy

$$\mathcal{A}_1 j^1 - \mathcal{A}_2 j^2 - \mathcal{A}_3 j^3 + \mathcal{A}_4 j^4 - \mathcal{A}_5 j^5 \geq 0. \quad (26)$$

C. Entropy

The thermodynamic entropy per unit volume of open CRNs,

$$S = S_Z - S_{\text{ch}} - S_{\text{ph}}, \quad (27)$$

is given in general by the contribution of the chemical species S_Z , the chemostats S_{ch} , and the photons S_{ph} . The contribution carried by the chemical species reads [74]

$$S_Z = (s_Z^\circ - R \ln z) \cdot z - R \|z\| = s_Z^\circ \cdot z - S_Z^{\text{H}}, \quad (28)$$

where $\|z\| = \sum_\alpha [\alpha]$ and we introduced a Shannon-like entropy for macroscopic non-normalized concentration distributions, $S_Z^{\text{H}} := -R \ln z \cdot z - R \|z\|$. The latter will play a crucial role in recovering the information thermodynamic framework for deterministic CRNs in Sec. IV.

Similarly, the contribution of the chemostats is given by

$$S_{\text{ch}} = (s_{\text{ch}}^\circ - R \ln c) \cdot c - R \|c\|, \quad (29)$$

with $s_{\text{ch}}^\circ = -\partial \mu_{\text{ch}}^\circ / \partial T$ the vector of the standard molar entropies of the chemostats. On the other hand, the contribution due to the photons is expressed as the entropy of an ideal Bose gas [79]

$$S_{\text{ph}} = R \int \left[(f_\nu - n_\nu) \ln (f_\nu - n_\nu) - n_\nu \ln n_\nu - f_\nu \ln f_\nu \right] d\nu, \quad (30)$$

where the integral runs in general over the whole spectrum.

The time derivative of the total entropy (27), i.e., $d_t S = s_Z \cdot \mathbb{S} j$ (with $s_Z = s_Z^\circ - R \ln z$), is given by the sum of the entropy production rate $\dot{\Sigma} \geq 0$, accounting for the dissipation, and of the entropy flow \dot{S}_e , accounting for the reversible exchange of entropy with the reservoirs. It provides the nonequilibrium formulation of the second law, which can be written as

$$\dot{\Sigma} = d_t S - \dot{S}_e \geq 0 \quad (31)$$

where, in autonomous CRNs (i.e., with c and n_ν being constant in time),

$$\dot{\Sigma} = \mathcal{A} \cdot j \geq 0, \quad (32)$$

$$\dot{S}_e = \frac{\mathbf{h}_Z^\circ}{T} \cdot \mathbb{S} j - F \cdot j, \quad (33)$$

with $\mathcal{A} = (\dots, \mathcal{A}_\rho, \dots)^\top$ and $F = (\dots, F_\rho, \dots)^\top$.

In bipartite CRNs, we can apply the same splittings of the chemical species $Z = a \cup b$ to the total entropy carried by the chemical species S_Z (given in Eq. (28)) which leads to

$$S = S_a - S_b - S_{\text{ch}} - S_{\text{ph}}, \quad (34)$$

where, given a set of species $v = \{a, b\}$, $S_v = (s_v^\circ - R \ln v) \cdot v - R \|v\|$ with $v = (\dots, [\alpha], \dots)_{\alpha \in v}$ and $s_v^\circ = (\dots, s_{\alpha}^\circ, \dots)_{\alpha \in v}$. Furthermore, we can apply the splitting of the chemical reactions $\mathcal{R} = A \cup X \cup Y$ to the entropy production rate and the entropy flow:

$$\dot{\Sigma} = \dot{\Sigma}^A - \dot{\Sigma}^X - \dot{\Sigma}^Y, \quad (35)$$

$$\dot{S}_e = \dot{S}_e^A - \dot{S}_e^X - \dot{S}_e^Y, \quad (36)$$

where

$$\dot{\Sigma}^i = \sum_{\rho \in i} \mathcal{A}_\rho j^\rho \geq 0, \quad (37)$$

$$\dot{S}_e^i = \sum_{\rho \in i} \left(\frac{\mathbf{h}_Z^\circ}{T} \cdot \mathbb{S}_\rho - F_\rho \right) j^\rho, \quad (38)$$

with $i = \{A, X, Y\}$. Note that the inequality in Eq. (37) holds because the reactions in the sets $\{A, X, Y\}$ either satisfy the local detailed balance condition (19), or are independent sets of coarse-grained reactions (see Subs. II).

IV. INFORMATION THERMODYNAMICS

We now formulate a second law for each subnetwork (defined in Sec. II) which, compared to the second law (31) of the whole CRN, is (mainly) modified by an information flow term accounting for the information exchange between the two subnetworks. This information flow arises due to the fact that the two subnetworks share some chemical species which are involved in both X and Y reactions.

A. Mutual Information

We start by introducing the probability that bipartite species $\alpha \in b$ are in a specific double state (x, y) as

$$p_{x,y} = \frac{[(x, y)]}{\|\mathbf{b}\|}, \quad (39)$$

with $\|\mathbf{b}\| = \sum_{\alpha \in b} [\alpha]$ and the marginal probabilities

$$p_x = \sum_y p_{x,y} = \frac{[x]}{\|\mathbf{b}\|}, \quad (40a)$$

$$p_y = \sum_x p_{x,y} = \frac{[y]}{\|\mathbf{b}\|}, \quad (40b)$$

that bipartite species are in state x (resp. y) irrespectively of their state y (resp. x). In analogy to information theory when dealing with the joint probability of a pair of random variables [41], this allows us to introduce the following Shannon entropies for the bipartite species

$$H_{x,y} = - \sum_{x,y} p_{x,y} \ln p_{x,y}, \quad (41a)$$

$$H_x = - \sum_x p_x \ln p_x, \quad (41b)$$

$$H_y = - \sum_y p_y \ln p_y, \quad (41c)$$

and the *mutual information*

$$\mathcal{I}_b = \sum_{x,y} p_{x,y} \ln \frac{p_{x,y}}{p_x p_y}, \quad (42)$$

which by construction satisfies

$$H_{x,y} = H_x + H_y - \mathcal{I}_b. \quad (43)$$

Crucially, by applying the definition of probabilities in Eqs. (39) and (40), the mutual information (42) can be directly expressed in terms of the macroscopic non-normalized concentration distributions

$$\mathcal{I}_b = \sum_{x,y} \frac{[(x, y)]}{\|\mathbf{b}\|} \ln \frac{[(x, y)] \|\mathbf{b}\|}{[x][y]}, \quad (44)$$

which still measures, as in information theory [33, 41], the reduction in uncertainty about the states $\{x\}$ when the states $\{y\}$ are known, and vice versa. This physically means that \mathcal{I}_b quantifies, in the framework of this paper, to which extent being in a specific state x correlates with being in a specific state y for the bipartite species and it can be determined by measuring the concentrations $[(x, y)]$.

Example. Consider the network in Fig. 1 with arbitrary concentrations $[(x, y)]$. On the one hand, if $[(x_1, y_1)]/[(x_2, y_1)] = [(x_1, y_2)]/[(x_2, y_2)]$, the y state is not correlated with the x state, as the relative concentration between species in states x_1 and x_2 is independent of state y . Analogously, if $[(x_1, y_1)]/[(x_1, y_2)] = [(x_2, y_1)]/[(x_2, y_2)]$,

the x state is not correlated with the y state, as the relative concentration between species in states y_1 and y_2 is independent of state x . In such a case, one cannot gain information about one state (e.g., the concentration of molecules binding the triangular substrate) by measuring the other (e.g., the concentration of molecules binding the small circular substrate). This reflects into a vanishing mutual information (44). On the other hand, if $[(x_1, y_1)]/[(x_2, y_1)] \neq [(x_1, y_2)]/[(x_2, y_2)]$, the y state correlates with the x state, as the relative concentration between species in states x_1 and x_2 now depends state y . For analogous reasons, if $[(x_1, y_1)]/[(x_1, y_2)] \neq [(x_2, y_1)]/[(x_2, y_2)]$, the x state correlates with the y state. In such a case, one can thus gain information on one state by measuring the other. For instance, in the presence of positive correlations between x_1 and y_1 and between x_2 and y_2 (namely, $[(x_1, y_1)] > [(x_1, y_2)]$ and $[(x_2, y_2)] > [(x_2, y_1)]$), knowing that the marginal concentration $[y]$ is shifted towards state y_2 (most of the molecules bind the small circular substrate) informs about the marginal concentration $[x]$ being shifted towards state x_2 (most of the molecules bind the triangular substrate). This reflects into a non-null mutual information (44).

B. Entropy Decomposition

In standard information thermodynamics, the thermodynamic entropy of a system coincides with the Shannon entropy. Thus, Eq. (43) splits the former into contributions from subsystems and their mutual information. On the other hand, as shown in Subs. III C, the entropy (28) of deterministic CRNs is not given by the Shannon entropy. Nevertheless, the bipartite structure of the network allows us to split the Shannon-like entropy of the bipartite species $S_b^H = -R(\ln \mathbf{b}) \cdot \mathbf{b} - R\|\mathbf{b}\|$ by following the same logic employed for the standard derivation of Eq. (43) [41], but exploiting the expression in Eq. (44) for the mutual information. In this way, we get

$$S_b^H = S_x^H + S_y^H - R\|\mathbf{b}\|\mathcal{I}_b - R\|\mathbf{b}\|(\ln\|\mathbf{b}\| - 1), \quad (45)$$

where

$$S_x^H = -R(\ln \mathbf{x}) \cdot \mathbf{x} - R\|\mathbf{x}\|, \quad (46a)$$

$$S_y^H = -R(\ln \mathbf{y}) \cdot \mathbf{y} - R\|\mathbf{y}\|, \quad (46b)$$

are the Shannon-like entropy of states $\{x\}$ and $\{y\}$, respectively, and $\|\mathbf{x}\| = \|\mathbf{y}\| = \|\mathbf{b}\|$. Note that the last term in Eq. (45), i.e., $\|\mathbf{b}\|(\ln\|\mathbf{b}\| - 1)$, emerges because of the non-normalized concentration distributions.

By using Eq. (45) to express the entropy of the bipartite species $S_b = S_b^H - s_b^\circ \cdot \mathbf{b}$, the total entropy (27) becomes

$$S = S_x^H + S_y^H - R\|\mathbf{b}\|\mathcal{I}_b - R\|\mathbf{b}\|(\ln\|\mathbf{b}\| - 1) - s_b^\circ \cdot \mathbf{b} + S_a + S_{\text{ch}} + S_{\text{ph}}, \quad (47)$$

where $s_b^\circ \cdot \mathbf{b}$ accounts for the standard molar entropy of the bipartite species and cannot be split into a contribution due

to the species in state x and a contribution due to the species in state y . This term is absent in the decomposition of entropy of bipartite Markov jump processes [31] as long as the states do not have an internal entropy.

C. Entropy Production Decomposition

We now decompose the entropy production rate in such a way as to formulate the second law for each subnetwork \mathcal{X} and \mathcal{Y} . We start by rewriting the second law (31) specifying the entropy flow according to Eq. (36):

$$\dot{\Sigma} = d_t S - \dot{S}_e^A - \dot{S}_e^X - \dot{S}_e^Y. \quad (48)$$

From Eq. (47), we further derive

$$\begin{aligned} d_t S = & d_t S_x^H - d_t S_y^H - \mathbf{s}_b^\circ \cdot d_t \mathbf{b} - R d_t (\|\mathbf{b}\| \mathcal{I}_b) \\ & R \ln \|\mathbf{b}\| d_t \|\mathbf{b}\| - \mathbf{s}_a \cdot d_t \mathbf{a}, \end{aligned} \quad (49)$$

where $\mathbf{s}_a = \mathbf{s}_a^\circ - R \ln \mathbf{a}$ and the time derivative of the information term (44) can be specified as

$$d_t \mathcal{I}_b = \dot{\mathcal{I}}_x - \dot{\mathcal{I}}_y - (\ln \|\mathbf{b}\| - \mathcal{I}_b) d_t \ln \|\mathbf{b}\| \quad (50)$$

with

$$\dot{\mathcal{I}}_x := \frac{1}{\|\mathbf{b}\|} \left\{ (\ln \mathbf{b}) \cdot \mathbb{S}_X^b \mathbf{j}^X - (\ln \mathbf{x}) \cdot d_t \mathbf{x} \right\}, \quad (51a)$$

$$\dot{\mathcal{I}}_y := \frac{1}{\|\mathbf{b}\|} \left\{ (\ln \mathbf{b}) \cdot \mathbb{S}_Y^b \mathbf{j}^Y - (\ln \mathbf{y}) \cdot d_t \mathbf{y} \right\}. \quad (51b)$$

Crucially, the information flows $\dot{\mathcal{I}}_x$ and $\dot{\mathcal{I}}_y$ quantify the changes in the mutual information (44) per units of concentration, namely, the changes in the relative uncertainty (correlations) between the $\{x\}$ and the $\{y\}$ states, due to the dynamics of subnetworks \mathcal{X} and \mathcal{Y} , respectively. The third term in Eq. (50) accounts for the non-normalized nature of the concentration distributions, vanishes at steady state, and does not contribute to the entropy production decomposition. Indeed, by using Eq. (49) and Eq. (50) into Eq. (48) and by splitting $d_t \mathbf{a}$ and $d_t \mathbf{b}$ into the contributions due to the different reactions $\mathcal{R} = A \cup X \cup Y$ as in Eqs. (9a) and (9b), we obtain

$$\begin{aligned} \dot{\Sigma}(t) = & d_t S_x^H - \dot{S}_e^X - R \|\mathbf{b}\| \dot{\mathcal{I}}_x - \mathbf{s}_b^\circ \cdot \mathbb{S}_X^b \mathbf{j}^X - \mathbf{s}_a \cdot \mathbb{S}_X^a \mathbf{j}^X \\ & d_t S_y^H - \dot{S}_e^Y - R \|\mathbf{b}\| \dot{\mathcal{I}}_y - \mathbf{s}_b^\circ \cdot \mathbb{S}_Y^b \mathbf{j}^Y - \mathbf{s}_a \cdot \mathbb{S}_Y^a \mathbf{j}^Y \\ & - \dot{S}_e^A - \mathbf{s}_a \cdot \mathbb{S}_A^a \mathbf{j}^A \end{aligned} \quad (52)$$

which, by using Eq. (37), allows us to identify

$$\dot{\Sigma}^X = d_t S_x^H - \dot{S}_e^X - R \|\mathbf{b}\| \dot{\mathcal{I}}_x - \mathbf{s}_b^\circ \cdot \mathbb{S}_X^b \mathbf{j}^X - \mathbf{s}_a \cdot \mathbb{S}_X^a \mathbf{j}^X \geq 0, \quad (53a)$$

$$\dot{\Sigma}^Y = d_t S_y^H - \dot{S}_e^Y - R \|\mathbf{b}\| \dot{\mathcal{I}}_y - \mathbf{s}_b^\circ \cdot \mathbb{S}_Y^b \mathbf{j}^Y - \mathbf{s}_a \cdot \mathbb{S}_Y^a \mathbf{j}^Y \geq 0, \quad (53b)$$

$$\dot{\Sigma}^A = -\dot{S}_e^A - \mathbf{s}_a \cdot \mathbb{S}_A^a \mathbf{j}^A \geq 0. \quad (53c)$$

Equation (53c) has the same form as the second law (31): it explicitly expresses the dissipation of the A reactions (which do not involve the bipartite species) in terms of the entropy

flow with the corresponding reservoirs \dot{S}_e^A and the variation of the total entropy due to the interconversion of the a species $\mathbf{s}_a \cdot \mathbb{S}_A^a \mathbf{j}^A$.

Equations (53a) and (53b) are the key results of our work. They provide a formulation of the second law for the subnetworks \mathcal{X} and \mathcal{Y} , respectively. Let us focus in the following on the subnetwork \mathcal{X} , but analogous comments also hold for the subnetwork \mathcal{Y} . Equation (53a) shows that the dissipation $\dot{\Sigma}^X$ is balanced by different mechanisms: i) the variation of the subnetwork Shannon-like entropy, $d_t S_x^H$; ii) the entropy flow with the reservoirs coupled to the X reactions, \dot{S}_e^X ; iii) the information flow accounting for the information exchange with the \mathcal{Y} subnetwork, $\dot{\mathcal{I}}_x$; iv) the variation to the subnetwork entropy due to the standard molar entropy carried by the chemical species, $\mathbf{s}_b^\circ \cdot \mathbb{S}_X^b \mathbf{j}^X$; v) the variation to the subnetwork entropy due to the consumption/production of the a species, $\mathbf{s}_a \cdot \mathbb{S}_X^a \mathbf{j}^X$. Note that the first three mechanisms in Eq. (53a) also appear in an analogous decomposition of entropy production for bipartite Markov jump processes [31], while the last two only emerge in bipartite deterministic CRNs.

Each mechanism can act as a *source* (of entropy), when it contributes positively to the entropy production, or as an *output* (of entropy), when it contributes negatively. The sources must always dominate the outputs to maintain the subnetwork out of equilibrium because of the unavoidable dissipation. For instance, when $\dot{\mathcal{I}}_x < 0$, the subnetwork \mathcal{X} is decreasing correlation with the subnetwork \mathcal{Y} : the mutual information (44) decreases (see Eq. (50)). This consumption of mutual information corresponds to a source of entropy ($-R \|\mathbf{b}\| \dot{\mathcal{I}}_x > 0$) that can be used to sustain the output mechanisms and/or to balance the dissipation. On the other hand, when $\dot{\mathcal{I}}_x > 0$ the subnetwork is increasing correlation with the subnetwork \mathcal{Y} . This corresponds to an output of entropy ($-R \|\mathbf{b}\| \dot{\mathcal{I}}_x < 0$) which has to be sustained by an entropy source to balance the dissipation ($\dot{\Sigma}^X \geq 0$).

V. APPLICATIONS

We now apply our framework to two paradigmatic examples of chemical engines with nonlinear dynamics: a model for chemically-driven self-assembly [13, 42] and a light-driven bimolecular motor [24, 27, 43, 44]. We characterize their functioning at steady state for various operating regimes.

A. Driven self-assembly

As a first application, we analyze a minimalist model (see Fig. 2a) epitomizing the basic working principles of chemically-driven self-assembly [13, 42, 80, 81]: an external force is exploited to increase the concentration of a target species with respect to the equilibrium one. Prominent examples of this kind of mechanism are the formation of microtubules out of tubulin dimers fueled by guanosine 5'-triphosphate (GTP) [82, 83] and the ATP-driven self-assembly

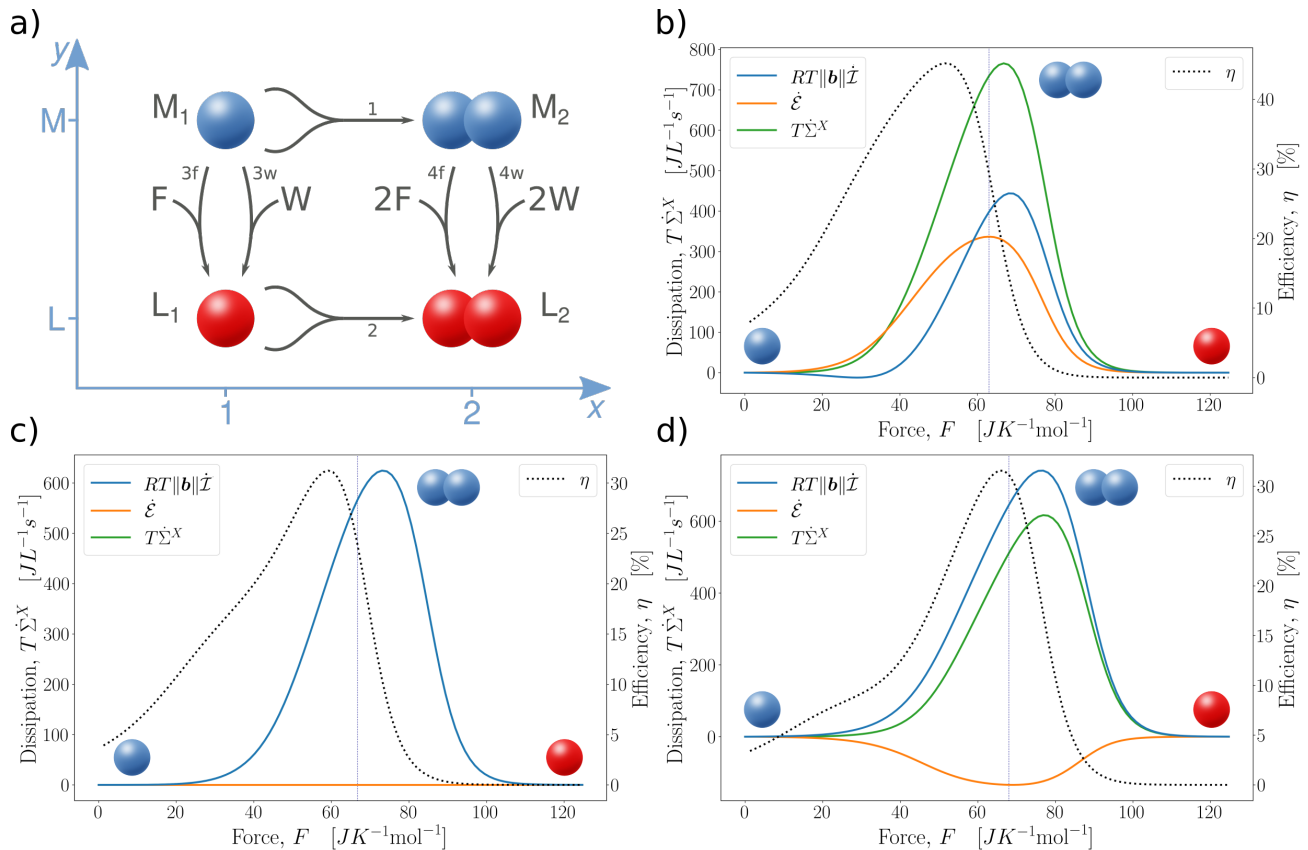


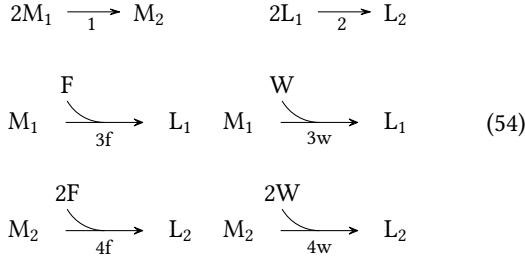
FIG. 2. **Chemically-driven self-assembly of monomers into dimers.** a) Hypergraph representation of the bipartite chemical reaction network. Arrows are used to indicate the conventional direction of the reactions, which are all reversible. Reactions are labelled according to the CRN (54); b-d) Numerical simulations of the model with three different sets of parameters. Information flow ($RT\|\mathbf{b}\|\dot{\mathcal{I}}$), energy flow ($\dot{\mathcal{E}}$), and dissipation of the self-assembly subnetwork ($T\dot{\Sigma}^X = \dot{\mathcal{E}} - RT\|\mathbf{b}\|\dot{\mathcal{I}}$) at steady state are plotted against the net thermodynamic force acting on the system ($F = (\mu_F - \mu_W)/T$, with $T = 298K$, $\mu_F^\circ = 11$ kJ/mol and $\mu_W^\circ = -11$ kJ/mol, up to an arbitrary constant) in a range corresponding to $[F]$ from $1 \cdot 10^{-4}M$ to $4 \cdot 10^2M$, with $[W] = 1M$. The efficiency of the internal free-energy transduction (η) is computed according to Eq. (64). The standard chemical potentials ruling the relative thermodynamic stability of the four species are, up to an arbitrary constant: $\mu_{M_1}^\circ = -2$ kJ/mol, $\mu_{L_1}^\circ = -3$ kJ/mol, $\mu_{L_2}^\circ = -4$ kJ/mol, and $\mu_{M_2}^\circ = 9$ kJ/mol for case b); $\mu_{M_1}^\circ = 2$ kJ/mol, $\mu_{L_1}^\circ = -3$ kJ/mol, $\mu_{L_2}^\circ = -6$ kJ/mol, and $\mu_{M_2}^\circ = 4$ kJ/mol for case c); and $\mu_{M_1}^\circ = 2$ kJ/mol, $\mu_{L_1}^\circ = -3$ kJ/mol, $\mu_{L_2}^\circ = -2$ kJ/mol, and $\mu_{M_2}^\circ = 4$ kJ/mol for case d). The independent kinetic parameters are the same as in Ref. [80]. Colored spheres indicate which is the most populated species at steady state for low, intermediate and high forces regimes, while dotted vertical lines mark the value of the force maximizing the concentration of the target species M_2 . Note that, in panel c), the information flow ($RT\|\mathbf{b}\|\dot{\mathcal{I}}$, blue solid line) coincides exactly with the dissipation of the self-assembly subnetwork ($T\dot{\Sigma}^X$, green solid line).

of actin filaments [84]. Driven self-assembly has also been exploited in experiments such as the controlled gelation of dibenzoyl-L-cysteine to form nanofibers [15] and the chemically fueled transient self-assembly of fibrous hydrogel materials [85].

In the model depicted in Fig. 2a, the direct aggregation of two monomers M_1 to form the dimer M_2 is coupled with the exergonic conversion of a high energetic species F into a low energetic one W . In particular, both the monomer M_1 and the dimer M_2 can catalyze the F -to- W conversion via their activated species L_1 and L_2 (e.g., $F + M_1 \rightarrow L_1 \rightarrow M_1 + W$). This leads, by properly fixing the concentrations of F and W , to a nonequilibrium steady state enriched in the dimer M_2 . Unlike conventional equilibrium self-assembly, the efficacy

of this synthetic procedure is not determined by the relative thermodynamic stability of the species, but rather by a kinetic asymmetry [86] in how the monomers M_1 and the dimer M_2 react with F and W [13, 42]. This built-in kinetic asymmetry is often referred to as an information ratchet mechanism because the rate at which F and W react with the system is somehow dependent on information about the reacting species: that is, a species will react more quickly with fuel if doing so enables forward cycling or prevents backward cycling [42, 46].

The model depicted in Fig. 2a corresponds to the open CRN



where there are four bipartite species $b = \{M_1, M_2, L_1, L_2\}$ (with monomeric/dimeric states $x \in \{1, 2\}$ and non-activated/activated states $y \in \{M, L\}$) and two chemostats F and W. No a species are present. The marginal concentrations read

$$[M] = [M_1] \quad [M_2], \quad (55a)$$

$$[L] = [L_1] \quad [L_2], \quad (55b)$$

$$[1] = [M_1] \quad [L_1], \quad (55c)$$

$$[2] = [M_2] \quad [L_2], \quad (55d)$$

Here, all the reactions satisfy mass action kinetics. The self-assembly reactions $\{1, 2\}$ are nonlinear with respect to the chemical species M_1 and L_1 , and only change their x state, i.e., $X = \{1, 2\}$. They thus define the self-assembly subnetwork \mathcal{X} . Conversely, the fueling and waste-forming reactions $\{3f, 3w, 4f, 4w\}$, accounting for the coupling with the chemostats, only change the y state of the species, i.e., $Y = \{3f, 3w, 4f, 4w\}$. They thus define the fueling subnetwork \mathcal{Y} . We stress that here, due to the nonlinearity of the X reactions, $\|b\| = [M] \quad [L] = [1] \quad [2] = [M_1] \quad [M_2] \quad [L_1] \quad [L_2]$ is not a conserved quantity of the dynamics, which prevents direct application of previous formulations of information thermodynamics for bipartite networks [31, 40]. In addition, the CRN in Eq. 54 provides an example where the marginal concentrations of y states ($[M]$ and $[L]$) change due to X reactions ($\{1, 2\}$).

We now characterize the energetics of the nonequilibrium steady state which will eventually be reached by the dynamics in the long time limit. We do so by computing the various contributions to the total dissipation (52), thus specializing Eqs. (53a) and (53b) for the case at study (notice that Eq. (53c) as well as all the terms involving a species do not play any role here). At steady state, the time derivatives of Shannon-like entropies of states $\{x\}$ and $\{y\}$ vanish ($d_t S_x^H = d_t S_y^H = 0$) and the dynamics is fully characterized by the net current $j = -j^1 = j^2 = (j^{3f} - j^{3w})/2 = -(j^{4f} - j^{4w})$, which is positive when flowing anticlockwise across the hyper-graph in Fig. 2a according to the sign convention in Eq. (54). Furthermore, the rate at which the fuel species F is injected into the system at steady state to keep its concentration constant ($I_F = j^{3f} - 2j^{4f}$) corresponds to the rate at which the waste species W is extracted ($I_W = -I_W = -j^{3w} - 2j^{4w}$). As a consequence, Eqs. (53a) and (53b) boil down to:

$$T\dot{\Sigma}^X = \underbrace{j(\mu_{M_2}^\circ - \mu_{L_2}^\circ - 2\mu_{L_1}^\circ + 2\mu_{M_1}^\circ)}_{=:\dot{\mathcal{E}}} \quad \underbrace{RT j \ln \frac{[L_1]^2 [M_2]}{[M_1]^2 [L_2]}}_{=:RT\|b\|\dot{J}}, \quad (56a)$$

$$T\dot{\Sigma}^Y = I_F(\mu_F - \mu_W) - \dot{\mathcal{E}} - RT\|b\|\dot{J}, \quad (56b)$$

where $\dot{J} = -\dot{J}_x = \dot{J}_y$ follows directly from Eqs. (51a), (51b) and (55), and we introduced the energy flow $\dot{\mathcal{E}}/T = -\dot{\Sigma}_e^X \quad s_b^\circ \cdot \mathbb{S}_X^b J^X$ accounting for the standard contribution to the variation of the CRN free energy due to X reactions, with $I_F(\mu_F - \mu_W) - \dot{\mathcal{E}} = -\dot{\Sigma}_e^Y \quad s_b^\circ \cdot \mathbb{S}_Y^b J^Y$ following directly from Eq. (38) (after identifying $F_{3f} = \mu_F/T$, $F_{3w} = \mu_W/T$, $F_{4f} = 2\mu_F/T$, and $F_{4w} = 2\mu_W/T$).

Equations (56a) and (56b) fully characterize the free-energy exchanges between the CRN and the chemostats, on the one hand, and between the subnetworks \mathcal{X} and \mathcal{Y} , on the other hand. The former is accounted for by $I_F(\mu_F - \mu_W)$, quantifying the fueling work performed by the external force $F = (\mu_F - \mu_W)/T$ [80]. This work is entirely delivered to the subnetwork \mathcal{Y} since $I_F(\mu_F - \mu_W)$ appears only in Eq. (56b). The free-energy exchange between \mathcal{Y} and \mathcal{X} is accounted for by the energy and information flows, which together ($\dot{\mathcal{E}} \quad RT\|b\|\dot{J}$) constitute the only possible source of free energy for the subnetwork \mathcal{X} (see Eq. (56a)). This implies that the subnetwork \mathcal{X} can be kept out of equilibrium only when part of the work $I_F(\mu_F - \mu_W)$ is not dissipated by the subnetwork \mathcal{Y} , but output to the subnetwork \mathcal{X} via $\dot{\mathcal{E}}$ and \dot{J} . We view this process as an internal free-energy transduction. Based on this understanding, one can define the efficiency of the internal free-energy transduction as the ratio between the free energy transferred to the subnetwork \mathcal{X} and the work performed on the subnetwork \mathcal{Y}

$$0 \leq \eta = \frac{RT\|b\|\dot{J} \quad \dot{\mathcal{E}}}{I_F(\mu_F - \mu_W)} = 1 - \frac{T\dot{\Sigma}^Y}{I_F(\mu_F - \mu_W)} \leq 1, \quad (57)$$

which is bounded by zero and one due to the non-negativity of $\dot{\Sigma}^Y$. Physically, η measures the fraction of the fueling work devoted to keep the subnetwork \mathcal{X} out of equilibrium, which is exactly the goal of driven self-assembly.

Such a level of resolution on how free energy is dissipated leads to a refinement of a previous analysis of this model [80]. First, Eq. (57) characterizes the steady state performance of the self-assembly (what is called ‘‘maintenance phase’’ in Ref. [80]). Second, if the subnetwork \mathcal{X} performed work $\mathcal{W}_{ext} < 0$ against the environment (as in the driven synthesis setup of Ref. [80]), Eq. (57) would allow us to split the overall efficiency $\eta_{ds} = -\mathcal{W}_{ext}/I_F(\mu_F - \mu_W)$ into two contributions:

$$\eta_{ds} = \frac{RT\|b\|\dot{J} \quad \dot{\mathcal{E}}}{I_F(\mu_F - \mu_W)} \cdot \frac{-\mathcal{W}_{ext}}{\dot{\mathcal{E}} \quad RT\|b\|\dot{J}} = \eta \cdot \eta_{ext} \leq 1, \quad (58)$$

where η_{ext} measures the fraction of the free-energy transduced towards the subnetwork \mathcal{X} ($\dot{\mathcal{E}} \quad RT\|b\|\dot{J}$) which is fruitfully converted into useful power delivered to the environment ($-\mathcal{W}_{ext}$).

As a further comment, we note that η can vanish because of two reasons: a thermodynamic and a kinetic one. The former occurs when the external force is null ($F = (\mu_F - \mu_W)/T = 0$), which directly implies that both subnetworks reach equilibrium ($\dot{\Sigma}^Y = \dot{\Sigma}^X = 0$). The latter occurs when the external force is not null ($F = (\mu_F - \mu_W)/T \neq 0$), but the sum of $\dot{\mathcal{E}}$ and $RT\|b\|\dot{J}$ vanishes, thus implying $\dot{\Sigma}^X = 0$, but $\dot{\Sigma}^Y > 0$. From a kinetic standpoint, this second case is often referred as a scenario where the system does not display

kinetic asymmetry [13, 42, 46]. From an information thermodynamic standpoint, this happens when there is no internal free-energy transduction between the two subnetworks and, consequently, the subnetwork \mathcal{X} remains at equilibrium even though \mathcal{Y} is maintained in a nonequilibrium steady state by the chemostats. This also implies that all the dissipation happens at the level of the subnetwork \mathcal{Y} ($\dot{\Sigma} = \dot{\Sigma}^{\mathcal{Y}}$).

To better illustrate our results, we examine three different scenarios. The first one (Fig. 2b) is the same as in Refs. [80] and [81], where the dimeric state is energetically highly unfavored in the non-activated state ($\mu_{M_2}^\circ - 2\mu_{M_1}^\circ \gg 0$), but favored in the activated state ($\mu_{L_2}^\circ - 2\mu_{L_1}^\circ < 0$). Under these conditions, the energy flow is positive for any value of the external force and has a maximum at $F = 63 \text{ J K}^{-1} \text{ mol}^{-1}$ (which also corresponds to the value of the force maximizing the concentration of M_2). This happens because $\dot{\mathcal{E}}$ given in Eq. (56) is proportional to i) the steady state current $j > 0$ and ii) $(\mu_{M_2}^\circ - \mu_{L_2}^\circ - 2\mu_{L_1}^\circ + 2\mu_{M_1}^\circ) > 0$. Thus, the energy flow always contributes positively to the internal free-energy transduction from the subnetwork \mathcal{Y} to the subnetwork \mathcal{X} . In contrast, the information flow $\dot{\mathcal{I}}$ is negative for small values of the external force ($F \lesssim 35 \text{ J K}^{-1} \text{ mol}^{-1}$), thus representing a cost for the free-energy transduction. Physically, this means that the subnetwork \mathcal{X} is spending a fraction of the free energy provided by the energy flow to create correlations with the subnetwork \mathcal{Y} . For intermediate values of the external force ($35 \text{ J K}^{-1} \text{ mol}^{-1} \lesssim F \lesssim 100 \text{ J K}^{-1} \text{ mol}^{-1}$), the information flow has a finite positive value meaning that correlations are generated by the subnetwork \mathcal{Y} and used as a source by the subnetwork \mathcal{X} to stay out of equilibrium.

In the second scenario (Fig. 2c), there is no energetic preference between monomeric and dimeric states: $\mu_{M_2}^\circ - 2\mu_{M_1}^\circ = 0$ and $2\mu_{L_2}^\circ - \mu_{L_1}^\circ = 0$. As a consequence, the energy flow vanishes at any value of the force, and the only source of free energy for the subnetwork \mathcal{X} is the (always positive) information flow. In such a case, the subnetwork \mathcal{Y} acts as a Maxwell Demon powered by the chemostats and transducing free-energy towards the subnetwork \mathcal{X} purely in form of information. The value of the force F maximizing the concentration of the target species M_2 turns out to be $67 \text{ J K}^{-1} \text{ mol}^{-1}$.

In the third scenario (Fig. 2d), there still is no energetic preference between the non-activated monomer and non-activated dimer ($\mu_{M_2}^\circ - 2\mu_{M_1}^\circ = 0$), but the activated dimer is unfavored with respect to the activated monomer ($\mu_{L_2}^\circ - 2\mu_{L_1}^\circ > 0$). As a consequence, the energy flow is always negative, and the term $RT\|\mathbf{b}\|\dot{\mathcal{I}}$ must always be positive and larger than $|\dot{\mathcal{E}}|$ in order to keep the subnetwork \mathcal{X} out of equilibrium. The value of the force F maximizing the concentration of the target species M_2 turns out to be $68 \text{ J K}^{-1} \text{ mol}^{-1}$.

In all the three scenarios, the internal free-energy transduction mechanism gets stalled for large values of the external force ($F \gtrsim 100 \text{ J K}^{-1} \text{ mol}^{-1}$): $\dot{\mathcal{E}}$ and $RT\|\mathbf{b}\|\dot{\mathcal{I}}$ go to zero. Thus, the subnetwork \mathcal{X} approaches equilibrium ($\dot{\Sigma}^{\mathcal{X}} \approx 0$), despite the subnetwork \mathcal{Y} being far from equilibrium. This can also be understood in terms of the so-called negative differential response of the steady state current [81]: j decreases when increasing F .

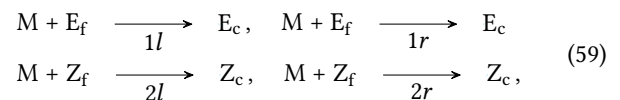
In all the three scenarios, the efficiency η is finite for small values of the external force, presents a maximum for intermediate values and then vanishes at large values. As shown in Appendix B, η is constant in the linear regime for very small forces. The fact that the efficiency is maximized for finite values of the force is therefore a truly far-from-equilibrium feature, which in this specific case can be ascribed to the negative differential response of the current j with respect to the force. Interestingly, in the third scenario (Fig. 2d), the value of the force maximizing the efficiency gets close to the one maximizing the concentration of the target species M_2 , indicating the possibility to fine-tune the parameters in order to optimize both the kinetic and the thermodynamic performance.

As a final comment, we notice that in the second and third scenarios the information flow constitutes the only free-energy source keeping the subnetwork \mathcal{X} out of equilibrium. This is in line with the fact that the functioning of this model is often described in terms of an information ratchet mechanism, hinting that the generation of information is the key working principle. However, the first scenario shows that there may be regimes where the information flow is negative or null, meaning that the information flow does not power the assembling reactions, but rather represents a cost. We therefore conclude that the thermodynamic concept of information flow is not synonymous with the kinetic concept of information ratcheting, but they are related. In particular, the connection appears to be vague only in those conditions where a positive energy flow (i.e., an energy ratchet like contribution) is present.

B. Light driven directional bimolecular motor

As a second application, we consider a class of synthetic bimolecular motors powered by light [24, 27, 45]. Their functioning consists in the autonomous directional threading and dethreading of a crown-ether macrocycle through an asymmetric molecular axle existing in two different conformations [43, 44].

This class of systems can be represented by the CRN in Fig. 3a. The threading of the macrocycle (or ring) M through the free axle, either in the conformation E_f or Z_f , can be described by four different self-assembly reactions forming the ring-axle complex, either in the conformation E_c or Z_c :



where labels l and r discriminate over which side of the axle the threading/de-threading of the ring occurs (see Fig. 3a). The reactions (59) follow mass action kinetics. The switching of the axle between its conformations in the free state as well as in the complex state can happen via four (in practice irreversible) coarse-grained photochemical reactions (see Appendix A 1 for the expression of their currents), all triggered

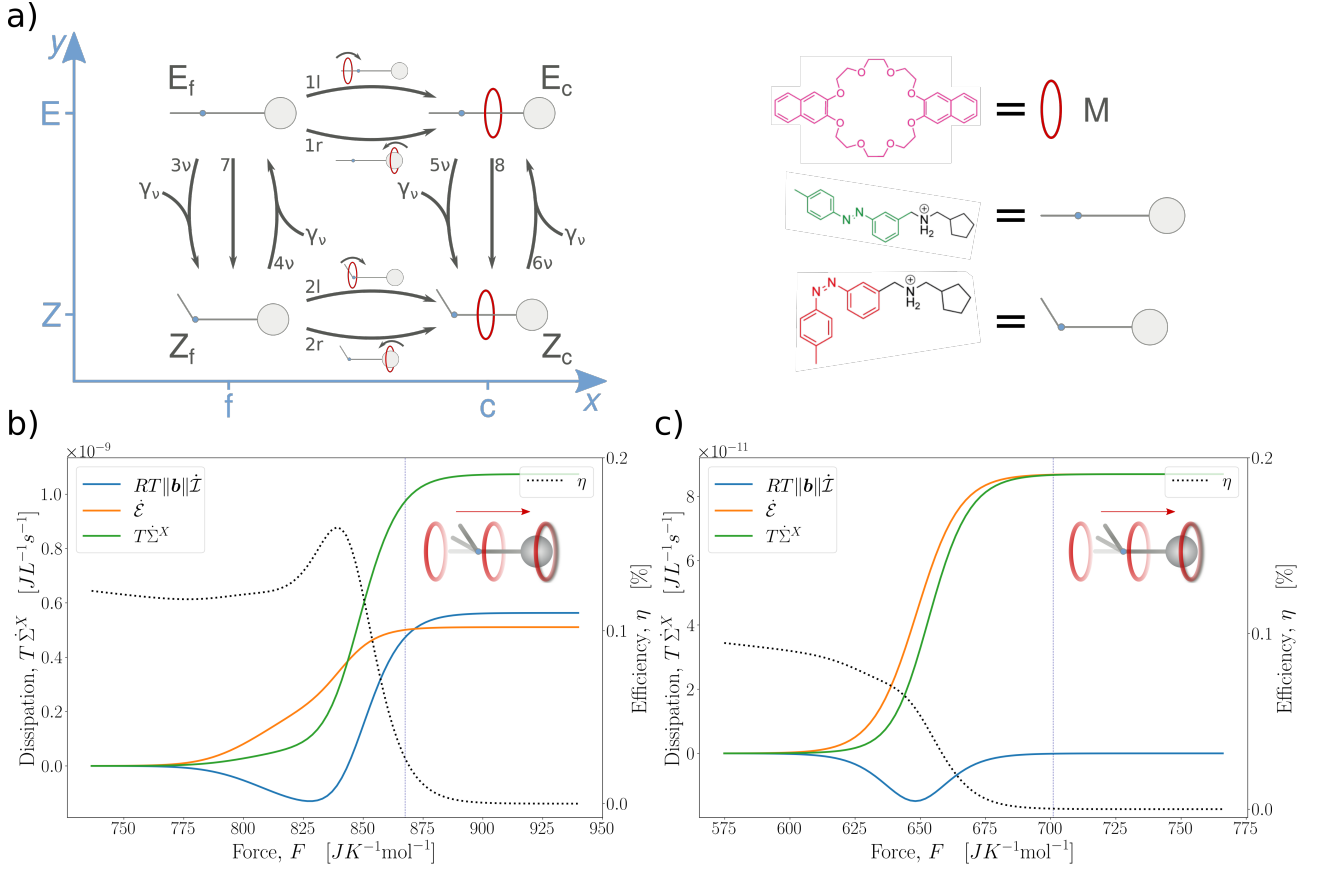
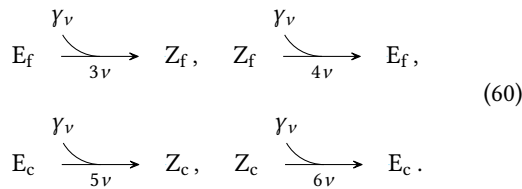
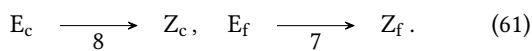


FIG. 3. **Light-driven bimolecular motor.** a) Hypergraph representation of the bipartite chemical reaction network and chemical structures of the involved species. Arrows are used to indicate the conventional direction of the reactions, which are all reversible (photochemical reactions are theoretically reversible [69], but practically irreversible as explained in A 1). Reactions are labeled according to equations (59), (60), and (61); b,c) Numerical simulations of the motor operated at 365 nm (b) and 436 nm (c). Information flow ($RT\|b\|\dot{I}$), energy flow ($\dot{\Sigma}$), and dissipation of the self-assembly reactions ($T\dot{\Sigma}^X = \dot{\Sigma} - RT\|b\|\dot{I}$) at steady state are plotted against the net thermodynamic force acting on the system ($F = \mu_v/T$, with $T = 298K$) in a range corresponding to a photon flow from 10^{-15} to 10^{-7} moles of photons per second. Experimentally probed conditions are marked by the vertical dotted lines. The efficiency of the internal free-energy transduction (η) is computed according to Eq. (64). Experimental parameters in both the regimes are taken from the original experimental data in Ref. [24] and Refs. [43, 44]. At steady state, on-average directed motion of the macrocycles M with respect to the axes is achieved, as a clockwise current arises with respect to the hypergraph and reactions $1r$ and $2l$ are kinetically unfavorable compared to reactions $1l$ and $2r$.

by photons γ_v at frequency ν :

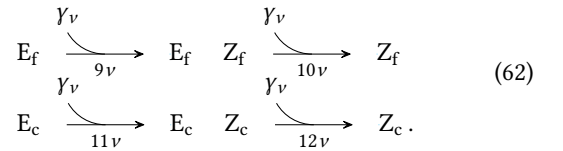


Changes of the axle's conformation are also possible via thermal reactions following mass-action kinetics:



Finally, four "futile" photochemical reactions are also possible, where photons are absorbed without leading to a confor-

mation change of the axle:



For this CRN, we define the set of bipartite species as $b = \{E_f, Z_f, E_c, Z_c\}$ and the set of ancillary species as $a = \{M\}$. Accordingly, reactions can be split into three sets. The futile photochemical reactions $A = \{9\nu, 10\nu, 11\nu, 12\nu\}$ do not lead to any change in the concentrations, but they contribute to the dissipation of the CRN, as shown in the following. The reactions $X = \{1l, 1r, 2l, 2r\}$ interconvert the assembled state of the bipartite species $x \in \{f, c\}$ between the free and the complex state. They thus define the self-assembly subnetwork \mathcal{X} . The reactions $Y = \{3\nu, 4\nu, 5\nu, 6\nu, 7, 8\}$ interconvert

the isomerization state of the bipartite species $y \in \{E, Z\}$. They thus define the isomerization subnetwork \mathcal{Y} .

As done for the previous example, we study the network's energetics at steady state by specializing Eqs. (53a), (53b) and (53c)

$$T\dot{\Sigma}^X = \underbrace{j(\mu_{Z_c}^\circ - \mu_{E_c}^\circ \quad \mu_{E_f}^\circ - \mu_{Z_f}^\circ)}_{=: \dot{\mathcal{E}}} \underbrace{RTj \ln \frac{[E_f][Z_c]}{[Z_f][E_c]}}_{=: RT\|\mathbf{b}\|\dot{\mathcal{I}}}, \quad (63a)$$

$$T\dot{\Sigma}^Y = \dot{W}_v^Y - \dot{\mathcal{E}} - RT\|\mathbf{b}\|\dot{\mathcal{I}}, \quad (63b)$$

$$T\dot{\Sigma}^A = \dot{W}_v^A, \quad (63c)$$

with $j = j_{1l} \quad j_{1r} = -(j_{2l} \quad j_{2r}) = -j_{3v} - j_7 \quad j_{4v} = j_{5v} \quad j_7 - j_{4v}$ the steady state net current (considered as positive when flowing clockwise across the hyper-graph in Fig. 3a), $\dot{\mathcal{I}} = -\dot{\mathcal{I}}_x = \dot{\mathcal{I}}_y$, and $\dot{\mathcal{E}}$ the energy flow. Furthermore, we introduced the two work sources \dot{W}_v^A and \dot{W}_v^Y quantifying the rates at which free-energy is provided by the radiation to futile and photoisomerization reactions, respectively (see Eq. (A11) and (A10) for their expressions). The former is fully dissipated by the futile reactions (see Eq. (63c)) and, consequently, it is useless to the motor's functioning. The latter accounts for the fraction of the free-energy absorbed from the radiation which can actually be used to sustain the motor's functioning [45]. Similarly to the previous example, the \mathcal{Y} subnetwork (involving the photoisomerization reactions) is the only one coupled to the external force $F = \mu_v/T$, but part of the free-energy exchanged with the radiation can be transduced towards subnetwork \mathcal{X} (involving self-assembly reactions) through the energy $\dot{\mathcal{E}}$ and information $\dot{\mathcal{I}}$ flows. These two mechanisms are the only possible free-energy sources for subnetwork \mathcal{X} which can be used at steady state to sustain the self-assembly reactions, i.e., the motor's functioning.

Based on the above understanding, the efficiency at which the free-energy harvested by subnetwork \mathcal{Y} from the radiation is transduced into free-energy available to subnetwork \mathcal{X} can be defined as:

$$0 \leq \eta = \frac{RT\|\mathbf{b}\|\dot{\mathcal{I}} \quad \dot{\mathcal{E}}}{\dot{W}_v^Y} = 1 - \frac{T\dot{\Sigma}^X}{\dot{W}_v^Y} \leq 1. \quad (64)$$

which is analogous to Eq. (57) except that the source work \dot{W}_v^Y is performed by the radiation instead of the chemostats. Similar comments on the properties of the efficiency η as those for the case of driven self-assembly apply here as well.

To illustrate our results, we simulated the CRN in Fig. 3a under two previously characterized experimental regimes which just differ by the wavelength (λ , assumed as monochromatic) of the radiation [24]. Based on experimental measures, the threading of the ring through the axles is always energetically favored independently of the isomerization state, but it is more favored in the E state, namely $\mu_{E_c}^\circ - \mu_{E_f}^\circ - \mu_M^\circ < \mu_{Z_c}^\circ - \mu_{Z_f}^\circ - \mu_M^\circ < 0$. Furthermore, because of steric hindrance, threading in the E state happens preferentially via reaction $1l$ in Eq. (59), and de-threading in the Z state happens preferentially via reaction $2r$. Such a kinetic preference favors the net directed motion of the ring with respect to the axles when the CRN is brought out of

equilibrium. This is often recognized as an energy ratchet mechanism in the literature of light-driven molecular motors, and provides the kinetic asymmetry allowing the the motor's functioning [10, 24, 27, 87, 88]. From a thermodynamic standpoint, this corresponds to a positive energy flow $\dot{\mathcal{E}}$ for any value of the external force since i) $(\mu_{Z_c}^\circ - \mu_{E_c}^\circ \quad \mu_{E_f}^\circ - \mu_{Z_f}^\circ) > 0$ and ii) $j > 0$ (see Eq. (63)).

In the first regime ($\lambda = 365\text{nm}$, Fig. 3b), the complexes E_c and Z_c are more effective than their free counterparts E_f and Z_f in absorbing light, thus causing photochemical reactions $5v$ and $6v$ in Eq. (60) to occur faster than the photochemical reactions $3v$ and $4v$ in Eq. (60). This is often recognized as an information ratchet mechanism providing an additional kinetic bias to the energy ratchet mechanism [10, 24, 27, 87, 88]. In the second regime ($\lambda = 436\text{nm}$, Fig. 3c), the photochemical properties of the axles are independent of the assembled state and, therefore, the motor's functioning is achieved by virtue of the energy ratchet mechanism alone [24]. In both cases, we find that relatively high values of the force are needed to take the self-assembly subnetwork out of equilibrium ($T\dot{\Sigma}^X > 0$) and, in contrast to the previous example, the dissipation increases monotonically up to a plateau, as well as the energy flow. On the other hand, the information flow is initially negative and reaches a plateau only after displaying a minimum. The plateau value is positive when the irradiation wavelength is 365 nm, while the information flow is always negative when the irradiation wavelength is 436 nm. As in the previous example, connections can be drawn between information/energy flow and information/energy ratcheting, but clearly the two concepts do not coincide. In particular, when the mechanism is a pure energy ratchet ($\lambda = 436$ nm, Fig. 3c), the only positive contribution to $T\dot{\Sigma}^X$ comes from the energy flow. When information ratcheting is added ($\lambda = 365$ nm, Fig. 3b) regimes with positive information flow arise. However, when both energy and information ratchet mechanisms are active, regimes where the information flow is null or negative are observed, and when there is no information ratchet mechanism, the information flow is always not null and negative.

Finally, in both cases the efficiency stays finite in the linear regime (see Appendix B) and drops to zero at high forces as $T\dot{\Sigma}^X$ saturates, with values of $3 \times 10^{-2}\%$ for $\lambda = 365$ nm and $3 \times 10^{-4}\%$ for $\lambda = 436$ nm at the experimentally probed conditions [24] (marked by the vertical dotted lines in Fig. 3b,c). Interestingly, while at $\lambda = 436$ nm the efficiency drops to zero monotonically, it presents a maximum at intermediate forces for $\lambda = 365$ nm.

VI. CONCLUSION

In this work, we formulated information thermodynamics for deterministic bipartite CRNs with nonlinear dynamics. The resulting picture is that these networks can be divided into two subnetworks made of different sets of reactions, which interact (due to the presence of shared species) by exchanging free-energy in the form of information and/or energy flows. To get there, we defined the mutual information

in terms of the concentrations of bipartite species (Eq. (44)), showing that it can be expressed in terms of Shannon-like contributions to the total entropy (Eq. (45)). Furthermore, we specialized the second law for each subnetwork (Eq. (53)), in analogy with information thermodynamics of Markov jump processes. Crucially, informational terms appear and may play the role of free-energy sources sustaining reactions which would not occur spontaneously. At odds with information thermodynamics of stochastic processes, information flow terms in deterministic CRNs must account for non-normalized concentration distributions. Also, variation of the subnetworks entropy due to the standard molar entropy carried by the chemical species and the contribution of non-bipartite species which may be present have to be taken into account.

Our framework generalizes a previous formulation developed for linear dynamics [40]. It directly applies to most of the artificial molecular machines and motors reported up to now, and to many relevant models in systems chemistry and biochemistry too (e.g., biomolecular motors, signal transduction). Here, we analyzed two epitomes of nonequilibrium supramolecular chemistry where the nonlinear self-assembly steps explicitly require leveraging our new results, as standard methods of information thermodynamics fail to treat non-normalized concentration distributions. Unprecedented insights were provided on i) how free-energy is used to power the chemically-driven self-assembly of monomers and an experimental light-driven bimolecular motor, and ii) on the corresponding performance evaluated according to their thermodynamic efficiencies that we introduced (Eqs. (57) and (64)). Furthermore, the role of information as a physical quantity is clarified and compared with the concept of information ratcheting. Albeit related, we found counterexamples to a general one-to-one correspondence between the two, thus establishing that the two concepts leverage two different notions of information, thus pertaining to distinct levels of descriptions (a thermodynamic level and a kinetic one).

Our results can be straightforwardly extended to cases where the external parameters c and n_v change in time according to a time dependent protocol [47], as well as when fluxes between the system and the reservoirs are controlled instead of the concentrations [51]. Also, if nonideal solutions were considered [50], additional terms would appear in the expression of the subnetworks entropy production without qualitatively altering the results. Furthermore, our approach should apply to multipartite CRNs too, namely when species and reactions can be split into more than two subnetworks exchanging information (e.g., a bipartite network where the species can also diffuse between two compartments defining an additional label for both species and reactions) [89, 90]. However, additional work is required to meaningfully explore this direction, especially concerning the notions of thermodynamic efficiency.

It is now clear that information processing is a key feature of living systems which is prone to quantitative analyses [91, 92]. With our work, we developed some of the theoretical tools needed to extend the approach of information thermodynamics towards the deterministic domain. As con-

centration distributions are often easier to characterize than probability distributions, we foresee that the current understanding of the role of information in chemical systems can benefit to some extent from the results that we derived.

ACKNOWLEDGMENTS

This research was funded by the Luxembourg National Research Fund, grant ChemComplex (C21/MS/16356329), and by the FQXi Foundation, grant ‘Information as a fuel in colloids and superconducting quantum circuits’ (FQXi-IAF19-05).

Appendix A: Photochemistry

In this appendix we derive the main quantities needed to reproduce the results in Section V B, namely, the expressions of photochemical reaction currents and the photochemical work.

1. Photochemical currents

In a typical photochemical experiment of the kind considered in Section V B, the solution is contained in a square *cuvette* and irradiated from one side with a fixed photon flow I_0 quantifying the amount of moles of photons impinging on the system’s surface Ω per unit time [93]. The radiation frequency is selected through interference filters with a bandwidth $\Delta\nu$ of approximately 10 nm in unit of wavelength [93], and all the frequency-dependent properties of the system are considered as constants over this interval. This approximation is equivalent to assuming the radiation to be monochromatic at the chosen frequency. According to the Beer-Lambert law, the intensity of the photon flow decreases along the direction of propagation r (with $r = 0$ being the point where the radiation starts interacting with the solution and $r = r_f$ being the optical path) as:

$$I(r) = I_0 \cdot 10^{-r\epsilon \cdot b}, \quad (\text{A1})$$

where ϵ is the vector collecting the molar extinction coefficients of each species. Note that the concentrations \mathbf{b} are considered uniform in the sample (i.e., we consider a well-mixed solution). To express the current of the photoisomerization reaction $\text{E} \xrightarrow{\text{light}} \text{Z}$, we compute the number of photons absorbed per unit of time and volume by the species E, and we multiply it by the quantum yield $\phi_{\text{E} \rightarrow \text{Z}}$ of the process, accounting for the probability that the photoisomerization happens once E absorbed a photon [94]:

$$j^{\text{E} \rightarrow \text{Z}} = \frac{I_0}{V} (1 - 10^{-r_f \epsilon \cdot b}) \frac{\epsilon_{\text{E}} [\text{E}]}{\epsilon \cdot \mathbf{b}} \phi_{\text{E} \rightarrow \text{Z}}. \quad (\text{A2})$$

The above expression simplifies under the the working conditions of interest [24], i.e., the total absorbance of the solution

at steady state is small ($r_f \epsilon \cdot \mathbf{b} \ll 1$), to

$$j^{E \rightarrow Z} = \frac{\ln(10) r_f \epsilon_E \phi_{E \rightarrow Z}}{V} I_0[E] = k_{E \rightarrow Z} I_0[E], \quad (\text{A3})$$

which is the expression we used in our simulations in Sec. VB. The currents of other photoisomerization reactions are derived in the same way, as well as the currents of futile photochemical reactions, like for $E \xrightarrow{\text{light}} E$

$$j^{E \rightarrow E} = \frac{\ln(10) r_f \epsilon_E \phi_{E \rightarrow E}}{V} I_0[E] = k_{E \rightarrow E} I_0[E], \quad (\text{A4})$$

with $\phi_{E \rightarrow E} = 1 - \phi_{E \rightarrow Z}$. Equations (A3) and (A4) show that photochemical reactions in low-absorbance regimes under controlled constant monochromatic irradiation can be treated as irreversible unimolecular mass action reactions [93, 94] by introducing effective kinetic constants, e.g., $k_{E \rightarrow Z}$ and $k_{E \rightarrow E}$. We note that a more detailed analysis taking into account the underlying excited state dynamics (photoisomerization reactions usually follow the so-called diabatic mechanism [95]) would show that this kind of reactions are in principle reversible, but the backward pathways (e.g., Z gets thermally excited and decays towards E by emitting a photon) are so unlikely that they can be safely neglected for any practical purpose, thus further justifying the above derivation [69].

2. Photochemical work

The concentration distribution of photons in the system can be expressed as $n_\nu(r) = n_\nu^T n_0 10^{-r\epsilon \cdot \mathbf{b}}$, where the black body distribution n_ν^T at the temperature T and n_0 account for the contribution due to the thermal bath and the external radiation, respectively. Within the experimental approximations, the latter can still be expressed in terms of a black body distribution ($n_0 = n_\nu^T$) by assigning to the external radiation the temperature T_r at which a black body would emit the same number of photons in the interval selected by the interference filter. To find T_r , we therefore impose

$$I_0 = \Omega c \int_{\nu - \Delta\nu/2}^{\nu + \Delta\nu/2} n_\nu^T d\nu' \approx \Omega c n_\nu^T \Delta\nu, \quad (\text{A5})$$

with Ω the irradiated area and c the speed of light. The approximation of treating the the black body distribution as almost constant over the integration interval $\Delta\nu$ is not necessary, but we mention it because it simplifies the identification of T_r without qualitatively changing the results. For the plots shown in Fig. 3, Eq. (A5) has been solved numerically without employing the mentioned approximation. In the low absorbance regime of interest in Ref. [24], we have that:

$$\begin{aligned} n_\nu(r) &= n_\nu^T n_\nu^T 10^{-r\epsilon \cdot \mathbf{b}} \\ &\approx n_\nu^T n_\nu^T (1 - \ln(10) r \epsilon \cdot \mathbf{b}). \end{aligned} \quad (\text{A6})$$

As a consequence, the chemical potential of the radiation (15) will be a function of r too. By considering $\mu_\nu(r)$ to be constant in the region of space from r to $r + dr$, the free energy

absorbed per unit of time from the radiation in the same region and in the low absorbance regime can be expressed as

$$\begin{aligned} &\mu_\nu(r) (I_0 (10^{-r\epsilon \cdot \mathbf{b}} - 10^{-(r+dr)\epsilon \cdot \mathbf{b}})) \\ &\approx \mu_\nu(r) I_0 \ln(10) \epsilon \cdot \mathbf{b} dr. \end{aligned} \quad (\text{A7})$$

The free energy harvested from the radiation by the whole system per unit time and volume is thus given by

$$\dot{W}_\nu = \frac{1}{V} \int_0^{r_f} \mu_\nu(r) I_0 \ln(10) \epsilon \cdot \mathbf{b} dr, \quad (\text{A8})$$

and it can be split into the contributions of Y and A reactions as

$$\dot{W}_\nu = \dot{W}_\nu^Y - \dot{W}_\nu^A, \quad (\text{A9})$$

with

$$\dot{W}_\nu^Y = \sum_{\rho \in Y_{\text{ph}}} \frac{j^\rho}{r_f} \int_0^{r_f} \mu_\nu(r) dr = \sum_{\rho \in Y} j^\rho \langle \mu_\nu(r) \rangle, \quad (\text{A10})$$

and

$$\dot{W}_\nu^A = \sum_{\rho \in A} \frac{j^\rho}{r_f} \int_0^{r_f} \mu_\nu(r) dr = \sum_{\rho \in A} j^\rho \langle \mu_\nu(r) \rangle, \quad (\text{A11})$$

with $Y_{\text{ph}} = \{3\nu, 4\nu, 5\nu, 6\nu\} \subset Y$ the set of photoisomerization reactions in Eq. (60). Equations (A10) and (A11) allow us to identify the average chemical potential $\langle \mu_\nu(r) \rangle$ as the photochemical force associated to the radiation, i.e., $F = \langle \mu_\nu(r) \rangle / T$, which is the expression used for the results reported in Fig. 3. We stress that this is a consequence of the low absorbance approximation, which makes the photochemical currents j_ρ independent of r . If this approximation did not hold, the currents j_ρ would enter the integrals in Eqs. (A10) and (A11) and the average chemical potential would not play the role of the net thermodynamic force [45].

Appendix B: Linear regime

In this appendix, we compute the efficiencies (57) and (64) in the linear regime, i.e., when the external force acting on the CRN is small enough ($F \ll R$) that the concentrations at steady state are well approximated by linear shifts from the equilibrium ones proportional to the external force according to $[\alpha]_{\text{ss}} = [\alpha]_{\text{eq}} (1 - m_\alpha F/R)$ with m_α the coefficient of proportionality. In these conditions, the chemical potentials (12) read $\mu_Z = \mu_Z^{\text{eq}} - \mathbf{m} T F$ with $\mathbf{m} = (\dots, m_\alpha, \dots)^\top$, while the current of a generic chemical reaction ρ following mass action kinetics can be expressed at the first order in F as $j^\rho = -j_{\text{eq}}^\rho (\mathbb{S}_\rho \cdot \mathbf{m}) F/R$ (where we used $j_{\text{eq}}^\rho = j_{\text{eq}}^{-\rho}$).

1. Driven self-assembly

Without loss of generality, we consider $[F] = [F]_{\text{eq}} (1 - F/R)$ and $[W] = [W]_{\text{eq}}$, which is consistent with $(\mu_F -$

$\mu_W)/T = F$ in the first order in F . Under these conditions, η as computed with Eq. (57) is constant up to the second order in F . Indeed, for $T\dot{\Sigma}^X$ we have

$$\begin{aligned} T\dot{\Sigma}^X &= j(\mu_{M_2} - \mu_{L_2} \quad 2\mu_{L_1} - 2\mu_{M_1}) = \\ &= \frac{TF^2}{R} j_{\text{eq}}^2 (2m_{L_1} - m_{L_2})(m_{M_2} - m_{L_2} \quad 2m_{L_1} - 2m_{M_1}), \end{aligned} \quad (\text{B1})$$

where the steady state current has been computed from reaction 2. For $I_F(\mu_F - \mu_W)$, we have

$$\begin{aligned} I_F(\mu_F - \mu_W) &= TF(j^{3F} \quad 2j^{4F}) = \\ &= \frac{TF^2}{R} (j_{\text{eq}}^{3F}(m_{M_1} \quad 1 - m_{L_1}) \quad 2j_{\text{eq}}^{4F}(m_{M_2} \quad 2 - m_{L_2})). \end{aligned} \quad (\text{B2})$$

Therefore:

$$\eta = \frac{j_{\text{eq}}^2 (2m_{L_1} - m_{L_2})(m_{M_2} - m_{L_2} \quad 2m_{L_1} - 2m_{M_1})}{j_{\text{eq}}^{3F}(m_{M_1} \quad 1 - m_{L_1}) \quad 2j_{\text{eq}}^{4F}(m_{M_2} \quad 2 - m_{L_2})}. \quad (\text{B3})$$

After analytically computing the coefficients \mathbf{m} by solving the steady state of the linearized dynamics, Eq. (B3) yields efficiencies of 7.5%, 3.7%, and 3.2% for the linear regimes of scenarios b, c and d, respectively.

2. Light driven bimolecular motor

In the case of the light driven bimolecular motor, we take $n_v(r) = n_v^T(1 - F10^{-r\epsilon \cdot \mathbf{b}_{\text{eq}}}/R)$. By doing so, the chemical potential (15) at first order in F reads $\mu_v(r) = TF10^{-r\epsilon \cdot \mathbf{b}_{\text{eq}}}$ provided that $f_v \gg n_v^T$, as it is the case at both the frequencies considered in Ref. [24]. For the equilibrium concentrations \mathbf{b}_{eq} , the low absorbance approximation does not hold, and therefore the free energy absorbed per unit of time from the radiation in the region of space between r and $r + dr$ reads

$$\begin{aligned} &\mu_v(r) \left(I_0^T \frac{F}{R} 10^{-r\epsilon \cdot \mathbf{b}_{\text{eq}}} - I_0^T \frac{F}{R} 10^{-(r+dr)\epsilon \cdot \mathbf{b}_{\text{eq}}} \right) \\ &= I_0^T \frac{TF^2}{R} 10^{-2r\epsilon \cdot \mathbf{b}_{\text{eq}}} \ln(10) \epsilon \cdot \mathbf{b}_{\text{eq}} dr, \end{aligned} \quad (\text{B4})$$

with $I_0^T = \Omega c \int_{v-\Delta v/2}^{v+\Delta v/2} n_{v'}^T dv'$. The free energy harvested from the radiation by the whole system per unit time and volume is obtained by integrating Eq. (B4) over r_f ,

$$\begin{aligned} \dot{W}_v &= \frac{I_0^T TF^2 \epsilon \cdot \mathbf{b}_{\text{eq}} \ln(10)}{VR} \int_0^{r_f} 10^{-2r\epsilon \cdot \mathbf{b}_{\text{eq}}} dr \\ &= \frac{I_0^T TF^2}{2VR} (1 - 10^{-2r_f \epsilon \cdot \mathbf{b}_{\text{eq}}}), \end{aligned} \quad (\text{B5})$$

and it can be decomposed in the contributions of Y and A reactions as

$$\dot{W}_v = \dot{W}_v^Y \quad \dot{W}_v^A, \quad (\text{B6})$$

with

$$\begin{aligned} \dot{W}_v^Y &= \dot{W}_v \Gamma_{\text{eq}}^Y = \\ &= \dot{W}_v \left(\epsilon_{E_f} [E_f]_{\text{eq}} \phi_{E_f \rightarrow Z_f} \quad \epsilon_{Z_f} [Z_f]_{\text{eq}} \phi_{Z_f \rightarrow E_f} \right. \\ &\quad \left. \epsilon_{E_c} [E_c]_{\text{eq}} \phi_{E_c \rightarrow Z_c} \quad \epsilon_{Z_c} [Z_c]_{\text{eq}} \phi_{Z_c \rightarrow E_c} \right) / \epsilon \cdot \mathbf{b}_{\text{eq}} \end{aligned} \quad (\text{B7})$$

and

$$\begin{aligned} \dot{W}_v^A &= \dot{W}_v \Gamma_{\text{eq}}^A = \\ &= \dot{W}_v \left(\epsilon_{E_f} [E_f]_{\text{eq}} \phi_{E_f \rightarrow E_f} \quad \epsilon_{Z_f} [Z_f]_{\text{eq}} \phi_{Z_f \rightarrow Z_f} \right. \\ &\quad \left. \epsilon_{E_c} [E_c]_{\text{eq}} \phi_{E_c \rightarrow E_c} \quad \epsilon_{Z_c} [Z_c]_{\text{eq}} \phi_{Z_c \rightarrow Z_c} \right) / \epsilon \cdot \mathbf{b}_{\text{eq}}, \end{aligned} \quad (\text{B8})$$

where Γ_{eq}^Y and Γ_{eq}^A quantify the fractions of the power which are provided by the radiation to photoisomerizations and futile reactions, respectively (with $\Gamma_{\text{eq}}^Y + \Gamma_{\text{eq}}^A = 1$). Finally, for the sum of the energy and information flows we have:

$$\begin{aligned} T\dot{\Sigma}^X &= j(\mu_{Z_c} - \mu_{E_c} \quad \mu_{E_f} - \mu_{Z_f}) = \\ &= \frac{TF^2}{R} (j_{\text{eq}}^{2I} \quad j_{\text{eq}}^{2r})(m_{Z_c} - m_{Z_f} - m_M)(m_{Z_c} - m_{E_c} \quad m_{E_f} - m_{Z_f}). \end{aligned} \quad (\text{B9})$$

Therefore, Eq. (64) in the linear regime boils down to

$$\eta = \frac{2V(j_{\text{eq}}^{2I} \quad j_{\text{eq}}^{2r})(m_{Z_c} - m_{Z_f} - m_M)(m_{Z_c} - m_{E_c} \quad m_{E_f} - m_{Z_f})}{I_0^T (1 - 10^{-2r_f \epsilon \cdot \mathbf{b}_{\text{eq}}}) \Gamma_{\text{eq}}^Y}, \quad (\text{B10})$$

which verifies that the efficiency is constant up to the second order in F . Here, solving the steady state of the linearized dynamics to compute the coefficients \mathbf{m} and then find the actual values of η in the two experimental regimes would require to know the rates of the backward photochemical reactions, which cease to be negligible at very small forces.

[1] Shuntaro Amano, Stefan Borsley, David A. Leigh, and Zhanhu Sun, "Chemical engines: driving systems away from equilibrium through catalyst reaction cycles," *Nature Nanotechnology*

16, 1057–1067 (2021).

[2] R. Dean Astumian and Martin Bier, "Fluctuation driven ratchets: Molecular motors," *Phys. Rev. Lett.* 72, 1766–1769 (1994).

- [3] Marcelo O. Magnasco, "Molecular combustion motors," *Phys. Rev. Lett.* **72**, 2656–2659 (1994).
- [4] David Keller and Carlos Bustamante, "The mechanochemistry of molecular motors," *Biophys. J.* **78**, 541–556 (2000).
- [5] Carlos Bustamante, David Keller, and George Oster, "The physics of molecular motors," *Acc. Chem. Res.* **34**, 412–420 (2001).
- [6] Anatoly B. Kolomeisky and Michael E. Fisher, "Molecular motors: A theorist's perspective," *Annu. Rev. Phys. Chem.* **58**, 675–695 (2007).
- [7] Reinhard Lipowsky and Steffen Liepelt, "Chemomechanical coupling of molecular motors: Thermodynamics, network representations, and balance conditions," *J. Stat. Phys.* **130**, 39–67 (2008).
- [8] Mauro L. Mugnai, Changbong Hyeon, Michael Hinczewski, and D. Thirumalai, "Theoretical perspectives on biological machines," *Rev. Mod. Phys.* **92**, 025001 (2020).
- [9] Aidan I. Brown and David A. Sivak, "Theory of nonequilibrium free energy transduction by molecular machines," *Chemical Reviews* **120**, 434–459 (2020).
- [10] Euan R. Kay, David A. Leigh, and Francesco Zerbetto, "Synthetic molecular motors and mechanical machines," *Angew. Chem. Int. Ed.* **46**, 72–191 (2007).
- [11] R. Dean Astumian, "Design principles for brownian molecular machines: how to swim in molasses and walk in a hurricane," *Phys. Chem. Chem. Phys.* **9**, 5067–5083 (2007).
- [12] Elio Mattia and Sijbren Otto, "Supramolecular systems chemistry," *Nat. Nanotechnol.* **10**, 111–119 (2015).
- [13] Krishnendu Das, Luca Gabrielli, and Leonard J. Prins, "Chemically fueled self-assembly in biology and chemistry," *Angew. Chem. Int. Ed.* **60**, 20120–20143 (2021).
- [14] Manfred Schliwa and Günther Woehlke, "Molecular motors," *Nature* **422**, 759–765 (2003).
- [15] Job Boekhoven, Aurelie M. Brizard, Krishna N. K. Kowligi, Ger J. M. Koper, Rienk Eelkema, and Jan H. van Esch, "Dissipative self-assembly of a molecular gelator by using a chemical fuel," *Angew. Chem. Int. Ed.* **49**, 4825–4828 (2010).
- [16] Chuyang Cheng, Paul R. McGonigal, Severin T. Schneebeli, Hao Li, Nicolaas A. Vermeulen, Chenfeng Ke, and J. Fraser Stoddart, "An artificial molecular pump," *Nat. Nanotechnol.* **10**, 547–553 (2015).
- [17] Miriam R. Wilson, Jordi Solà, Armando Carlone, Stephen M. Goldup, Nathalie Lebrasseur, and David A. Leigh, "An autonomous chemically fuelled small-molecule motor," *Nature* **534**, 235–240 (2016).
- [18] Sundus Erbas-Cakmak, Stephen D. P. Fielden, Ulvi Karaca, David A. Leigh, Charlie T. McTernan, Daniel J. Tetlow, and Miriam R. Wilson, "Rotary and linear molecular motors driven by pulses of a chemical fuel," *Science* **358**, 340–343 (2017).
- [19] Julien Robert-Paganin, Olena Pylypenko, Carlos Kikutu, H. Lee Sweeney, and Anne Houdusse, "Force generation by myosin motors: A structural perspective," *Chem. Rev.* **120**, 5–35 (2020).
- [20] Shuntaro Amano, Stephen D. P. Fielden, and David A. Leigh, "A catalysis-driven artificial molecular pump," *Nature* **594**, 529–534 (2021).
- [21] Nagatoshi Koumura, Robert W. J. Zijlstra, Richard A. van Delden, Nobuyuki Harada, and Ben L. Feringa, "Light-driven unidirectional molecular rotor," *Nature* **401**, 152–155 (1999).
- [22] Viviana Serreli, Chin-Fa Lee, Euan R. Kay, and David A. Leigh, "A molecular information ratchet," *Nature* **445**, 523–527 (2007).
- [23] Lutz Greb and Jean-Marie Lehn, "Light-driven molecular motors: Imines as four-step or two-step unidirectional rotors," *J. Am. Chem. Soc.* **136**, 13114–13117 (2014).
- [24] Giulio Ragazzon, Massimo Baroncini, Serena Silvi, Margherita Venturi, and Alberto Credi, "Light-powered autonomous and directional molecular motion of a dissipative self-assembling system," *Nat. Nanotechnol.* **10**, 70–75 (2015).
- [25] Quan Li, Gad Fuks, Emilie Moulin, Mounir Maaloum, Michel Rawiso, Igor Kulic, Justin T. Foy, and Nicolas Giuseppone, "Macroscopic contraction of a gel induced by the integrated motion of light-driven molecular motors," *Nat. Nanotechnol.* **10**, 161–165 (2015).
- [26] Justin T. Foy, Quan Li, Antoine Goujon, Jean-Rémy Colard-Itté, Gad Fuks, Emilie Moulin, Olivier Schiffmann, Damien Dattler, Daniel P. Funeriu, and Nicolas Giuseppone, "Dual-light control of nanomachines that integrate motor and modulator subunits," *Nat. Nanotechnol.* **12**, 540–545 (2017).
- [27] Martina Canton, Jessica Groppi, Lorenzo Casimiro, Stefano Corra, Massimo Baroncini, Serena Silvi, and Alberto Credi, "Second-generation light-fueled supramolecular pump," *J. Am. Chem. Soc.* **143**, 10890–10894 (2021).
- [28] Guillaume Vives, Henri-Pierre Jacquot de Rouville, Alexandre Carella, Jean-Pierre Launay, and Gwénaél Rapenne, "Prototypes of molecular motors based on star-shaped organometallic ruthenium complexes," *Chem. Soc. Rev.* **38**, 1551–1561 (2009).
- [29] Cristian Pezzato, Minh T. Nguyen, Dong Jun Kim, Ommid Anamimoghadam, Lorenzo Mosca, and J. Fraser Stoddart, "Controlling dual molecular pumps electrochemically," *Angew. Chem. Int. Ed.* **57**, 9325–9329 (2018).
- [30] D Hartich, A C Barato, and U Seifert, "Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a maxwell's demon interpretation," *J. Stat. Mech.: Theory Exp.* **2014**, P02016 (2014).
- [31] Jordan M. Horowitz and Massimiliano Esposito, "Thermodynamics with continuous information flow," *Phys. Rev. X* **4**, 031015 (2014).
- [32] Takahiro Sagawa and Masahito Ueda, "Generalized jarzynski equality under nonequilibrium feedback control," *Phys. Rev. Lett.* **104**, 090602 (2010).
- [33] Juan M. R. Parrondo, Jordan M. Horowitz, and Takahiro Sagawa, "Thermodynamics of information," *Nat. Phys.* **11**, 131 (2015).
- [34] Jordan M. Horowitz, Takahiro Sagawa, and Juan M. R. Parrondo, "Imitating chemical motors with optimal information motors," *Phys. Rev. Lett.* **111**, 010602 (2013).
- [35] Andre C. Barato and Udo Seifert, "Stochastic thermodynamics with information reservoirs," *Phys. Rev. E* **90**, 042150 (2014).
- [36] Dimitri Loutchko, Maximilian Eisbach, and Alexander S. Mikhailov, "Stochastic thermodynamics of a chemical nanomachine: The channeling enzyme tryptophan synthase," *J. Chem. Phys.* **146**, 025101 (2017).
- [37] Ryota Takaki, Mauro L. Mugnai, and D. Thirumalai, "Information flow, gating, and energetics in dimeric molecular motors," *BioRxiv:2021.12.30.474541* (2021).
- [38] Thomas McGrath, Nick S. Jones, Pieter Rein ten Wolde, and Thomas E. Ouldridge, "Biochemical machines for the interconversion of mutual information and work," *Phys. Rev. Lett.* **118**, 028101 (2017).
- [39] Emma Lathouwers and David A. Sivak, "Internal energy and information flows mediate input and output power in bipartite molecular machines," *Phys. Rev. E* **105**, 024136 (2022).
- [40] Shuntaro Amano, Massimiliano Esposito, Elisabeth Kreidt, David A. Leigh, Emanuele Penocchio, and Benjamin M. W. Roberts, "Insights from an information thermodynamics analysis of a synthetic molecular motor," *Nat. Chem.* (2022), 10.1038/s41557-022-00899-z.

- [41] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory* (John Wiley & Sons, Inc., Hoboken, 2012).
- [42] G. Ragazzon and L. J. Prins, "Energy consumption in chemical fuel-driven self-assembly," *Nat. Nanotechnol.* **13**, 882–889 (2018).
- [43] Andrea Sabatino, Emanuele Penocchio, Giulio Ragazzon, Alberto Credi, and Diego Frezzato, "Individual-molecule perspective analysis of chemical reaction networks: The case of a light-driven supramolecular pump," *Angew. Chem. Int. Ed.* **58**, 14341–14348 (2019).
- [44] Daniele Asnicar, Emanuele Penocchio, and Diego Frezzato, "Sample size dependence of tagged molecule dynamics in steady-state networks with bimolecular reactions: Cycle times of a light-driven pump," *The Journal of Chemical Physics* **156**, 184116 (2022).
- [45] S. Corrà, Marina Tranfić Bakić, Jessica Groppi, Serena Silvi, Emanuele Penocchio, Massimiliano Esposito, and Alberto Credi, "Kinetic and energetic insights in the dissipative non-equilibrium operation of an autonomous light-powered supramolecular pump," *ChemRxiv:2021-8pwcc* (2021).
- [46] R. Dean Astumian, "Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet," *Nat. Commun.* **10**, 3837 (2019).
- [47] Danilo Forastiere, Gianmaria Falasco, and Massimiliano Esposito, "Strong current response to slow modulation: A metabolic case-study," *J. Chem. Phys.* **152**, 134101 (2020).
- [48] Gianmaria Falasco, Riccardo Rao, and Massimiliano Esposito, "Information thermodynamics of turing patterns," *Phys. Rev. Lett.* **121**, 108301 (2018).
- [49] Francesco Avanzini, Gianmaria Falasco, and Massimiliano Esposito, "Thermodynamics of chemical waves," *J. Chem. Phys.* **151**, 234103 (2019).
- [50] Francesco Avanzini, Emanuele Penocchio, Gianmaria Falasco, and Massimiliano Esposito, "Nonequilibrium thermodynamics of non-ideal chemical reaction networks," *J. Chem. Phys.* **154**, 094114 (2021).
- [51] Francesco Avanzini and Massimiliano Esposito, "Thermodynamics of concentration vs flux control in chemical reaction networks," *J. Chem. Phys.* **156**, 014116 (2022).
- [52] Yiyu Xiang, Neil Dalchau, and Baojun Wang, "Scaling up genetic circuit design for cellular computing: advances and prospects," *Nat. Comput.* **17**, 833–853 (2018).
- [53] Lewis Grozinger, Martyn Amos, Thomas E. Gorochowski, Pablo Carbonell, Diego A. Oyarzún, Ruud Stoof, Harold Fellermann, Paolo Zuliani, Huseyin Tas, and Angel Goñi-Moreno, "Pathways to cellular supremacy in biocomputing," *Nat. Commun.* **10**, 5250 (2019).
- [54] Alexander A. Green, Jongmin Kim, Duo Ma, Pamela A. Silver, James J. Collins, and Peng Yin, "Complex cellular logic computation using ribocomputing devices," *Nature* **548**, 117–121 (2017).
- [55] Fei Wang, Hui Lv, Qian Li, Jiang Li, Xueli Zhang, Jiye Shi, Lihua Wang, and Chunhai Fan, "Implementing digital computing with dna-based switching circuits," *Nat. Commun.* **11**, 1–8 (2020).
- [56] Lulu Qian, Erik Winfree, and Jehoshua Bruck, "Neural network computation with dna strand displacement cascades," *Nature* **475**, 368–372 (2011).
- [57] J.M. Deutsch, "Computational mechanisms in genetic regulation by rna," *J. Theor. Biol.* **458**, 156–168 (2018).
- [58] William Poole, Thomas Ouldridge, Manoj Gopalkrishnan, and Erik Winfree, "Detailed balanced chemical reaction networks as generalized boltzmann machines," Preprint at <https://arxiv.org/abs/2205.06313> (2022).
- [59] Giulia Malaguti and Pieter Rein ten Wolde, "Theory for the optimal detection of time-varying signals in cellular sensing systems," *eLife* **10**, e62574 (2021).
- [60] David Andrieux and Pierre Gaspard, "Nonequilibrium generation of information in copolymerization processes," *Proc. Natl. Acad. Sci.* **105**, 9516–9521 (2008).
- [61] Pablo Sartori and Simone Pigolotti, "Thermodynamics of error correction," *Phys. Rev. X* **5**, 041039 (2015).
- [62] Thomas E. Ouldridge, Christopher C. Govern, and Pieter Rein ten Wolde, "Thermodynamics of computational copying in biochemical systems," *Phys. Rev. X* **7**, 021004 (2017).
- [63] Jenny M. Poulton, Pieter Rein ten Wolde, and Thomas E. Ouldridge, "Nonequilibrium correlations in minimal dynamical models of polymer copying," *Proc. Natl. Acad. Sci.* **116**, 1946–1951 (2019).
- [64] Adrien Padirac, Teruo Fujii, and Yannick Rondelez, "Bottom-up construction of in vitro switchable memories," *Proc. Natl. Acad. Sci.* **109**, E3212–E3220 (2012).
- [65] Fabian Schnitter, Benedikt Rieß, Christian Jandl, and Job Boekhoven, "Memory, switches, and an or-port through bistability in chemically fueled crystals," *Nat. Commun.* **13**, 2816 (2022).
- [66] Svehla G., "Nomenclature of kinetic methods of analysis (iupac recommendations 1993)," *Pure and Applied Chemistry*, **65**, 2291 (1993).
- [67] Artur Wachtel, Riccardo Rao, and Massimiliano Esposito, "Thermodynamically consistent coarse graining of biocatalysts beyond michaelis-menten," *New J. Phys.* **20**, 042002 (2018).
- [68] Francesco Avanzini, Gianmaria Falasco, and Massimiliano Esposito, "Thermodynamics of non-elementary chemical reaction networks," *New J. Phys.* **22**, 093040 (2020).
- [69] Emanuele Penocchio, Riccardo Rao, and Massimiliano Esposito, "Nonequilibrium thermodynamics of light-induced reactions," *J. Chem. Phys.* **155**, 114101 (2021).
- [70] S. R. de Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (Dover, 1984).
- [71] K. J. Laidler, *Chemical Kinetics* (Harper Collins Publishers, 1987).
- [72] Miloslav Pekař, "Thermodynamics and foundations of mass-action kinetics," *Prog. React. Kinet. Mech.* **30**, 3–113 (2005).
- [73] Stefan Borsley, Elisabeth Kreidt, David A. Leigh, and Benjamin M. W. Roberts, "Autonomous fuelled directional rotation about a covalent single bond," *Nature* **604**, 80–85 (2022).
- [74] Riccardo Rao and Massimiliano Esposito, "Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics," *Phys. Rev. X* **6**, 041064 (2016).
- [75] Robert Holyst and Andrzej Poniewierski, *Thermodynamics for Chemists, Physicists and Engineers* (Springer, Dordrecht (NL), 2012).
- [76] Harald Ries and A.J. McEvoy, "Chemical potential and temperature of light," *J. Photochem. Photobiol. A* **59**, 11–18 (1991).
- [77] Peter Gräber and Giulio Milazzo, *Bioenergetics* (Birkhäuser Basel, 1997).
- [78] I. Prigogine, *Introduction to Thermodynamics of Irreversible Processes*, 2nd ed. (Interscience Publishers, New York, 1961).
- [79] Ilya Prigogine Dilip Kondepudi, *Modern Thermodynamics. From heat engines to dissipative structures*, 2nd ed. (Wiley, 2015).
- [80] Emanuele Penocchio, Riccardo Rao, and Massimiliano Esposito, "Thermodynamic efficiency in dissipative chemistry," *Nat. Commun.* **10**, 3865 (2019).
- [81] Gianmaria Falasco, Tommaso Cossetto, Emanuele Penocchio, and Massimiliano Esposito, "Negative differential response in chemical reactions," *New J. Phys.* **21**, 073005 (2019).

- [82] Arshad Desai and Timothy J. Mitchison, "Microtubule polymerization dynamics," *Annu. Rev. Cell Dev. Biol.* **13**, 83–117 (1997).
- [83] H. Hess and Jennifer L. Ross, "Non-equilibrium assembly of microtubules: from molecules to autonomous chemical robots," *Chem. Soc. Rev.* **46**, 5570–5587 (2017).
- [84] J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton* (Sunderland, MA: Sinauer Associates, 2001).
- [85] Job Boekhoven, Wouter E. Hendriksen, Ger J. M. Koper, Rienk Eelkema, and Jan H. van Esch, "Transient assembly of active materials fueled by a chemical reaction," *Science* **349**, 1075–1079 (2015).
- [86] R.D. Astumian and M. Bier, "Mechanochemical coupling of the motion of molecular motors to atp hydrolysis," *Biophys. J.* **70**, 637–653 (1996).
- [87] R. D. Astumian, "Optical vs. chemical driving for molecular machines," *Faraday Discuss.* **195**, 583–597 (2016).
- [88] Stefano Corrà, Lorenzo Casimiro, Massimo Baroncini, Jessica Groppi, Marcello La Rosa, Marina Tranfić Bakić, Serena Silvi, and Alberto Credi, "Artificial supramolecular pumps powered by light," *Chem. Eur. J.* **27**, 11076–11083 (2021).
- [89] Jordan M Horowitz, "Multipartite information flow for multiple maxwell demons," *J. Stat. Mech.* **2015**, P03006 (2015).
- [90] David H Wolpert, "Minimal entropy production rate of interacting systems," *New Journal of Physics* **22**, 113013 (2020).
- [91] Gašper Tkačik and William Bialek, "Information processing in living systems," *Annu. Rev. Condens. Matter Phys.* **7**, 89–117 (2016).
- [92] H. H. Mattingly, K. Kamino, B. B. Machta, and T. Emonet, "Escherichia coli chemotaxis is information limited," *Nature Physics* **17**, 1426–1431 (2021).
- [93] Marco Montalti, Alberto Credi, Luca Prodi, and Teresa Gandolfi, *Handbook of Photochemistry*, 3rd ed. (CRC Press, 2006).
- [94] H. Mauser and G. Gauglitz, *Photokinetics : theoretical fundamentals and applications* (Elsevier, 1998).
- [95] Vincenzo Balzani, Paola Ceroni, and Alberto Juris, *Photochemistry and photophysics : concepts, research, applications*, 1st ed. (Wiley-VCH Verlag, 2014).

The following article is reprinted from
[S. AMANO, M. ESPOSITO, E. KREIDT, D. A. LEIGH, E. PENOCCHIO and B. M. W.
ROBERTS, *Nat. Chem.* **14**. (2022), 530–537]
under the conditions of the Springer Nature Limited license².

The page numbers placed in the outer margins provide a continuous pagination throughout the thesis.

² <https://www.nature.com/nature-portfolio/reprints-and-permissions/permissions-requests>



Insights from an information thermodynamics analysis of a synthetic molecular motor

Shuntaro Amano¹, Massimiliano Esposito², Elisabeth Kreidt¹, David A. Leigh¹✉, Emanuele Penocchio²✉ and Benjamin M. W. Roberts¹

Information is physical, a realization that has transformed the physics of measurement and communication. However, the flow between information, energy and mechanics in chemical systems remains largely unexplored. Here we analyse a minimalist autonomous chemically driven molecular motor in terms of information thermodynamics, a framework that quantitatively relates information to other thermodynamic parameters. The treatment reveals how directional motion is generated by free energy transfer from chemical to mechanical (conformational and/or co-conformational) processes by ‘energy flow’ and ‘information flow’. It provides a thermodynamic level of understanding of molecular motors that is general, complements previous analyses based on kinetics and has practical implications for machine design. In line with kinetic analysis, we find that power strokes do not affect the directionality of chemically driven machines. However, we find that power strokes can modulate motor velocity, the efficiency of free energy transfer and the number of fuel molecules consumed per cycle. This may help explain the role of such (co-)conformational changes in biomachines and illustrates the interplay between energy and information in chemical systems.

Understanding how and why a machine works in the way it does is crucial for optimizing designs and inventing new ones. For macroscopic machines, such an understanding can be deduced from engineering principles and Newtonian physics. In contrast, at the molecular level there is no simple explanation for why individual components of biomachines move in a particular way. Is the movement of a specific amino acid in adenosine triphosphate synthase a key requisite for the mechanism, or does it occur incidentally as part of an evolutionary pathway that was overall successful? Biomolecular machines operate autonomously, apparently through Brownian ratchet mechanisms^{1–5}. They use energy and information to rectify the directionality of random thermal movements of their components so that work can be performed. These types of machines include biological pumps and motors, in which the energy is generally provided in the form of a chemical potential gradient^{6,7}. Chemists are learning how to design synthetic analogues of such machines^{8–13}, with examples including a minimalist autonomous chemically fuelled molecular rotary motor¹⁴ (Fig. 1).

Kinetic models have proved useful^{15–17} in describing the behaviour of molecular machines, but such analysis is inherently unable to give an account of free energy transfer from the fuel to the machine. The consideration of thermodynamics is essential to understand free energy transduction, how this drives directional motion and generates the capacity to perform work¹⁸ and, hence, how to optimize the motor’s design. Up to now, attempts^{12,19–24} to design molecular motors have been led by chemical intuition, with little opportunity to reliably judge the effectiveness of a machine’s design or performance unless and until it has been realized experimentally.

Recently, stochastic thermodynamics has emerged as a method for studying systems that operate at energies of the order of thermal fluctuations^{25–28}. The theory is valid even when systems are driven far from equilibrium^{25,29,30} and has been used to study nonequilibrium nanoscale systems^{31–37}. A major achievement in this field was the connection made between stochastic thermodynamics and

information theory³⁸, which gave birth to ‘information thermodynamics’^{39,40}. Information thermodynamics relates information to other thermodynamic quantities, such as free energy and entropy, and has proved particularly successful in resolving apparent thermodynamic paradoxes, such as Maxwell’s demon⁴¹.

In this Article we develop a quantitative understanding of the processes that drive an autonomous chemically fuelled molecular motor¹⁴ (Fig. 1) using an analysis that incorporates concepts from information thermodynamics within the framework of nonequilibrium thermodynamics of open chemical reaction networks^{42,43}. Contrary to common models in stochastic thermodynamics^{25,30,37}, our framework neglects fluctuations and describes macroscopic ensembles of chemical species characterized by experimentally measurable concentrations. The approach is consistent with kinetic models^{16,17,44}, but goes further by introducing a quantitative thermodynamic understanding of how autonomous molecular motors work. Two features—information flow and energy flow—contribute to the transfer of free energy from the fuel to the machine that is the origin of current (that is, the net rate of displacement of the macrocycle directionally along the track; Box 1) in the motor. The effect of changing chemical gating¹⁵, power strokes^{45–47} (a viscoelastic, free energy-releasing, large-amplitude conformational change^{45,47}) and overall rates on current and efficiency (Box 1) are examined through simulations, revealing design principles for molecular motors. Particular insight is gained in terms of the role of power strokes in tuning a motor’s performance while remaining consistent with core aspects of kinetic models^{16,17,44–48}, informing the current debate concerning the role of power strokes in biomolecular machines^{44,46,47}.

As a result, the mechanism of operation of the rotary motor can be understood in several different ways: through chemical design¹⁴, reaction kinetics^{15,49}, molecular dynamics⁵⁰ and, now, nonequilibrium information thermodynamics. Accordingly, this minimalist molecular motor can act as a Rosetta Stone for relating these disparate frameworks^{16,17,35,44,49–51}, aiding the translation

¹Department of Chemistry, University of Manchester, Manchester, UK. ²Department of Physics and Materials Science, University of Luxembourg, Luxembourg City, Luxembourg. ✉e-mail: david.leigh@manchester.ac.uk; emanuele.penocchio@uni.lu

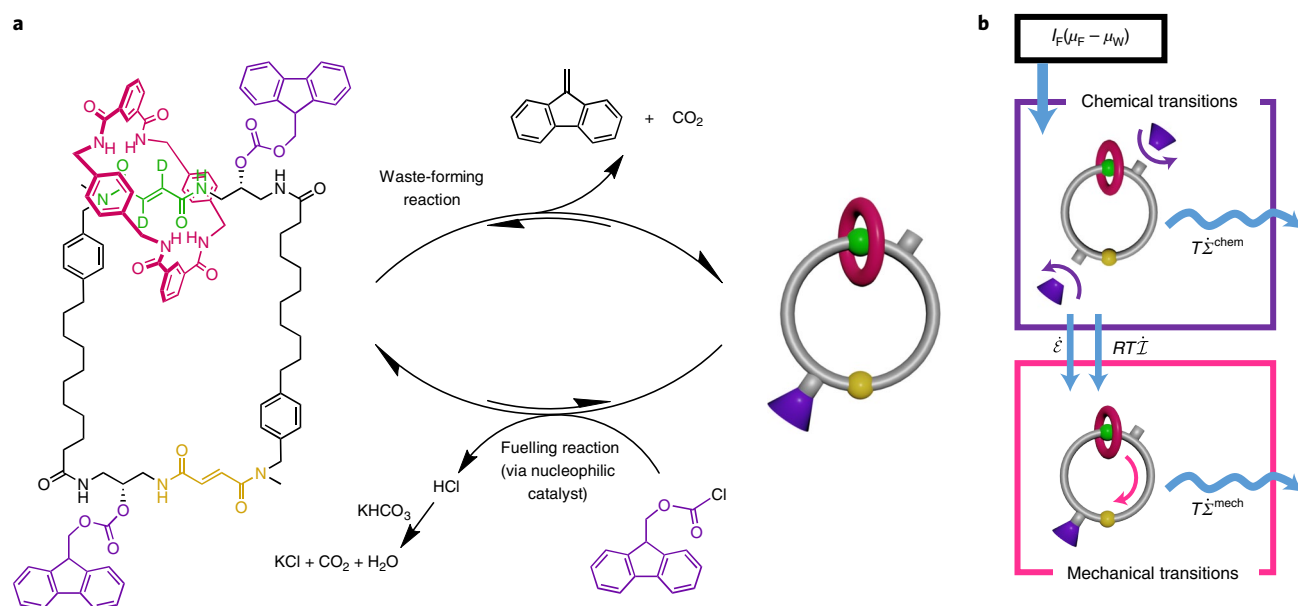


Fig. 1 | A Rosetta Stone for chemical (reactions and co-conformational dynamics) and information thermodynamics descriptions of a molecular motor: two distinct but complementary accounts of the processes involved in a minimalist, autonomous, chemically fuelled, molecular rotary motor.

a. Chemical structure of the rotary motor and the chemical reactions involved in its operation. The motor comprises a benzylic amide macrocycle (magenta) and a track with two fumaramide binding sites (yellow, non-deuterated; green, deuterated for analytical purposes). The macrocycle randomly shuttles between the two fumaramide sites when its path is not blocked by Fmoc groups (purple). The fuelling reaction consumes the fuel (Fmoc-Cl) and attaches an Fmoc group to the track, and the waste-forming reaction removes the Fmoc group (allowing passage of the macrocycle) and generates waste species (dibenzofulvene and CO_2). The fuelling reaction is catalysed by a pyridine-based nucleophilic catalyst (for example, 4-dimethylaminopyridine or the bulky catalyst shown in Fig. 4a). Note that both reactions are considered reversible, even when the backward reactions (that is, regeneration of fuel via barrier removal and waste products reacting to give the barrier) are extremely rare events⁶⁶. **b.** Information thermodynamics description of free energy transduction in the rotary motor. In the chemical transitions, free energy is supplied to the motor ($I_F(\mu_F - \mu_W)$), part of which is dissipated ($T\dot{\Sigma}^{\text{chem}}$). I_F is the rate at which the motor reacts with the fuel, $(\mu_F - \mu_W)$ is the chemical potential difference between the fuel and waste species, T is the temperature and $\dot{\Sigma}^{\text{chem}}$ is the entropy production rate of the chemical transitions. The rest of the free energy is supplied to the mechanical transitions either as energy ($\dot{\mathcal{E}}$) or information ($RT\dot{\mathcal{I}}$) flow, where R is the gas constant. Mechanical transitions dissipate this free energy ($T\dot{\Sigma}^{\text{mech}}$), generating directional motion of the macrocycle. When $\dot{\mathcal{E}} + RT\dot{\mathcal{I}} = 0$, mechanical transitions are at equilibrium and no net mechanical displacement of the macrocycle can arise.

of concepts and relationships among energy, information, kinetics and molecular structure.

Results and discussion

A bipartite chemical reaction network for the minimalist rotary motor. The rotary motor in Fig. 1 comprises a cyclic track with two degenerate binding sites for a macrocycle¹⁴. One site is deuterium labelled to distinguish it from the other by ^1H NMR spectroscopy, although deuteration does not influence the chemical properties. Fluorenylmethoxycarbonyl (Fmoc) groups, which sterically prevent passage of the macrocycle, can be attached to hydroxy residues on the track (Fig. 2). When only one barrier is in place, macrocycle shuttling enables exchange between two co-conformers (co-conformers are structures that differ in the relative positions of the components⁵²) that have the macrocycle either proximal (p) or distal (d) to the free hydroxy group. The fuelling reaction kinetically discriminates between the two co-conformers, favouring the reaction of the distal co-conformer over the proximal co-conformer. Under basic conditions, the waste-forming reaction removes barriers without any chemical gating. The chemical gating afforded by the biased fuelling reaction, and the free energy supplied by the fuel-to-waste conversion, result in directional movement of the macrocycle around the track. The rotary motor can be represented by a chemomechanical network of reactions (Fig. 2) in which mechanical and chemical transitions are coupled, as in common models for biological molecular motors⁴⁴ and minimal Brownian motors with external dichotomous noise⁵³. Because no transitions

enable simultaneous change in the mechanical and chemical states, this network is said to be bipartite^{40,54}.

Information thermodynamic analysis. As detailed in Supplementary Section I, the rotary motor is modelled as the isothermal open bipartite chemical reaction network^{40,42,43} shown in Fig. 2. The concentrations of the six motor species, 2_{H} , 2_{D} , 1_{H}^{H} , 1_{D}^{D} , 1_{H}^{D} and 1_{D}^{H} (see Fig. 2 caption) evolve according to the rate constants of each reaction following mass-action kinetics. The system is open because the concentrations of fuel (Fmoc-Cl) and waste (HCl—liberated as part of the fuelling reaction and subsequently neutralized by KHCO_3 , which is present to produce KCl, CO_2 and H_2O —and dibenzofulvene and CO_2 produced during the waste-forming reaction) species are kept constant through addition from, or removal to, an external source. An important quantity for our analysis is the chemical potential gradient between the fuel and waste species ($\mu_{\text{Fmoc-Cl}} - \mu_{\text{HCl}} - \mu_{\text{dibenzofulvene}} - \mu_{\text{CO}_2}$) which we denote $\mu_F - \mu_W$. For any thermodynamically consistent set of parameters (Methods), the system will evolve towards a stationary state in which the concentrations of all the motor species (that is, all co-conformations and chemical states of the motor) are constant in time, as are the thermodynamic properties of the system. The entropy production rate (the entropy changes in the system and in the reservoirs per unit time), denoted $\dot{\Sigma}$, measures how far from equilibrium the motor operates and has to be non-negative by virtue of the second law of thermodynamics⁵⁵. When multiplied by temperature (T), this corresponds to the amount of free energy that is instantaneously

Box 1 | Definitions as applied to the minimalist rotary motor

Fuelling reaction	The reaction that consumes the fuel (Fmoc-Cl) and attaches an Fmoc group to the motor's track, generating HCl as waste product. Forward ($k_{+F}^{p/d}$) and backward ($k_{-F}^{p/d}$) rate constants may depend on the macrocycle being proximal (p) or distal (d) to the reaction site.
Waste-forming reaction	The reaction that removes the Fmoc group from the motor's track, generating dibenzofulvene and CO ₂ as waste products. Forward ($k_{-W}^{p/d}$) and backward ($k_{+W}^{p/d}$) rate constants may depend on the macrocycle being proximal (p) or distal (d) to the reaction site.
Chemical gating	The bias of distal over proximal rate constants in fuelling and waste-forming reactions, quantified by k_{+F}^d/k_{+F}^p (fuelling gating) and k_{-W}^d/k_{-W}^p (waste-forming gating).
Overall reaction rate	The sum of the rate constants for a particular process: for example, Fmoc addition through fuelling reactions ($k_{+F}^d + k_{+F}^p$), Fmoc removal through waste-forming reactions ($k_{-W}^p + k_{-W}^d$), mechanical shuttling ($k_{+\Delta} + k_{-\Delta}$ or $k'_{+\Delta} + k'_{-\Delta}$, where the prime symbol indicates the case when the barrier close to the deuterated station is removed to allow for the shuttling).
Current	The net rate of directional displacement of the macrocycle along the track, quantified by equation (4).
Efficiency	The free energy available to the motor to dissipate as mechanical motion compared to the free energy available from the fuel-to-waste reaction, quantified by equation (5). This definition of efficiency does not take into account the background fuel decomposition.
Power stroke	The viscoelastic, free energy-releasing, mechanical shuttling of the macrocycle along the track. The free energy released is quantified by the change in standard chemical potential due to a net mechanical displacement of the macrocycle from the station proximal to the reacting hydroxy group ($1_{Hr}^D, 1_{Hr}^H$) to the one distal ($1_{D}^D, 1_{D}^H$): $\mu_{1_{Hr}^D}^o - \mu_{1_{D}^H}^o$ and $\mu_{1_{Hr}^H}^o - \mu_{1_{D}^H}^o \neq 0$.
Kinetic asymmetry	The kinetic preference for one direction over the other in a chemomechanical cycle, embodied by the ratcheting constant: $K_r = \frac{k_{+\Delta} k'_{+\Delta} (k_{-F}^p [HCl] + k_{-W}^p)^2 ([Fmoc-Cl] k_{+F}^p + [CO_2] [DBF] k_{+W}^p)^2}{k_{-\Delta} k'_{-\Delta} ([Fmoc-Cl] k_{+F}^p + [CO_2] [DBF] k_{+W}^p)^2 (k_{-F}^d [HCl] + k_{-W}^d)^2}$, where DBF stands for dibenzofulvene. When $K_r > 1$, the motor preferentially cycles in the forward direction, as is apparent from equation (4).

dissipated by the motor ($T\dot{\Sigma}$). In a stationary state (the only type of state considered in this Article), all the free energy from the conversion of the fuel to waste is dissipated by the motor: $T\dot{\Sigma} = I_F(\mu_F - \mu_W) \geq 0$, where I_F is the rate at which the fuel is consumed by the motor (we neglect the fuel-to-waste background reaction)⁴³. This indicates that a non-null chemical potential gradient between fuel and waste species (we consider the case where $\mu_F > \mu_W$) is necessary to drive the system out of equilibrium and produce directed motion. However, this expression gives no information as to how free energy is consumed nor the amount of dissipation specifically devoted to sustain directional motion of the components.

To obtain a more in-depth understanding, we use information thermodynamics to split the free energy dissipation rate into two separately non-negative contributions (Fig. 1b), one due to the chemical transitions ($T\dot{\Sigma}^{\text{chem}}$) and the other due to the mechanical transitions ($T\dot{\Sigma}^{\text{mech}}$). This is possible because the Fmoc-motor chemical reaction network is bipartite⁴⁰:

$$T\dot{\Sigma} = \overbrace{I_F(\mu_F - \mu_W) - \dot{\mathcal{E}} - RT\dot{\mathcal{I}}}_{T\dot{\Sigma}^{\text{chem}} \geq 0} + \overbrace{\dot{\mathcal{E}} + RT\dot{\mathcal{I}}}_{T\dot{\Sigma}^{\text{mech}} \geq 0} \geq 0. \quad (1)$$

This dissection of terms in equation (1) underlines that the free energy supplied by the fuel is only partially dissipated via the chemical processes as $T\dot{\Sigma}^{\text{chem}}$, while the remaining part is transferred to the mechanical processes and dissipated as $T\dot{\Sigma}^{\text{mech}}$. This transfer of free energy is composed of a standard free energy part, denoted 'energy flow':

$$\dot{\mathcal{E}} = J(\mu_{1_{Hr}^D}^o - \mu_{1_{D}^D}^o + \mu_{1_{D}^H}^o - \mu_{1_{Hr}^H}^o), \quad (2)$$

and a mutual information part³⁸, denoted 'information flow':

$$RT\dot{\mathcal{I}} = JRT \log \frac{[1_{Hr}^D][1_{D}^H]}{[1_{D}^D][1_{Hr}^H]}, \quad (3)$$

where R is the gas constant and J is the stationary clockwise (as viewed in Fig. 2) current at which the motor operates (Box 1). This current can be expressed as

$$J = k_{+\Delta}[1_{Hr}^D] - k_{-\Delta}[1_{D}^D] = k'_{+\Delta}[1_{D}^H] - k'_{-\Delta}[1_{Hr}^H] = \Gamma(K_r - 1) \quad (4)$$

where both Γ and K_r are positive quantities (for the derivation of equation (4) and the definition of Γ , see Supplementary Section V-B4 and Supplementary equation (69)), the latter denoting the ratcheting constant (recently applied in the context of dissipative self-assembly^{17,49}), which quantifies the kinetic asymmetry of the motor^{15,16,56} (Box 1).

Under the experimental conditions in which the motor was originally operated¹⁴, the rotary motor is driven purely by information flow as the macrocycle binds with equal affinity to fumaramide stations adjacent to a hydroxy group and an Fmoc group (within the detection limits of ¹H NMR measurements), so there is no energy flow. However, when the standard chemical potentials of the distal and proximal co-conformers differ, energy flow arises according to equation (2). An example of this could arise if there was, say, a stabilizing interaction between the macrocycle and the Fmoc group, which would lead to an increase in energy of the macrocycle upon removal of the Fmoc-barrier close to it. As a consequence, the standard chemical potential would decrease, leading to the release of heat, each time net mechanical displacement occurs in the forward direction. This fits the definition of a 'power stroke'⁴⁵⁻⁴⁷ (Box 1). Therefore, according to equation (2), the energy flow accounts for the part of the fuelling free energy that contributes to destabilizing the macrocycle during chemical transitions and which is subsequently dissipated in a power stroke.

Mutual information quantifies the correlation between the two parts of a bipartite system^{40,54}, here the chemical and mechanical states. For example, when $[1_{D}^H]$ (the concentration of species 1_{D}^H) and $[1_{Hr}^D]$ are larger than $[1_{Hr}^H]$ and $[1_{D}^D]$, respectively, a correlation is present between the mechanical and chemical states: when the motor's

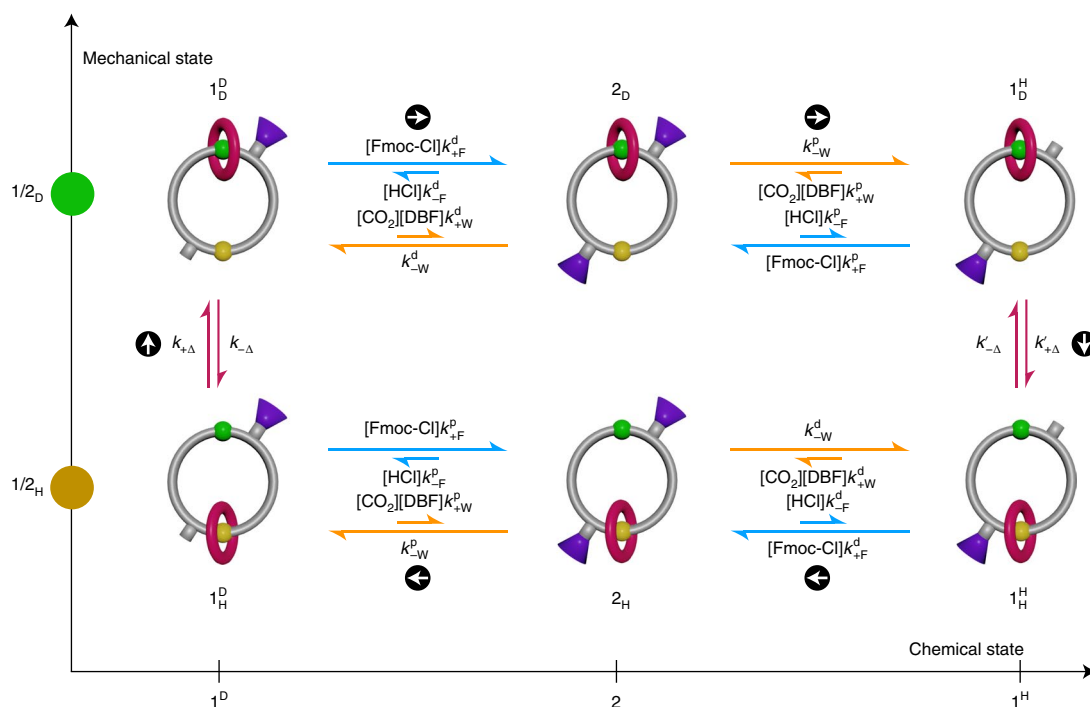


Fig. 2 | Rotary motor as an open and bipartite chemical reaction network. The state of the motor can be represented as a combination of two mechanical states (subscript D or H, expressing whether the macrocycle is binding to the deuterated or non-deuterated site on the track) and three chemical states (1^D , 2 or 1^H). The number in the chemical state shows the number of the Fmoc groups attached to the track, and the superscript H or D denotes the binding site close to the attached Fmoc group. Note that chemical states without any Fmoc group on the track are neglected as these are present in negligible concentration during machine operation due to the faster barrier formation than removal. The mechanical transitions involve displacement of the macrocycle, the rate constants for which are denoted $k_{+\Delta}$, $k_{-\Delta}$, $k'_{+\Delta}$ and $k'_{-\Delta}$. Subscript Δ indicates that they are the rate constants of mechanical transitions, which are only coupled to the thermal reservoir. Rate constants with and without the prime (') are for the interconversion between 1^D_D and 1^H_D and the interconversion between 1^D_H and 1^H_H respectively. The sign of the subscripts shows the direction of the transition (+ for the clockwise and – for the anticlockwise direction). The rate constants of the fuelling reaction and its reverse are denoted $k_{+F}^{p/d}$ and $k_{-F}^{p/d}$, respectively. The superscript shows whether the macrocycle is proximal (p) or distal (d) to the reacting hydroxy group. The rate constants of the waste-forming reaction and its reverse are denoted $k_{-W}^{p/d}$ and $k_{+W}^{p/d}$, respectively. In the experimental rotary motor, a clockwise current across the network is generated because $k_{+F}^p < k_{+F}^d$, $k_{-W}^p = k_{-W}^d$ and free energy is supplied to the motor by the fuel-to-waste conversion. The chemical gating in the fuelling reaction arises due to the steric crowding in the transition states of the proximal co-conformers (1^D_D and 1^D_H) compared to the distal co-conformers (1^D_D and 1^H_D). Arrows in circles represent the most probable (that is, most frequent) clockwise path. DBF, dibenzofulvene.

chemical state is 1^H , its mechanical state is more likely to be D than H. Similarly, when the motor's chemical state is 1^D , its mechanical state is more likely to be H than D. A concentration distribution with this kind of correlation has smaller (Shannon-like³⁸) entropy than one without ($[1^D_H]$ and $[1^H_H]$ equal to $[1^D_D]$ and $[1^H_D]$, respectively). Therefore, correlation between the mechanical and chemical states (mutual information) generates an entropic driving force for a directional current (from $[1^D_H]$ to $[1^H_H]$ and from $[1^H_D]$ to $[1^D_D]$). From the thermodynamic viewpoint, mutual information constitutes the entropic contribution of the free energy that comes from the fuel. As mutual information is constant in the stationary state, changes in mutual information due to the chemical processes and mechanical processes are balanced⁴⁰. If, as in the above situation, $\dot{I} > 0$, the chemical transitions are producing mutual information that is consumed by the mechanical ones. Therefore, according to equation (3), the information flow accounts for the part of fuelling free energy that contributes to increasing the system's mutual information during chemical transitions and that is subsequently erased by mechanical shuttling.

Regimes where the free energy supplied by the chemical to the mechanical processes is exclusively due to the information flow ($RT\dot{I}$), and thus lack any energy flow (\dot{E}), are denoted as pure Maxwell demon regimes⁴⁰.

This analysis demonstrates that the free energy supplied to the mechanical processes by the energy and information flows,

$RT\dot{I} + \dot{E}$, is the origin of net directional motion of the macrocycle around the track. Indeed, in the absence of such flow ($RT\dot{I} + \dot{E} = 0$), the mechanical transitions are at thermodynamic equilibrium ($\Sigma^{\text{mech}} = 0$), meaning a zero directional shuttling current ($J = 0$; equation (4)). From an information thermodynamics perspective, the molecular motor operates by using chemical processes to transduce the free energy supplied by the fuel into the free energy supplied to the mechanical processes. The efficiency of this transduction is the ratio of the latter to the former and is bounded between zero and one due to the non-negativity of Σ^{chem} :

$$0 \leq \eta = \frac{RT\dot{I} + \dot{E}}{I_F(\mu_F - \mu_W)} = 1 - \frac{T\Sigma^{\text{chem}}}{I_F(\mu_F - \mu_W)} \leq 1. \quad (5)$$

The efficiency of the transduction of the free energy of fuelling to the mechanical processes (η) is not directly comparable with the efficiencies usually reported for biological motors, which are often defined with respect to the work performed by the motor against an external force^{35,47,57} or as the fraction of fuel molecules that are productively consumed, on average, over an operational cycle^{50,58,59} (the latter measure of performance is computed in Supplementary Section VI for the rotary motor). In this set-up, the energy and information flows are entirely dissipated by the shuttling of the macrocycle as $T\Sigma^{\text{mech}}$. If the mechanical steps of the motor were

to work against a force (for example, if a load were attached to the macrocycle), $T\dot{\Sigma}^{\text{mech}}$ would incorporate a negative work term in addition to $RT\dot{\mathcal{I}} + \dot{\mathcal{E}}$, which could serve to define the efficiency of the energy and information flows being converted into output work, instead of just being dissipated. The energy and information flows would thus constitute the maximum work output that can be delivered by the motor. A traditional thermodynamic analysis of such a motor^{25,26,60} would exclusively focus on the efficiency with which the input free energy supplied by the fuel-to-waste chemical potential gradient is converted into output work, thus over-estimating the maximum work output as the overall free energy input $I_F(\mu_F - \mu_W)$. The present approach refines this analysis by showing how the input-to-output transduction is mediated by the free energy transfer within the motor, whose efficiency η (equation (5)) limits the maximum work output potentially deliverable by the motor. It also formally defines a thermodynamic efficiency that can be applied to motors while they perform no appreciable output work, as is the case for most of the synthetic molecular motors made so far, and can serve to compare the efficiencies of their operation.

The framework we have outlined can also be used to re-derive previous results obtained using kinetic arguments as a consequence of the second law of thermodynamics in bipartite systems. In Supplementary Section V-B, we show that the condition $T\dot{\Sigma}^{\text{mech}} = 0$ implies $K_r = 1$, whereas the condition $T\dot{\Sigma}^{\text{mech}} > 0$ implies $K_r \neq 1$, with forward movement when $K_r > 1$. This shows how the nonequilibrium thermodynamic framework, which focuses on energetic aspects quantified by the dissipation $T\dot{\Sigma}^{\text{mech}}$, is consistent with previous analysis^{15–17,49} focusing on kinetic aspects quantified by K_r , which determines the sign of the current J according to equation (4). This reiterates the effectiveness of this information thermodynamics-based approach and, again, demonstrates the usefulness of this minimalist molecular motor as a Rosetta Stone for the translation of meaning and understanding between different frameworks for describing phenomena.

Previously¹⁷, the ratcheting constant has also been related to the ability of a dissipative self-assembly system to store free energy, but this connection is only valid in an operating regime where chemical transitions are substantially faster than mechanical ones. The information thermodynamics framework offers a general understanding of dissipative chemical systems and establishes limits to the maximum work deliverable by them that are valid in any operating regime, thanks to concepts such as the efficiency η introduced in equation (5).

Design principles for molecular motors. To demonstrate the use of our framework as a design tool, we explored the effects of altering the design features of the rotary motor on its current and efficiency (Supplementary Section VI). Basing the simulations on experimentally derived parameters¹⁴ (Supplementary Section VI-A), we established that, under the experimental conditions employed, the rotary motor is driven purely by information flow as only the macrocycle distribution, rather than binding site affinity (within the detection limits of ¹H NMR measurements), is altered during operation. Under experimental conditions ([motor] = 10 mM, [Fmoc-Cl] = 30 mM, which is sustained by constant addition, [bulky catalyst] = 50 mM, [Et₃N] = 15 mM, [KHCO₃] = 200 mM, CH₂Cl₂, room temperature)¹⁴, the current was estimated to be 2.1×10^{-8} mol dm⁻³ s⁻¹, requiring an average of seven fuel molecules per cycle per motor, although only 10⁻⁶% of the free energy provided by the fuel is used to sustain the current (efficiency, $\eta = 10^{-8}$; equation (5)).

Varying the model parameters allows consideration of the effects of potential structural and chemical changes on the rotary motor (Supplementary Section VI-B1). Greater chemical gating for either the fuelling (Fig. 3a,b) or waste-forming reaction (Supplementary Section VI-B2) increases the current and efficiency by increasing information flow. The former has been achieved by increasing

the steric bulk of the barrier-formation catalyst^{14,61,62} (Fig. 4a) and the latter by catalysis of barrier removal by a proximal macrocycle (Fig. 4b)²⁴. Gating of both fuelling and waste-forming reactions (Supplementary Section VI-B3) was recently demonstrated in a rotaxane information ratchet²⁴. Inverting chemical gating is predicted to reverse the direction of the motor and could be achieved if the macrocycle activates, rather than hinders, proximal barrier formation (Fig. 4c). In the absence of a kinetic preference, or when gating from the fuelling and waste-forming reactions cancel out, the motor stops working, as this precludes information flow and hence mechanical dissipation ($\dot{\Sigma}^{\text{mech}} = 0$). These results are consistent with kinetic models where kinetic asymmetry predicts the direction of the current^{15–17}.

The relevance of power strokes in molecular machinery is contentious, as power strokes are often observed experimentally in biological molecular motors^{44–46}, but, according to analysis based on kinetic asymmetry, the magnitude of the free energy released by such conformational changes does not affect the properties of chemically driven molecular machines, such as directionality and stopping force, and cannot improve the efficiency of a motor to work against an external force⁴⁷. To reconcile differing viewpoints as to the importance of power strokes in molecular machines, we attempted to use our framework to understand the ways in which a power stroke can affect a molecular motor while staying kinetically and thermodynamically consistent. Power strokes can be used to induce energy flow (equation (2)) and, in principle, could be added to the rotary motor by introducing attractive interactions between the barrier and the macrocycle (Fig. 4d), stabilizing distal co-conformers 1_D^D and 1_H^H or, by adding repulsive interactions between the free barrier site and the macrocycle (Fig. 4e), destabilizing proximal co-conformers 1_H^D and 1_D^H. Our simulations show that power strokes can change the magnitude of the current and efficiency of internal energy transduction (as defined in equations (4) and (5), respectively), despite kinetic asymmetry remaining unaltered (Methods and Supplementary Section VI-B4). This is because, while not altering K_r , power strokes can still increase the value of Γ in equation (4), reflecting their ability to favour forward cycles by inducing energy flow out of equilibrium (Fig. 3c,d). However, power strokes cannot drive directional motion in the absence of kinetic asymmetry ($RT\dot{\mathcal{I}} = -\dot{\mathcal{E}}$ when $K_r = 1$), nor can they invert directionality while the kinetic asymmetry remains constant. In these simulations, altering the power strokes and kinetic asymmetry together (Supplementary Section VI-B5) gave the greatest simulated efficiency (up to ~1%), suggesting that modifying both aspects may be important for optimizing the design of synthetic molecular motors. We note that improvements that occur through power strokes arise from the induced energy flow rather than from any special role for the energetically downhill nature of the power stroke in determining the motor's behaviour. As our analysis shows, a prerequisite for the motor's operation is the availability of free energy to dissipate through mechanical motion. Therefore, any design feature that enhances free energy transduction from the chemical to the mechanical transitions could equally well foster improvements in performance.

If power strokes cancel out over the motor cycle, then no net energy flow is introduced: free energy gained from one mechanical transition is lost in the other (Methods and Supplementary Section VI-B6). This could be realized in molecular form by using non-degenerate binding sites, for example, changing one fumaramide site to a more weakly binding succinamide unit (Figs. 3e,f and 4f), with a typical difference in binding energy of 23 kJ mol⁻¹ (pink + symbol, Fig. 3e,f, equilibrium distribution > 99:1)⁶³ under experimental conditions similar to those used for rotary motor operation. Kinetic asymmetry remains unaltered in this scenario and, correspondingly, the direction of the motor could not be inverted in the simulation^{15–17}. However, our analysis suggests that such a

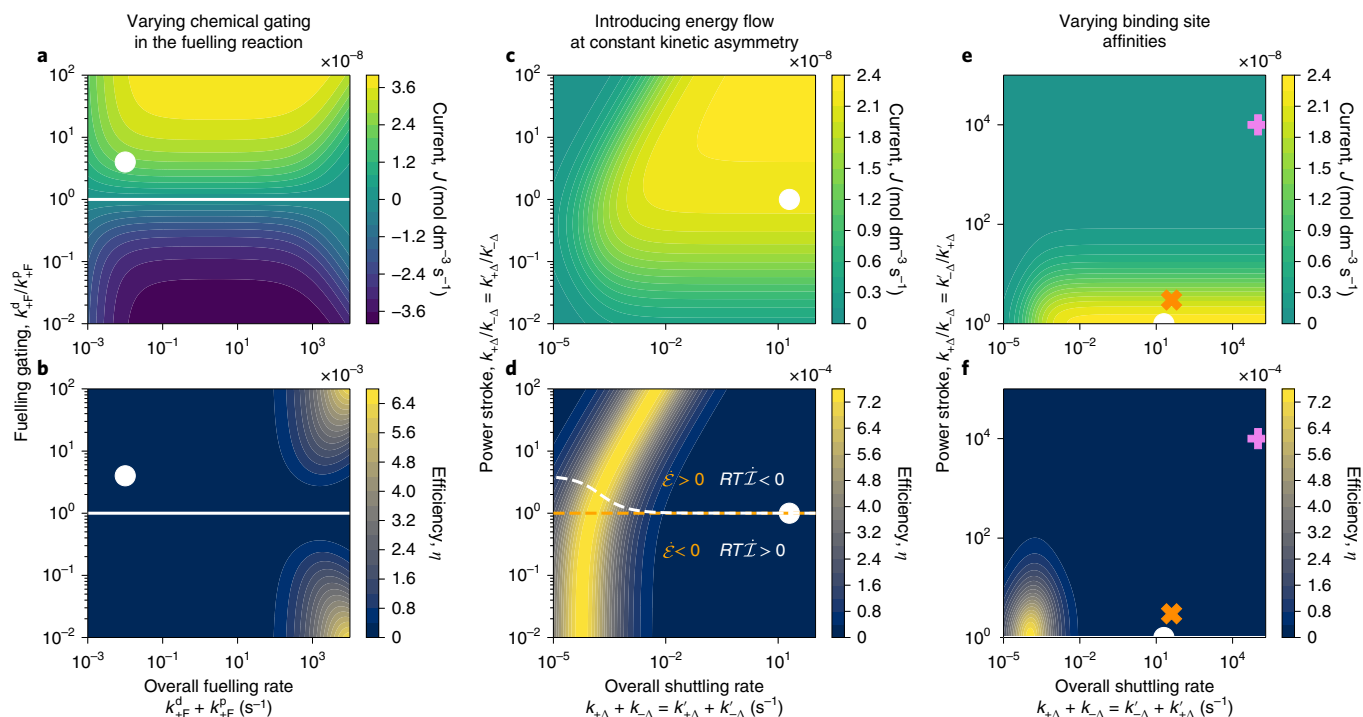


Fig. 3 | Numerical simulations of different molecular motor design modifications. a–f, Simulated current (J , top) and efficiency (η , bottom). Solid white lines indicate the regions of null current (and efficiency). White circles indicate the approximate parameters of the experimental rotary motor, giving a current of $2.1 \times 10^{-8} \text{ mol dm}^{-3} \text{ s}^{-1}$ and efficiency of $10^{-6}\%$. Directional motion in the experimental rotary motor arises because distal fuelling reactions are approximately four times faster than proximal fuelling reactions (chemical gating, $k_{+F}^d/k_{+F}^p \approx 4$) due to the steric hindrance of the proximal transition state, whereas there was no chemical gating in waste-forming reactions ($k_{-W}^d/k_{-W}^p \approx 1$). In **a** and **b**, the overall rate of the fuelling reactions (x axis) and the fuelling gating (y axis) are varied. In **c** and **d**, the overall shuttling rate (x axis) and the ratio of forward and backward shuttling rates (indicative of power stroke magnitude; Methods) are varied. Power strokes are introduced in a way that induces energy flow ($\dot{\mathcal{E}}$) while keeping K_r constant. Below the dashed white lines, information flow ($RT\dot{I}$) is positive and thus contributes to mechanical dissipation $\dot{\Sigma}^{\text{mech}}$. Above these lines, information flow is negative, requiring the motor to be driven by positive energy flow ($\dot{\mathcal{E}} > 0$). Negative energy flow is found below the orange dashed line. In **e** and **f**, the overall shuttling rate (x axis) and ratio of shuttling rate between stations (y axis) are varied. When this ratio diverges from one (indicating non-degenerate binding sites; Methods), both the current and efficiency drop. The pink + symbol indicates the position expected by changing one fumaramide group to a succinamide⁶³. The orange x symbol indicates a position with a 3:1 difference in binding site affinity for the macrocycle.

change would be sufficient to effectively stall the motor, if operated under the original experimental conditions, despite the unchanged kinetic asymmetry. The simulations predict that, with power strokes cancelling out, any change from degenerate binding sites lowers the current and efficiency, although a smaller difference would leave the motor functional, albeit less effective. A 3:1 bias⁶¹ (orange x symbol, Fig. 3e,f) is predicted to reduce the current by $\sim 20\%$, rendering a design with non-degenerate binding sites plausible but less effective than a motor with binding sites of equal affinities.

In all of the cases considered, the highest efficiencies are predicted for when the rates of all the forward processes are approximately equal, leading to a cycle with no single rate-limiting step. Rate-limiting mechanical steps promote futile cycles, in which fuel is consumed without taking a forward step, as the unfavourable fuelling reaction is kinetically favoured over shuttling, decreasing both the current and efficiency. Rate-limiting chemical reactions result in lower thermodynamic efficiency but do not reduce the current or substantially change the fuel consumption per cycle. This is because relatively fast shuttling hinders the generation of a concentration bias (relative to mechanical equilibrium), which decreases the information flow in the steady state. The strong dependence of efficiency on shuttling rate indicates that, like macroscopic engines, the efficiency of a motor will be dependent on the load against which it is working³⁵. As a consequence, instead of discussing a generic efficiency of a molecular motor, it is more proper to discuss the

efficiency to work in a specific range of force/attached load. To use molecular machines most efficiently, they must either be tailored to the job they perform—such as using diesel engines for heavy loads—or they must use the equivalent of gears for macroscopic engines, to ensure they are working under optimal conditions.

Conclusions

Information thermodynamics-based analysis of a minimalist autonomous chemically driven molecular motor shows how information and energy flow, the two components of free energy transfer from chemical to mechanical transitions, enable the generation of directional motion from free energy supplied by a chemical fuel. The experimental rotary motor is a pure ‘chemical Maxwell’s demon’, as information flow is the sole driving force. However, energy flow could potentially be introduced using power strokes, one of several design variations explored using our model. The predicted effect of energy flow is in line with observations made in biological motors⁴⁴ and contributes to the ongoing debate regarding the role of power strokes in molecular motors^{46–48}. Information thermodynamics confirms that, in line with kinetic analysis, power strokes do not affect some key properties of chemically driven molecular machines, such as directionality. However, the magnitude of power strokes is able to affect the magnitude of the current (how fast the motor components rotate), the efficiency in terms of how free energy is dissipated, and the number of fuel molecules consumed per cycle. However,

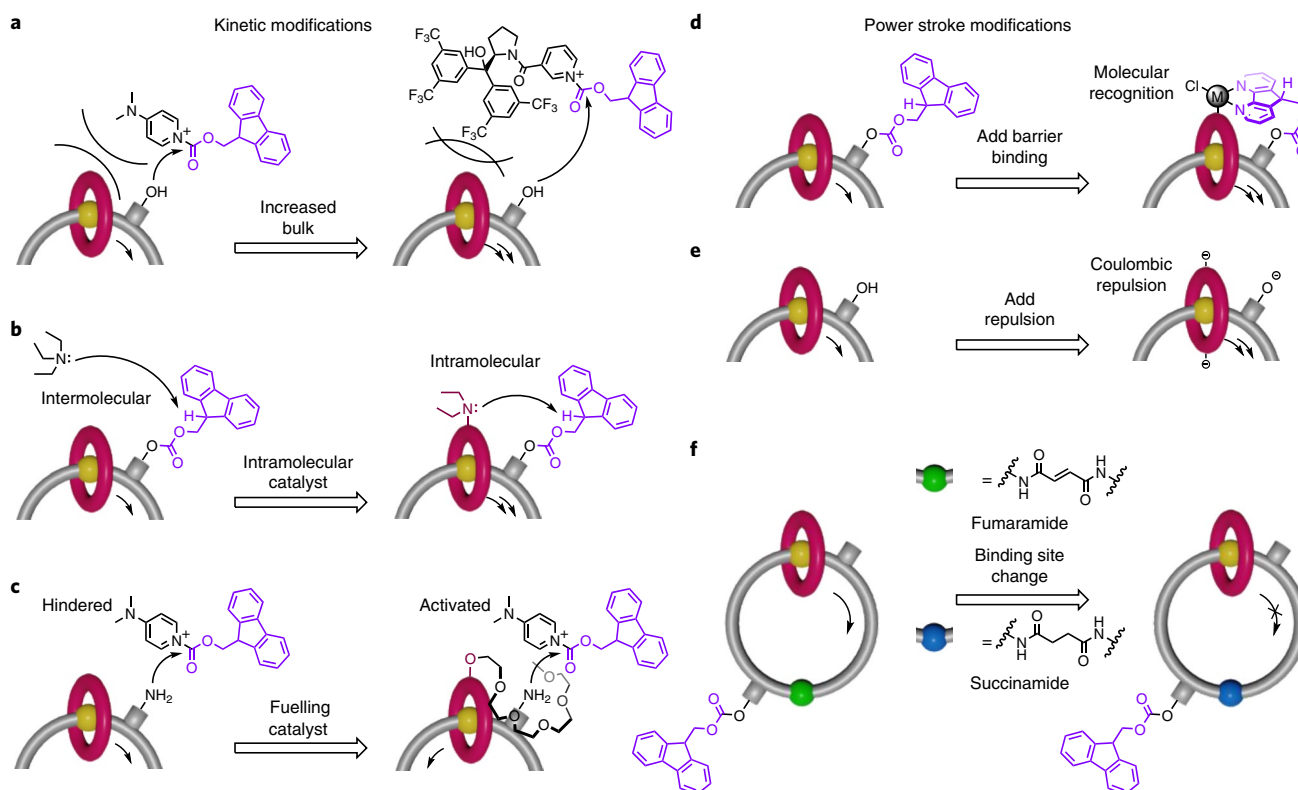


Fig. 4 | Potential ways of achieving different molecular motor design modifications. **a**, Increasing the steric bulk of the catalyst could be used to increase the fuelling chemical gating by slowing the undesired proximal reaction. **b**, Addition of a catalyst for the waste-forming reaction to the macrocycle could increase waste-formation chemical gating by increasing the rate of the proximal waste-forming reaction. **c**, Addition of a catalyst to the macrocycle that accelerates rather than hinders proximal barrier formation could negate chemical gating arising from steric hindrance, leading to an inversion of the directionality of the motor. **d**, Allowing the formation of a complex that binds the barrier and the macrocycle could allow the introduction of a power stroke by stabilizing the co-conformer in which the macrocycle is adjacent to the barrier. **e**, Conversely, a power stroke might be introduced by destabilizing the co-conformer with the macrocycle adjacent to the site with no barrier, for example, by the introduction of coulombic repulsion. **f**, Swapping one fumaramide binding site for a succinamide should, in principle, be sufficient to stall the rotary motor under the previous experimental operating conditions.

these results should not be misinterpreted as supporting a special importance of power strokes compared to other processes in the chemomechanical cycle. The information thermodynamics framework used in this Article should be generally applicable to other types of synthetic molecular machine, such as non-autonomous^{10,12} and light-driven^{8,9,11} motors, providing a quantitative basis through which to compare molecular machine designs. Additionally, it could, in principle, be extended to other types of (supra)molecular system (such as dissipative self-assembly¹⁷) powered by chemical engines⁶⁴. We have uncovered substantial roles for ‘energy flow’ and ‘information flow’ in the mechanism of the transduction of free energy from chemical reactions by molecular machinery, although the exact nature of the connection of energy and information flow to energy and information ratcheting^{1,65} remains to be clarified. The minimalist autonomous chemically driven molecular motor acts as a Rosetta Stone for relating energy, information, kinetics and molecular structure by aiding the translation of concepts and relationships between the ‘languages’ (that is, frameworks) of chemical kinetics, thermodynamics and chemical reactions.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41557-022-00899-z>.

Received: 4 July 2021; Accepted: 28 January 2022;
Published online: 17 March 2022

References

- Chatterjee, M. N., Kay, E. R. & Leigh, D. A. Beyond switches: ratcheting a particle energetically uphill with a compartmentalized molecular machine. *J. Am. Chem. Soc.* **128**, 4058–4073 (2006).
- Coskun, A., Banaszak, M., Astumian, R. D., Stoddart, J. F. & Grzybowski, B. A. Great expectations: can artificial molecular machines deliver on their promise? *Chem. Soc. Rev.* **41**, 19–30 (2012).
- Erbas-Cakmak, S., Leigh, D. A., McTernan, C. T. & Nussbaumer, A. L. Artificial molecular machines. *Chem. Rev.* **115**, 10081–10206 (2015).
- Qiu, Y., Feng, Y., Guo, Q.-H., Astumian, R. D. & Stoddart, J. F. Pumps through the ages. *Chem* **6**, 1952–1977 (2020).
- Aprahamian, I. The future of molecular machines. *ACS Cent. Sci.* **6**, 347–358 (2020).
- Howard, J. *Mechanics of Motor Proteins and the Cytoskeleton* (Oxford Univ. Press, 2001).
- Astumian, R. D., Mukherjee, S. & Warshel, A. The physics and physical chemistry of molecular machines. *Chemphyschem* **17**, 1719–1741 (2016).
- Koumura, N., Zijlstra, R. W. J., van Delden, R. A., Harada, N. & Feringa, B. L. Light-driven monodirectional molecular rotor. *Nature* **401**, 152–155 (1999).
- Serrelli, V., Lee, C.-F., Kay, E. R. & Leigh, D. A. A molecular information ratchet. *Nature* **445**, 523–527 (2007).
- Cheng, C. et al. An artificial molecular pump. *Nat. Nanotechnol.* **10**, 547–553 (2015).
- Ragazzon, G., Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Light-powered autonomous and directional molecular motion of a dissipative self-assembling system. *Nat. Nanotechnol.* **10**, 70–75 (2015).
- Erbas-Cakmak, S. et al. Rotary and linear molecular motors driven by pulses of a chemical fuel. *Science* **358**, 340–343 (2017).

13. Amano, S., Fielden, S. D. P. & Leigh, D. A. A catalysis-driven artificial molecular pump. *Nature* **594**, 529–534 (2021).
14. Wilson, M. R. et al. An autonomous chemically fuelled small-molecule motor. *Nature* **534**, 235–240 (2016).
15. Astumian, R. D. How molecular motors work—insights from the molecular machinist's toolbox: the Nobel Prize in Chemistry 2016. *Chem. Sci.* **8**, 840–845 (2017).
16. Astumian, R. D. Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet. *Nat. Commun.* **10**, 3837 (2019).
17. Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotechnol.* **13**, 882–889 (2018).
18. Esposito, M. Open questions on nonequilibrium thermodynamics of chemical reaction networks. *Commun. Chem.* **3**, 107 (2020).
19. Kelly, T. R., De Silva, H. & Silva, R. A. Unidirectional rotary motion in a molecular system. *Nature* **401**, 150–152 (1999).
20. Mock, W. L. & Ochwat, K. J. Theory and example of a small-molecule motor. *J. Phys. Org. Chem.* **16**, 175–182 (2003).
21. Fletcher, S. P., Dumur, F., Pollard, M. M. & Feringa, B. L. A reversible, unidirectional molecular rotary motor driven by chemical energy. *Science* **310**, 80–82 (2005).
22. Collins, B. S. L., Kistemaker, J. C. M., Otten, E. & Feringa, B. L. A chemically powered unidirectional rotary molecular motor based on a palladium redox cycle. *Nat. Chem.* **8**, 860–866 (2016).
23. Zhang, Y. et al. A chemically driven rotary molecular motor based on reversible lactone formation with perfect unidirectionality. *Chem* **6**, 2420–2429 (2020).
24. Borsley, S., Leigh, D. A. & Roberts, B. M. W. A doubly kinetically-gated information ratchet autonomously driven by carbodiimide hydration. *J. Am. Chem. Soc.* **143**, 4414–4420 (2021).
25. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **75**, 126001 (2012).
26. Parrondo, J. M. R. & de Cisneros, B. J. Energetics of Brownian motors: a review. *Appl. Phys. A* **75**, 179–191 (2002).
27. Esposito, M., Lindenberg, K. & Van den Broeck, C. Universality of efficiency at maximum power. *Phys. Rev. Lett.* **102**, 130602 (2009).
28. Benenti, G., Casati, G., Saito, K. & Whitney, R. S. Fundamental aspects of steady-state conversion of heat to work at the nanoscale. *Phys. Rep.* **694**, 1–124 (2017).
29. Jarzynski, C. Equalities and inequalities: irreversibility and the second law of thermodynamics at the nanoscale. *Annu. Rev. Condens. Matter Phys.* **2**, 329–351 (2011).
30. Peliti, L. & Pigolotti, S. *Stochastic Thermodynamics: An Introduction* (Princeton Univ. Press, 2021).
31. Andrieux, D. & Gaspard, P. Fluctuation theorems and the nonequilibrium thermodynamics of molecular motors. *Phys. Rev. E* **74**, 011906 (2006).
32. Lipowsky, R. & Liepelt, S. Chemomechanical coupling of molecular motors: thermodynamics, network representations, and balance conditions. *J. Stat. Phys.* **130**, 39–67 (2008).
33. Toyabe, S. et al. Nonequilibrium energetics of a single F_1 -ATPase molecule. *Phys. Rev. Lett.* **104**, 198103 (2010).
34. Ariga, T., Tomishige, M. & Mizuno, D. Nonequilibrium energetics of molecular motor kinesin. *Phys. Rev. Lett.* **121**, 218101 (2018).
35. Brown, A. I. & Sivak, D. A. Theory of nonequilibrium free energy transduction by molecular machines. *Chem. Rev.* **120**, 434–459 (2020).
36. Ciliberto, S. Experiments in stochastic thermodynamics: short history and perspectives. *Phys. Rev. X* **7**, 021051 (2017).
37. McGrath, T., Jones, N. S., ten Wolde, P. R. & Ouldrige, T. E. Biochemical machines for the interconversion of mutual information and work. *Phys. Rev. Lett.* **118**, 028101 (2017).
38. Cover, T. M. & Thomas, J. A. *Elements of Information Theory* (Wiley, 2012).
39. Parrondo, J. M. R., Horowitz, J. M. & Sagawa, T. Thermodynamics of information. *Nat. Phys.* **11**, 131–139 (2015).
40. Horowitz, J. M. & Esposito, M. Thermodynamics with continuous information flow. *Phys. Rev. X* **4**, 031015 (2014).
41. Leff, H. S. & Rex, A. F. (eds) *Maxwell's Demon: Entropy, Information, Computing* (Princeton Univ. Press, 1990).
42. Rao, R. & Esposito, M. Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. *Phys. Rev. X* **6**, 041064 (2016).
43. Penocchio, E., Rao, R. & Esposito, M. Thermodynamic efficiency in dissipative chemistry. *Nat. Commun.* **10**, 3865 (2019).
44. Wagoner, J. A. & Dill, K. A. Mechanisms for achieving high speed and efficiency in biomolecular machines. *Proc. Natl Acad. Sci. USA* **116**, 5902–5907 (2019).
45. Howard, J. Protein power strokes. *Curr. Biol.* **16**, R517–R519 (2006).
46. Hwang, W. & Karplus, M. Structural basis for power stroke vs. Brownian ratchet mechanisms of motor proteins. *Proc. Natl Acad. Sci. USA* **116**, 19777–19785 (2019).
47. Astumian, R. D. Irrelevance of the power stroke for the directionality, stopping force, and optimal efficiency of chemically driven molecular machines. *Biophys. J.* **108**, 291–303 (2015).
48. Pezzato, C., Cheng, C., Stoddart, J. F. & Astumian, R. D. Mastering the non-equilibrium assembly and operation of molecular machines. *Chem. Soc. Rev.* **46**, 5491–5507 (2017).
49. Das, K., Gabrielli, L. & Prins, L. J. Chemically-fueled self-assembly in biology and chemistry. *Angew. Chem. Int. Ed.* **60**, 20120–20143 (2021).
50. Albaugh, A. & Gingrich, T. R. Simulating a chemically-fueled molecular motor with nonequilibrium molecular dynamics. Preprint at <https://doi.org/10.48550/arXiv.2102.06298> (2021).
51. Horowitz, J. M., Sagawa, T. & Parrondo, J. M. R. Imitating chemical motors with optimal information motors. *Phys. Rev. Lett.* **111**, 010602 (2013).
52. Fyfe, M. C. T. et al. Anion-assisted self-assembly. *Angew. Chem. Int. Ed.* **36**, 2068–2070 (1997).
53. Astumian, R. D. Design principles for Brownian molecular machines: how to swim in molasses and walk in a hurricane. *Phys. Chem. Chem. Phys.* **9**, 5067–5083 (2007).
54. Hartich, D., Barato, A. C. & Seifert, U. Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a Maxwell's demon interpretation. *J. Stat. Mech. Theory Exp.* **2014**, P02016 (2014).
55. Kondepudi, D. K. & Prigogine, I. *Modern Thermodynamics: From Heat Engines to Dissipative Structures* (Wiley, 2015).
56. Astumian, R. & Bier, M. Mechanochemical coupling of the motion of molecular motors to ATP hydrolysis. *Biophys. J.* **70**, 637–653 (1996).
57. Kinoshita, K. Jr., Yasuda, R., Noji, H. & Adachi, K. A rotary molecular motor that can work at near 100% efficiency. *Philos. Trans. R. Soc. Lond. B* **355**, 473–489 (2000).
58. Watt, I. N., Montgomery, M. G., Runswick, M. J., Leslie, A. G. W. & Walker, J. E. Bioenergetic cost of making an adenosine triphosphate molecule in animal mitochondria. *Proc. Natl Acad. Sci. USA* **107**, 16823–16827 (2010).
59. Petersen, J., Förster, K., Turina, P. & Gräber, P. Comparison of the H^+ /ATP ratios of the H^+ -ATP synthases from yeast and from chloroplast. *Proc. Natl Acad. Sci. USA* **109**, 11150–11155 (2012).
60. Hill, T. L. *Free Energy Transduction in Biology* (Academic Press, 1977).
61. Alvarez-Pérez, M., Goldup, S. M., Leigh, D. A. & Slawin, A. M. Z. A chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **130**, 1836–1838 (2008).
62. Carlone, A., Goldup, S. M., Lebrasseur, N., Leigh, D. A. & Wilson, A. A three-compartment chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **134**, 8321–8323 (2012).
63. Lussis, P. et al. A single synthetic small molecule that generates force against a load. *Nat. Nanotechnol.* **6**, 553–557 (2011).
64. Amano, S., Borsley, S., Leigh, D. A. & Sun, Z. Chemical engines: driving systems away from equilibrium through catalyst reaction cycles. *Nat. Nanotechnol.* **16**, 1057–1067 (2021).
65. Kay, E. R., Leigh, D. A. & Zerbetto, F. Synthetic molecular motors and mechanical machines. *Angew. Chem. Int. Ed.* **46**, 72–191 (2007).
66. Blackmond, D. G. 'If pigs could fly' chemistry: a tutorial on the principle of microscopic reversibility. *Angew. Chem. Int. Ed.* **48**, 2648–2654 (2009).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Methods

Wegscheider's conditions. To ensure that the system reaches thermodynamic equilibrium (detailed balance), when there is no chemical potential gradient between fuel and waste species ($\mu_F - \mu_W = 0$), the rate constants must satisfy Wegscheider's conditions⁴². These are equivalent to the constraints on the rate constants imposed in previous kinetic analyses^{16,48} and dictate that the product of the forward rate constants along each independent cyclic pathway of reactions in the network, with neither net consumption nor net production of fuel or waste species, must equal the product of the corresponding backward rate constants:

$$k_{-F}^p k_{+\Delta}^d k_{+F}^p k'_{+\Delta}^d k_{+F}^d = k_{+F}^p k_{-\Delta}^d k_{-F}^p k'_{+\Delta}^d k_{-\Delta}^d \quad (6)$$

$$k_{-W}^p k_{+\Delta}^d k_{+W}^p k'_{+\Delta}^d k_{+W}^d = k_{+W}^p k_{-\Delta}^d k_{-W}^p k'_{+\Delta}^d k_{-\Delta}^d \quad (7)$$

See Supplementary Section III for the derivation. These conditions were always imposed in numerical simulations to guarantee thermodynamic consistency⁶⁶.

Local detailed balance. In Fig. 3, variations in the power stroke magnitude have been related to changes in the ratio of shutting rate constants by virtue of the so-called 'principle of local detailed balance' (see below), which relates the log-ratio of forward and backward rate constants of a single chemical reaction to the difference in standard chemical potentials between its reagents and products⁴² (Supplementary equation (33)). For example, it implies the relation

$$RT \log \left(\frac{k_{+\Delta} k'_{+\Delta}}{k_{-\Delta} k'_{-\Delta}} \right) = \mu_{\text{H}}^{\circ} - \mu_{\text{D}}^{\circ} + \mu_{\text{D}}^{\circ} - \mu_{\text{H}}^{\circ} \quad (8)$$

which was employed in the numerical simulations.

In addition, Wegscheider's conditions (equations (6) and (7)) imply that a variation in the power stroke magnitude must always be compensated by a variation in the fuelling and waste-forming rate constants, because the following constraint (equation (9)) must always hold for thermodynamic consistency:

$$\frac{k_{+\Delta}}{k_{-\Delta}} \times \frac{k'_{+\Delta}}{k'_{-\Delta}} = \frac{k_{+F}^p k_{-F}^d}{k_{-F}^p k_{+F}^d} \times \frac{k_{+W}^p k_{-W}^d}{k_{-W}^p k_{+W}^d} \quad (9)$$

When energy flow-inducing power strokes were introduced in Fig. 3c,d, the constraint in equation (9) was imposed by changing rate constants k_{+W}^p and k_{+F}^p according to variations in the shuttling rate constants. By doing so, kinetic asymmetry (K_c) is not altered during the simulation, but the magnitude of the current in equation (4) can still change by virtue of alterations in the value of the positive factor Γ (for its mathematical expression, see Supplementary equation (69)). Instead, in Supplementary Fig. 5, rate constants k_{+W}^p and k_{-F}^p were changed to vary the energy flow and K_c together. Experimentally, this could correspond to introducing an interaction between the macrocycle and the Fmoc group that affects (Supplementary Fig. 5) or not (Fig. 3c,d) the transition state of the proximal fuelling reaction, without affecting the transition state of the proximal waste-forming reaction. Note that, when binding affinities are modified as in Fig. 3e,f, the left-hand side of equation (9) stays constant and the constraint is automatically satisfied.

We end by noting that the terminology 'local detailed balance' comes from statistical physics⁶⁷, where it has become the central concept to formulate thermodynamically consistent dynamics^{68,69}. Its chemical counterpart⁴² (Supplementary equation (33)) is fully equivalent to the usual conditions imposed

on the rate constants to ensure that microscopic reversibility holds⁴⁷ (see, for example, equation (5) in ref. 16). We note that use of the term 'local detailed balance' in this context is considered contentious by some¹⁶. For a more detailed discussion see Supplementary Section III.

Data availability

All data needed to reproduce the numerical results are reported in the Supplementary Information.

Code availability

The code that generated the plots is available at the following link: gitlab.com/emanuele.penocchio/infothermorotmot

References

- Katz, S., Lebowitz, J. L. & Spohn, H. Phase transitions in stationary nonequilibrium states of model lattice systems. *Phys. Rev. B* **28**, 1655–1658 (1983).
- Esposito, M. Stochastic thermodynamics under coarse graining. *Phys. Rev. E* **85**, 041125 (2012).
- Maes, C. Local detailed balance. *SciPost Phys. Lect. Notes* **32**, <https://doi.org/10.48550/arXiv.2102.06298> (2021).

Acknowledgements

We acknowledge support from the European Research Council (ERC Consolidator grant no. 681456 to M.E. and funding to E.P.; ERC Advanced grant no. 786630 to D.A.L.), the FQXi Foundation, project 'Information as a fuel in colloids and superconducting quantum circuits' (grant no. FQXi-IAF19-05 to M.E.), the Engineering and Physical Sciences Research Council (EPSRC; grant no. EP/P027067/1 to D.A.L.), the Deutsche Forschungsgemeinschaft (a postdoctoral fellowship to E.K.) and the University of Manchester and EPSRC for PhD studentships to S.A. and B.M.W.R. D.A.L. is a Royal Society Research Professor. We thank R. D. Astumian for valuable discussions regarding the science in this study as well as robust debate regarding the use of the term 'local detailed balance' within the stochastic thermodynamics community.

Author contributions

S.A., B.M.W.R. and E.K. proposed the collaboration. E.P. developed the theoretical model. S.A., B.M.W.R., E.P. and E.K. carried out the theoretical analysis and simulations. D.A.L. and M.E. directed the research. All authors contributed to the analysis of the results and the writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41557-022-00899-z>.

Correspondence and requests for materials should be addressed to David A. Leigh or Emanuele Penocchio.

Peer review information *Nature Chemistry* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Supplementary information

**Insights from an information
thermodynamics analysis of a synthetic
molecular motor**

In the format provided by the
authors and unedited

Supplementary information for *Insights from an information thermodynamics analysis of a synthetic molecular motor*

Shuntaro Amano,¹ Massimiliano Esposito,² Elisabeth Kreidt,¹ David
A. Leigh*,¹ Emanuele Penocchio*,² and Benjamin M. W. Roberts¹

¹*Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom*

²*Complex Systems and Statistical Mechanics, Department of Physics and Materials
Science, University of Luxembourg, L-1511 Luxembourg City, G.D. Luxembourg*

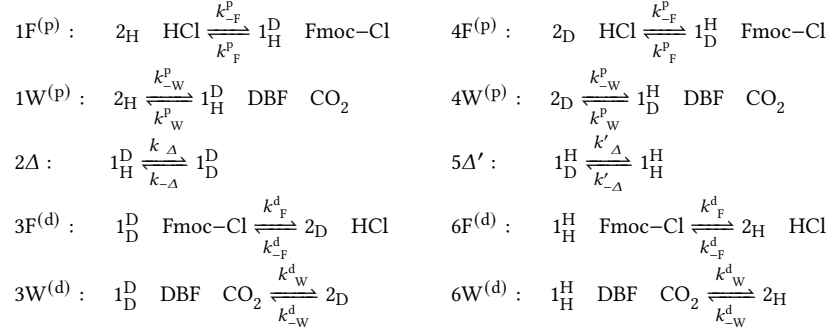
*email: david.leigh@manchester.ac.uk, emanuele.penocchio@uni.lu

CONTENTS

I. The model	2
A. Coarse-graining	2
II. Dynamics	4
A. Topological properties	5
B. Stationary state	6
III. Thermodynamics	7
A. Steady state entropy production and J^{ss}/I_F^{ss}	9
IV. Information Thermodynamics Perspective	9
V. Thermodynamic constraints on the stationary state dynamics and connection with kinetic asymmetry: the rotary motor as a Rosetta Stone for kinetic and thermodynamic analysis	12
A. Condition for directional J^{ss}	12
B. Connection to kinetic asymmetry and K_r	12
1. Equilibrium condition	13
2. Proof of the implication $\dot{\Sigma}^{\text{mech}} > 0 \rightarrow K_r \neq 1$	13
3. Special proof for the experimental case	14
4. On the magnitude of the stationary state current J^{ss}	14
VI. Numerical Simulations	15
A. Parameters	15
B. Supplementary analysis of the numerical simulations	15
1. Chemical gating of the fueling reaction	15
2. Chemical gating of the waste-forming reaction	17
3. Further exploration of chemical gating	17
4. Introducing energy flow with power strokes without varying kinetic asymmetry	18
5. Introducing energy flow with a power stroke together with varying kinetic asymmetry	19
6. Causing power strokes to cancel out	20
7. Conclusion	21
References	21

I. THE MODEL

As stated in the main text, we treat the autonomous chemically-fueled molecular motor¹ as an open bipartite chemical reaction network. We consider an isothermal, isobaric, and well-stirred ideal dilute solution containing species undergoing the 10 reactions represented in Fig. 2 of the main text, which we collect in the vector $\boldsymbol{\rho} = \{1F^{(p)}, 1W^{(p)}, 2\Delta, 3F^{(d)}, 3W^{(d)}, 4F^{(p)}, 4W^{(p)}, 5\Delta', 6F^{(d)}, 6W^{(d)}\}$:



Here and in the rest of the SI, the abbreviation DBF stands for dibenzofulvene. The symbols Δ and Δ' label thermal shuttlings (*i.e.*, mechanical processes) which change the mechanical state of the system, while all the other reactions constitute the chemical processes changing the chemical state of the system. Note that the latter are not elementary reactions, as detailed in the next section. The system is said to be open because during the experiment fuel ($F = \{\text{Fmoc-Cl}\}$) and waste ($W = \{\text{HCl, DBF, CO}_2\}$) species undergo other (not included in $\boldsymbol{\rho}$) processes: Fmoc-Cl is constantly added to the solution to compensate for its depletion, CO_2 exits the system as a gas, and HCl is readily buffered by the presence of KHCO_3 in the solution. With the exception of DBF, whose accumulation does not affect the rest of the system as it is unreactive under the experimental conditions, the concentrations of fuel and waste species are kept approximately constant by external processes. This is modeled with the concept of *chemostats*: large chemical reservoirs in contact with the system which fix the chemical potentials (and thus the concentrations in the ideal dilute solution setting) of F and W species in the system. These species are therefore said to be chemostatted.

In the experimental case, the following symmetries between rate constants hold:

$$k_{\Delta} = k_{-\Delta} = k'_{\Delta} = k'_{-\Delta} := k_{\Delta}, \quad (1a)$$

$$k_W^p = k_{-W}^d := k_W \quad \text{and} \quad k_{-W}^p = k_W^d := k_{-W}. \quad (1b)$$

A. Coarse-graining

Consider the fueling reaction pathway, catalyzed by a pyridine-based nucleophilic catalyst (denoted Cat):



which has the net effect of adding one Fmoc group to the motor. In principle, we could explicitly include the above steps in our analysis and this wouldn't change the main outcomes of the work, thus uselessly complicating the mathematical treatment. For the sake of simplifying our model without losing its relevant thermodynamic features, we then consider a coarse-grained version of the fueling process based on the assumption that k_{2F} and k_{3F} are much larger than k_{-2F} and k_{-3F} . In addition, given that in the original experiment Fmoc-Cl is continuously added to the system, the catalyst Cat, whose total amount is denoted by L_{Cat} can be assumed to be saturated by the Fmoc-Cl ($[\text{Fmoc-Cat} \cdot \text{Cl}] \approx L_{\text{Cat}}$). This is an additional level of control on the fueling process, since the availability of Fmoc-Cl to react with the motor

species is regulated by the amount of catalyst in the system. As a consequence of these assumptions, scheme (2) can be coarse-grained via a standard steady-state approximation of the intermediates to get²:



where we defined the rate constants

$$k_F = \Omega \cdot k_{1F} k_{2F} k_{3F}, \quad k_{-F} = \Omega \cdot k_{-1F} k_{-2F} k_{-3F}, \quad (4)$$

with the kinetic factor

$$\Omega = \frac{L_{\text{Cat}}}{[\text{Fmoc-Cl}] k_{1F} k_{3F} [\text{HCl}] [\text{Fmoc-Cl}] k_{1F} k_{-2F} [\text{HCl}] [\text{Fmoc-O-R}] k_{-2F} k_{-3F}} \approx \frac{L_{\text{Cat}}}{[\text{Fmoc-Cl}] k_{1F} k_{3F}}. \quad (5)$$

The final approximation highlights that, in the experimental conditions, the fueling rate is strongly determined by $[\text{Fmoc-Cat} \cdot \text{Cl}] \approx L_{\text{Cat}}$, regardless the actual value of $[\text{Fmoc-Cl}]$. Indeed, we could also have considered the species $\text{Fmoc-Cat} \cdot \text{Cl}$ as the actual chemostatted fuel without altering the results of our analysis. However, the way we proceeded here is more general and equation (3) holds also for situations where the catalysts is not saturated³.

Similar considerations apply to the waste-forming reaction pathway catalyzed by triethylamine (NET_3), which has the net effect of removing one Fmoc group from the track. Let us consider the following reaction pathway:



By considering the forward processes to be faster than the backward ones, here the coarse-graining of the intermediates through a stationary state approximation leads to



where

$$k_{-W} = \Psi \cdot k_{1W} k_{2W} k_{3W} k_{4W}, \quad k_W = \Psi \cdot k_{-1W} k_{-2W} k_{-3W} k_{-4W}, \quad (8)$$

with kinetic factor

$$\Psi = \frac{L_{\text{NET}_3}}{k_{2W} k_{3W} k_{4W} k_{-1W} (k_{3W} k_{4W} [\text{DBF}] k_{-2W} ([\text{CO}_2] k_{-3W} k_{4W}))} \approx \frac{L_{\text{NET}_3}}{k_{3W} k_{4W} (k_{-1W} k_{2W})}. \quad (9)$$

The symbol L_{NET_3} denotes the total amount of NET_3 added to the system to catalyze the waste-forming process. The final approximation is based on the assumption that k_{-2W} and k_{-3W} are small compared to the other rate constants.

Expressions (3) and (7) are commonly adopted in the literature to model this kind of fueling and waste-forming processes⁴⁻⁶. Here, we have shown under which (reasonable) assumptions they can be considered as the coarse-grained versions of elementary reaction pathways, whose fast dynamics can be safely neglected. We stress that the results in the following do not rely on the chosen elementary model and can be extended also to more complicated schemes. Crucially, it has been shown that the coarse-graining procedure described above is thermodynamically consistent, namely it exactly retains the stationary state entropy production of the system, which is the key observable in our analysis^{3,7}.

II. DYNAMICS

The evolution in time of the concentrations of the motor species $X = \{2_H, 2_D, 1_H^H, 1_D^H, 1_H^D, 1_D^D\}$ is ruled by the following rate equations. Note that, from now on, we will abbreviate Fmoc-Cl as F, as it is the only fuel species.

$$d_t \begin{pmatrix} [2_H] \\ [2_D] \\ [1_H^D] \\ [1_D^D] \\ [1_H^H] \\ [1_D^H] \end{pmatrix} = \underbrace{\begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & -1 & 0 & 0 \end{pmatrix}}_{\mathbb{S}^X} \cdot \underbrace{\begin{pmatrix} k_{-F}^p [2_H] [HCl] - k_F^p [F] [1_H^D] \\ k_{-W}^p [2_H] - k_W^p [CO_2] [DBF] [1_H^D] \\ k_{\Delta} [1_H^D] - k_{-\Delta} [1_D^D] \\ k_F^d [F] [1_D^D] - k_{-F}^d [HCl] [2_D] \\ k_W^d [CO_2] [DBF] [1_D^D] - k_{-W}^d [2_D] \\ k_{-F}^p [HCl] [2_D] - k_F^p [F] [1_H^H] \\ k_{-W}^p [2_D] - k_W^p [CO_2] [DBF] [1_H^H] \\ k'_{\Delta} [1_D^D] - k'_{-\Delta} [1_H^H] \\ k_F^d [F] [1_H^H] - k_{-F}^d [HCl] [2_H] \\ k_W^d [CO_2] [DBF] [1_H^H] - k_{-W}^d [2_H] \end{pmatrix}}_{J = J - J_-} \quad (10)$$

The balance equations for concentrations of chemostatted species $Y = \{F, HCl, DBF, CO_2\}$ read

$$0 = d_t \begin{pmatrix} [F] \\ [HCl] \\ [DBF] \\ [CO_2] \end{pmatrix} = \underbrace{\begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & -1 \end{pmatrix}}_{\mathbb{S}^Y} \cdot \underbrace{\begin{pmatrix} k_F^p [HCl] [2_H] - k_F^p [F] [1_H^D] \\ k_{-W}^p [2_H] - k_W^p [CO_2] [DBF] [1_H^D] \\ k_{\Delta} [1_H^D] - k_{-\Delta} [1_D^D] \\ k_F^d [F] [1_D^D] - k_{-F}^d [HCl] [2_D] \\ k_W^d [CO_2] [DBF] [1_D^D] - k_{-W}^d [2_D] \\ k_{-F}^p [HCl] [2_D] - k_F^p [F] [1_H^H] \\ k_{-W}^p [2_D] - k_W^p [CO_2] [DBF] [1_H^H] \\ k'_{\Delta} [1_D^D] - k'_{-\Delta} [1_H^H] \\ k_F^d [F] [1_H^H] - k_{-F}^d [HCl] [2_H] \\ k_W^d [CO_2] [DBF] [1_H^H] - k_{-W}^d [2_H] \end{pmatrix}}_{J = J - J_-} \begin{pmatrix} I_F \\ I_{HCl} \\ I_{DBF} \\ I_{CO_2} \end{pmatrix} \quad (11)$$

with I collecting currents of the chemostating processes which keep the corresponding concentrations constant. For instance, from equation (11) we obtain the following expression for the fuelling current:

$$I_F = J_{3F^{(d)}} - J_{6F^{(d)}} - J_{1F^{(p)}} - J_{4F^{(p)}}. \quad (12)$$

To conclude this part, we introduce an alternative representation of the dynamics which will prove useful in some of the following derivations. The reader who is not interested in detailed derivations may safely skip the following formula (as well as all the equations marked with an asterisk).

$$d_t \begin{pmatrix} [2_H] \\ [2_D] \\ [1_H^D] \\ [1_D^D] \\ [1_H^H] \\ [1_D^H] \end{pmatrix} = \underbrace{\begin{pmatrix} 2_H & 2_D & 1_H^D & 1_D^D & 1_H^H & 1_D^H \\ -(\tilde{k}^p & \tilde{k}^d) & 0 & \tilde{k}^p & 0 & \tilde{k}^d & 0 \\ 0 & -(\tilde{k}^d & \tilde{k}^p) & 0 & \tilde{k}^d & 0 & \tilde{k}^p \\ \tilde{k}^p & 0 & -(\tilde{k}^p & \tilde{k}_{\Delta}) & \tilde{k}_{-\Delta} & 0 & 0 \\ 0 & \tilde{k}^d & \tilde{k}_{\Delta} & -(\tilde{k}^d & \tilde{k}_{-\Delta}) & 0 & 0 \\ \tilde{k}^d & 0 & 0 & 0 & -(\tilde{k}^d & \tilde{k}'_{-\Delta}) & \tilde{k}'_{\Delta} \\ 0 & \tilde{k}^p & 0 & 0 & \tilde{k}'_{-\Delta} & -(\tilde{k}^p & \tilde{k}'_{\Delta}) \end{pmatrix}}_{\mathbb{W}} \cdot \begin{pmatrix} [2_H] \\ [2_D] \\ [1_H^D] \\ [1_D^D] \\ [1_H^H] \\ [1_D^H] \end{pmatrix} \quad (13^*)$$

The above equations are equivalent to equations (10), and we will make use of the \mathbb{W} matrix in the following. Note that we defined new pseudo first-order rate constants marked with a tilde: the ones referring to mechanical transitions are identical to the ones introduced before (e.g., $\tilde{k}_{\Delta} = k_{\Delta}$), while for chemical transitions fuelling and waste-forming rate constants are lumped together (e.g., $\tilde{k}^d = k_F^d[F] + k_W^d[\text{DBF}][\text{CO}_2]$ and $\tilde{k}^p = k_{-F}^p[\text{HCl}] + k_{-W}^p$). Despite symbols redundancy, the use of the latter rate constants will be convenient in § V B 2.

A. Topological properties

The stoichiometric matrix $\mathbb{S} \equiv (\mathbb{S}^X, \mathbb{S}^Y)^T$ in Equations (10) and (11) encodes the topological properties of the chemical reaction network (CRN). The first class of them can be accessed by determining the cokernel of \mathbb{S} , which is spanned by the following vectors

$$\boldsymbol{\ell}_M = \begin{pmatrix} 2_H & 2_D & 1_H^D & 1_D^D & 1_H^H & 1_D^H & F & \text{HCl} & \text{DBF} & \text{CO}_2 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (14)$$

$$\boldsymbol{\ell}_{\text{CO}_2} = \begin{pmatrix} 2_H & 2_D & 1_H^D & 1_D^D & 1_H^H & 1_D^H & F & \text{HCl} & \text{DBF} & \text{CO}_2 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad (15)$$

$$\boldsymbol{\ell}_{\text{DBF}} = \begin{pmatrix} 2_H & 2_D & 1_H^D & 1_D^D & 1_H^H & 1_D^H & F & \text{HCl} & \text{DBF} & \text{CO}_2 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}, \quad (16)$$

$$\boldsymbol{\ell}_{\text{Cl}} = \begin{pmatrix} 2_H & 2_D & 1_H^D & 1_D^D & 1_H^H & 1_D^H & F & \text{HCl} & \text{DBF} & \text{CO}_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}. \quad (17)$$

The first of these vectors identifies a conserved quantity

$$L_M = \boldsymbol{\ell}_M \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [2_H] \quad [2_D] \quad [1_H^D] \quad [1_D^D] \quad [1_H^H] \quad [1_D^H] \\ d_t L_M = 0, \quad (18)$$

which is proved through the rate equations (10) and (11). It expresses the fact that the total concentration of the motor species is determined by the initial conditions and cannot be altered by the dynamics. The other three vectors identify what we call *broken* conserved quantities

$$L_{\text{CO}_2} = \boldsymbol{\ell}_{\text{CO}_2} \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [2_H] \quad [2_D] \quad [F] \quad [\text{CO}_2], \quad (19)$$

$$L_{\text{DBF}} = \boldsymbol{\ell}_{\text{DBF}} \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [2_H] \quad [2_D] \quad [F] \quad [\text{DBF}], \quad (20)$$

$$L_{\text{Cl}} = \boldsymbol{\ell}_{\text{Cl}} \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [F] \quad [\text{HCl}]. \quad (21)$$

They express the total concentrations of moieties (specific groups of atoms) whose number in the system is not determined by the initial conditions, as they are exchanged with the exterior through the chemostatting. Indeed, by using again the rate equations, it can be shown that

$$d_t L_{\text{CO}_2} = I_F - I_{\text{CO}_2}, \quad (22)$$

$$d_t L_{\text{DBF}} = I_F - I_{\text{DBF}}, \quad (23)$$

$$d_t L_{\text{Cl}} = I_F - I_{\text{HCl}}. \quad (24)$$

Namely, quantities L_{CO_2} , L_{DBF} and L_{Cl} change only due to the exchange of fuel and waste species with the chemostats. If the CRN was closed, they would be conserved by the dynamics like L_M .

The other useful piece of information contained in the stoichiometric matrix \mathbb{S} is given by cycles. A cycle is a vector in the kernel of \mathbb{S}^X , which corresponds to a series of reactions which upon completion leaves the concentrations of each motor species unchanged. They play an important role at the steady state, where reaction currents can only flow across

cycles.

$$\mathbf{c}_F^T = \begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}, \quad (25a)$$

$$\mathbf{c}_W^T = \begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}, \quad (25b)$$

$$\mathbf{c}_a^T = \begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{pmatrix}, \quad (25c)$$

$$\mathbf{c}_b^T = \begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ -1 & 1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (25d)$$

$$\mathbf{c}_\epsilon^T = \begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}. \quad (25e)$$

In the above vectors, each entry represents the number of times the corresponding reaction has to be performed in order to complete the corresponding cycle. Minus signs indicate that the corresponding reaction is performed backwards in the cycle, considering the clockwise cycling in Figure 2 of the main text as the forward direction (the same convention is adopted for the reactions at the beginning of Sec. I, where left-to-right is considered as forward direction). Cycles \mathbf{c}_F , \mathbf{c}_W , \mathbf{c}_a and \mathbf{c}_b would be present also if the system was closed (fuel and waste species not chemostatted and free to vary) and we call them *internal* cycles (mathematically, they are also right-null vectors of the whole \mathbb{S} matrix⁸). They set thermodynamic constraints on the rate constants that will be introduced later (Wegscheider's conditions, equations (36)). The cycle \mathbf{c}_ϵ involves a net exchange of chemicals with chemostats, and we call it *emergent* cycle because it arises only when the system is open. Note that the set of vectors (25) is a complete basis for the cycles in the system, but it is not unique. Another useful basis that we will exploit in § III A is the following, obtained from linear combinations of vectors (25):

$$\mathbf{c}_\epsilon^T = \begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}, \quad (26a)$$

$$\mathbf{c}_{1p}^T = \frac{1}{2}(\mathbf{c}_\epsilon^T - \mathbf{c}_F^T - \mathbf{c}_a^T) = \begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (26b)$$

$$\mathbf{c}_{3d}^T = \mathbf{c}_{1p}^T - \mathbf{c}_b^T = \begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (26c)$$

$$\mathbf{c}_{4p}^T = \mathbf{c}_a^T - \mathbf{c}_{1p}^T = \begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{pmatrix}, \quad (26d)$$

$$\mathbf{c}_{6d}^T = \mathbf{c}_\epsilon^T - \mathbf{c}_W^T - \mathbf{c}_{3d}^T = \begin{pmatrix} 1F^{(p)} & 1W^{(p)} & 2\Delta & 3F^{(d)} & 3W^{(d)} & 4F^{(p)} & 4W^{(p)} & 5\Delta' & 6F^{(d)} & 6W^{(d)} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}. \quad (26e)$$

For instance, the vector \mathbf{c}_{1p} describes the sequence of reactions $2H \xrightarrow{1W^{(p)}} 1H^D \xrightarrow{1F^{(p)}} 2H$. Cycles \mathbf{c}_{1p} , \mathbf{c}_{3d} , \mathbf{c}_{4p} , and \mathbf{c}_{6d} are called *futile* cycles, since they consume fuel molecules without contributing to a forward step of the motor.

B. Stationary state

Given a set of values for the kinetic parameters and the total concentration of motor species L_M (defined in equation (18)), for any choice of the chemostatted concentrations in Y the system will reach a stationary state distribution (mathematically, this is because the network is complex balanced and confined⁹) such that $d_t[X]_{ss} = 0$ in equation (10). It can be analytically computed using graph-theoretical techniques¹⁰⁻¹² (see also Appendix A of reference 3 for a short introduction):

$$[X_\alpha]_{ss} = \frac{L_M}{\mathcal{N}} \sum_{t \in \mathcal{T}_\alpha} \prod_{\lambda \in t} \tilde{k}_\lambda. \quad (27^*)$$

In the above expression, $\mathcal{N} = \sum_\alpha \sum_{t \in \mathcal{T}_\alpha} \prod_{\lambda \in t} \tilde{k}_\lambda$ is a normalizing denominator and \mathcal{T}_α is the set of spanning trees rooted in vertex α of the chemical reaction network representing the motor. The index λ runs over all forward and backward reactions such that \tilde{k}_λ are the pseudo-first-order rate constants as defined in equation (13*). A rooted spanning tree is a spanning tree with its edges oriented such that all edges point towards the root.

For the experimental case where equations (1) hold, the above equations boil down to:

$$\begin{pmatrix} [2_H]_{ss} \\ [2_D]_{ss} \\ [1_H^D]_{ss} \\ [1_D^D]_{ss} \\ [1_H^H]_{ss} \\ [1_D^H]_{ss} \end{pmatrix} = \frac{L_M}{\mathcal{N}} \begin{pmatrix} (k_F^p[F] \quad k_w[CO_2][DBF])(k_F^d[F] \quad k_w[CO_2][DBF]) & (k_F^p[F] \quad k_F^d[F] \quad 2k_w[CO_2][DBF])k_\Delta \\ (k_F^p[F] \quad k_w[CO_2][DBF])(k_F^d[F] \quad k_w[CO_2][DBF]) & (k_F^p[F] \quad k_F^d[F] \quad 2k_w[CO_2][DBF])k_\Delta \\ (k_F^p[HCl] \quad k_{-w})(k_F^d[F] \quad k_w[CO_2][DBF]) & (k_F^p[HCl] \quad k_F^d[HCl] \quad 2k_{-w})k_\Delta \\ (k_F^d[HCl] \quad k_{-w})(k_F^p[F] \quad k_w[CO_2][DBF]) & (k_F^p[HCl] \quad k_F^d[HCl] \quad 2k_{-w})k_\Delta \\ (k_F^d[HCl] \quad k_{-w})(k_F^p[F] \quad k_w[CO_2][DBF]) & (k_F^p[HCl] \quad k_F^d[HCl] \quad 2k_{-w})k_\Delta \\ (k_F^p[HCl] \quad k_{-w})(k_F^d[F] \quad k_w[CO_2][DBF]) & (k_F^p[HCl] \quad k_F^d[HCl] \quad 2k_{-w})k_\Delta \end{pmatrix}. \quad (28)$$

It is easy to see that if the steric hindrance caused by the macrocycle did not affect the rates of the fueling reaction (*i.e.*, $k_F^p = k_F^d$ and $k_{-F}^p = k_{-F}^d$), the stationary state would be symmetric ($[1_H^H]_{ss} = [1_D^H]_{ss} = [1_H^D]_{ss} = [1_D^D]_{ss}$).

From equation (10), it follows that the stationary state current vector J_{ss} must be a right null vector of the matrix S^X . As anticipated in the previous section (§ II A), this implies that the stationary state currents can be decomposed on a basis of cycles:

$$J_{ss} = \sum_i \psi_i c_i, \quad (29)$$

where the index i runs over the elements of a complete basis of cycles and terms ψ_i are the currents flowing across each cycle. Considering vectors (25) and (26) as basis sets:

$$\begin{pmatrix} J_{1F(p)}^{ss} \\ J_{1W(p)}^{ss} \\ J_{2\Delta}^{ss} \\ J_{3F(d)}^{ss} \\ J_{3W(d)}^{ss} \\ J_{4F(p)}^{ss} \\ J_{4W(p)}^{ss} \\ J_{5\Delta'}^{ss} \\ J_{6F(d)}^{ss} \\ J_{6W(d)}^{ss} \end{pmatrix} = \begin{pmatrix} \psi_{c_f} & \psi_{c_a} - \psi_{c_b} \\ \psi_{c_w} - \psi_{c_a} & \psi_{c_b} & \psi_{c_\epsilon} \\ \psi_{c_f} & \psi_{c_w} & \psi_{c_\epsilon} \\ \psi_{c_f} - \psi_{c_b} & \psi_{c_\epsilon} \\ \psi_{c_w} & \psi_{c_b} \\ \psi_{c_f} - \psi_{c_a} \\ \psi_{c_w} & \psi_{c_a} & \psi_{c_\epsilon} \\ \psi_{c_f} & \psi_{c_w} & \psi_{c_\epsilon} \\ \psi_{c_f} & \psi_{c_\epsilon} \\ \psi_{c_w} \end{pmatrix} = \begin{pmatrix} -\psi_{c_{1p}} \\ \psi_{c_{1p}} & \psi'_{c_\epsilon} \\ \psi'_{c_\epsilon} \\ \psi_{c_{3d}} & \psi'_{c_\epsilon} \\ -\psi_{c_{3d}} \\ -\psi_{c_{4p}} \\ \psi_{c_{4p}} & \psi'_{c_\epsilon} \\ \psi'_{c_\epsilon} \\ \psi_{c_{6d}} & \psi'_{c_\epsilon} \\ -\psi_{c_{6d}} \end{pmatrix}. \quad (30)$$

Note that, depending on the chosen basis set, the same cycle may be associated with a different cycle current (*i.e.*, $\psi_{c_\epsilon} \neq \psi'_{c_\epsilon}$) in equation (30). We will come back to this point at the end of the next section (§ III A), when the set of fundamental cycles and corresponding currents will be identified. Here, we conclude by noticing that equation (30) implies the following relations between steady state currents:

$$J_{1F(p)}^{ss} := J_{1F(p)}^{ss} \quad J_{1W(p)}^{ss} = J_{2\Delta}^{ss} = J_{3F(d)}^{ss} \quad J_{3W(d)}^{ss} = J_{4F(p)}^{ss} \quad J_{4W(p)}^{ss} = J_{5\Delta'}^{ss} = J_{6F(d)}^{ss} \quad J_{6W(d)}^{ss} = \psi_{c_\epsilon} \quad \psi_{c_w} \quad \psi_{c_f} = \psi'_{c_\epsilon}, \quad (31a)$$

$$I_F^{ss} = J_{3F(d)}^{ss} \quad J_{6F(d)}^{ss} - J_{1F(p)}^{ss} - J_{4F(p)}^{ss} = 2\psi_{c_\epsilon} = 2\psi'_{c_\epsilon} \quad \psi_{c_{1p}} \quad \psi_{c_{3d}} \quad \psi_{c_{4p}} \quad \psi_{c_{6d}}, \quad (31b)$$

where J^{ss} corresponds to the rate of net displacement of the macrocycle along the track in the stationary state. Note that J^{ss} and I_F^{ss} are respectively expressed as J and I_F in the main text, where we unambiguously refer to the stationary state.

III. THERMODYNAMICS

Each species is thermodynamically characterized by chemical potentials of the form

$$\mu_X = \mu_X^\circ - RT \ln[X], \quad \mu_Y = \mu_Y^\circ - RT \ln[Y], \quad (32)$$

where μ_X° and μ_Y° are standard-state chemical potentials. As in the main text, we will denote the sum of the chemical potentials of waste species as $\mu_W = \mu_{HCl} + \mu_{DBF} + \mu_{CO_2}$. Note that in the experimental Fmoc rotary motor (see equations (1))

we have $\mu_{1D}^\circ = \mu_{1D}^\circ = \mu_{1H}^\circ = \mu_{1H}^\circ := \mu_1^\circ$, since the benzylic amide macrocycle only interacts with the two identical fumaramide sites. The equality $\mu_{2H}^\circ = \mu_{2D}^\circ = \mu_2^\circ$ also holds in the experimental case.

Dynamics and thermodynamics can be connected via the so-called “principle of local detailed balance”⁸ which relates the ratio of forward and backward rate constants of a single elementary reaction to the difference of standard-state chemical potentials between its reagents and products, and thus to the free-energy exchanged with the reservoirs (thermal bath and chemostats) each time the reaction happens:

$$RT \ln \frac{k_{\rho}}{k_{-\rho}} = -\sum_X \mu_X^\circ \mathbb{S}_\rho^X - \sum_Y \mu_Y^\circ \mathbb{S}_\rho^Y, \quad (33)$$

where indexes X and Y run over motor’s and chemostatted species, respectively, while index ρ runs over all the 10 reactions introduced in Section I. For closed systems, this relation is essentially the law of mass action which has been known since the nineteenth’s century in chemistry. Its open systems formulation was later extensively used and was derived to ensure that microreversibility holds at equilibrium^{11,13,14}. In the broader context of statistical mechanics, similar relations were explicitly formulated in reference 15 and to the best of our knowledge the terminology “local detailed balance” first appeared in reference 16 to denote a constructive principle for formulating physically sound models describing nonequilibrium phenomena. It is nowadays the central concept underlying stochastic thermodynamics, the theory that formulates the general principles for building nonequilibrium thermodynamics for a variety of dynamics,^{17,18} not always related to chemical processes. Under certain conditions, local detailed balance also holds for nonelementary reactions as shown in references 3 and 7. This is indeed the case for the coarse-grained rate constants of the fueling and waste-forming reaction (see § IA), as commonly assumed^{4,5}.

We note that the local detailed balance as stated in (33), is different from how local detailed balance is discussed in references 5 and 19. Indeed, only in contexts where there is only one possible transition (e.g. one chemical reaction) connecting a pair of states (e.g. two motor species), local detailed balance reduces to a constraint on the forward and backward transition rates connecting the two. But in the cases where there are multiple transitions between a couple of states (e.g. two reactions involving the same motor species but different chemostats), local detailed balance only holds at the level of each single transition^{16,17,20}, as done in equation (33), and not at the level of the net (i.e. summed) transition rates. This important point is also discussed in comment 21.

At equilibrium, the thermodynamic forces driving each reaction (also called *affinities* in the literature, a term not adopted here to avoid confusion with binding affinities) vanish

$$A_\rho^{\text{eq}} = -\sum_X \mu_X^{\text{eq}} \mathbb{S}_\rho^X - \sum_Y \mu_Y^{\text{eq}} \mathbb{S}_\rho^Y = 0, \quad (34)$$

as well as all reaction currents

$$J_\rho^{\text{eq}} = J_\rho^{\text{eq}} - J_{-\rho}^{\text{eq}} = 0. \quad (35)$$

Equations (33) imply the Wegscheider’s conditions for the internal cycles: the product of the forward rate constants along each internal cycle must be equal to that of the backward rate constants

$$\text{from } \mathbf{c}_F : \quad k_{-F}^p k_{\Delta}^d k_{-F}^p k'_{\Delta} k_{-F}^d = k_{-F}^p k_{-\Delta}^d k_{-F}^p k_{-F}^d k'_{-\Delta} k_{-F}^d, \quad (36a)$$

$$\text{from } \mathbf{c}_W : \quad k_{-W}^p k_{\Delta}^d k_{-W}^p k'_{\Delta} k_{-W}^d = k_{-W}^p k_{-\Delta}^d k_{-W}^p k_{-W}^d k'_{-\Delta} k_{-W}^d, \quad (36b)$$

$$\text{from } \mathbf{c}_a : \quad k_{-F}^p k_{-W}^p k_{-F}^p k_{-W}^p = k_{-F}^p k_{-W}^p k_{-F}^p k_{-W}^p, \quad (36c)$$

$$\text{from } \mathbf{c}_b : \quad k_{-F}^p k_{-W}^p k_{-F}^d k_{-W}^d = k_{-F}^p k_{-W}^p k_{-F}^d k_{-W}^d. \quad (36d)$$

They set thermodynamic constraints that the kinetic parameters must satisfy, thus reducing freedom in their choice. Note that, due to symmetries in the system (i.e., rate constants of reactions 1F^(p) and 4F^(p) are considered identical), condition (36c) is trivially satisfied and condition (36d) is implied by (36a) and (36b), which are therefore the only ones reported in the main text.

The dissipation of the process is captured by the entropy production rate times the temperature T :

$$T\dot{\Sigma} = RT \sum_\rho J_\rho \ln \frac{J_\rho}{J_{-\rho}} \geq 0, \quad (37)$$

which also vanishes at equilibrium. Equation (37) shows that each of the 10 reactions in the vector $\boldsymbol{\rho}$ (see Section I) brings a positive contribution to the total dissipation. We will use this property in Section IV to identify the specific contributions of chemical and mechanical processes.

By using the rate equations and the local detailed balance (33), the dissipation in equation (37) can be rewritten as

$$T\dot{\Sigma} = -d_t \mathcal{G} - \dot{W}_{\text{fuel}}, \quad (38)$$

where

$$\mathcal{G} = [1_{\text{H}}^{\text{D}}]\mu_{1_{\text{H}}^{\text{D}}} \quad [1_{\text{D}}^{\text{D}}]\mu_{1_{\text{D}}^{\text{D}}} \quad [1_{\text{H}}^{\text{H}}]\mu_{1_{\text{H}}^{\text{H}}} \quad [1_{\text{D}}^{\text{H}}]\mu_{1_{\text{D}}^{\text{H}}} \quad [2_{\text{H}}] \left(\mu_{2_{\text{H}}} - \mu_{\text{DBF}} - \mu_{\text{CO}_2} \right) \quad [2_{\text{D}}] \left(\mu_{2_{\text{D}}} - \mu_{\text{DBF}} - \mu_{\text{CO}_2} \right) \\ [F] (\mu_{\text{F}} - \mu_{\text{W}}) - RT (L_{\text{M}} \quad [F] \quad [\text{HCl}] \quad [\text{DBF}] \quad [\text{CO}_2]) \quad (39)$$

is the *semigrand* Gibbs potential of the system⁸, and

$$\dot{\mathcal{W}}_{\text{fuel}} = I_{\text{F}}(\mu_{\text{F}} - \mu_{\text{W}}) \quad (40)$$

is the fueling chemical work per unit of time (*i.e.*, the fueling power). If $\mu_{\text{F}} = \mu_{\text{W}}$, equation (38) shows that \mathcal{G} is a monotonically decreasing function in time, given that $T\dot{\Sigma} \geq 0$.

A. Steady state entropy production and $J^{\text{ss}}/I_{\text{F}}^{\text{ss}}$

As pointed out in the main text, at stationary state (see § II B) the dissipation in equation (38) boils down to:

$$T\dot{\Sigma}^{\text{ss}} = \dot{\mathcal{W}}_{\text{fuel}}^{\text{ss}} = I_{\text{F}}^{\text{ss}}(\mu_{\text{F}} - \mu_{\text{W}}). \quad (41)$$

It is instructive to look at the equivalent expressions of equation (41) which are obtained by decomposing the stationary I_{F}^{ss} current on both cycles basis (25) and (26):

$$T\dot{\Sigma}^{\text{ss}} = 2\psi_{c_{\epsilon}}(\mu_{\text{F}} - \mu_{\text{W}}) = (\psi_{c_{1p}} \quad \psi_{c_{3d}} \quad \psi_{c_{4p}} \quad \psi_{c_{6d}} \quad 2\psi'_{c_{\epsilon}})(\mu_{\text{F}} - \mu_{\text{W}}). \quad (42)$$

The first equality (decomposition according to the basis of vectors (25)) identifies the cycle current $\psi_{c_{\epsilon}}$ as the fundamental current in the stationary regime: it accounts for total dissipation once multiplied by its conjugated fundamental force $2(\mu_{\text{F}} - \mu_{\text{W}})$, also denoted *cycle's affinity*; this is a consequence of the presence of a unique emergent cycle. The second one (decomposition according to the basis of vectors (26)) highlights how the consumption of fuel can be thought of as happening due to 5 different processes: the directional cycling of the macrocycle ($\psi'_{c_{\epsilon}}$, which one would like to enhance), and four futile cycles consuming fuel without inducing directional movement of the macrocycle ($\psi_{c_{1p}}$, $\psi_{c_{3d}}$, $\psi_{c_{4p}}$, and $\psi_{c_{6d}}$ which one would like to suppress). From the latter perspective, we use relations (31) to introduce the following *coefficient of performance* ($J^{\text{ss}}/I_{\text{F}}^{\text{ss}}$) quantifying how many fuel molecules are needed on average to complete a directional cycle:

$$\frac{J^{\text{ss}}}{I_{\text{F}}^{\text{ss}}} = \frac{\psi'_{c_{\epsilon}}}{\psi_{c_{1p}} \quad \psi_{c_{3d}} \quad \psi_{c_{4p}} \quad \psi_{c_{6d}} \quad 2\psi'_{c_{\epsilon}}}. \quad (43)$$

In the best possible scenario, namely when futile currents vanish and the fuel consumption is tightly coupled with macrocycle directional movement, we have an average of 2 fuel molecules consumed per cycle, leading to an optimal $J^{\text{ss}}/I_{\text{F}}^{\text{ss}}$ of 0.5. Equation (43) allows to quantify how much the motor deviates from the optimal condition and will be employed in the supplementary simulations of § VIB to further compare different designs. In particular, the values of $J^{\text{ss}}/I_{\text{F}}^{\text{ss}}$ for the rotary motor can be directly compared with similar data measured for biological molecular motors. For instance, in reference 22 ATP-synthase in yeast mitochondria was found to use on average about 3.9 H⁺ ions to synthesize one ATP molecule, when the expected theoretical value predicted by models was 3.3 (85% of the theoretical limit). As anticipated in the main text, our minimalist rotary motor consumes on average seven fuel molecules per cycle, corresponding to a $J^{\text{ss}}/I_{\text{F}}^{\text{ss}}$ of 0.15 (30% of the theoretical limit) in the experimental condition.

IV. INFORMATION THERMODYNAMICS PERSPECTIVE

If we stick to the description above, the motor appears as an autonomous system exploiting the chemical gradient $\mu_{\text{F}} - \mu_{\text{W}}$ to sustain currents. We can evaluate its performance through equation (43), but questions such as how effectively the free energy harvested from the fuel is transduced by the motor to produce directional motion remain open. The detailed mechanism is better understood when the motor is thought of as a bipartite system, where we distinguish between “mechanical” states $\mathbf{mech} = \{\text{H}, \text{D}\}$ referring to the position of the macrocycle, and “chemical” ones $\mathbf{chem} = \{2, 1^{\text{H}}, 1^{\text{D}}\}$ indicating the state of the track (how many stoppers attached, if one, where). With this structure in mind, the entropy production in equation (37) can be split as:

$$\begin{aligned}
T\dot{\Sigma} &= RT \sum_{\rho} J_{\rho} \ln \frac{J_{\rho}}{J_{-\rho}} = TR \sum_{\rho} J_{\rho} \ln \left(\frac{k_{-\rho}}{k_{\rho}} Y_{\sigma}^{-S^{\nu\sigma}} \right) = TR \sum_{\rho} J_{\rho} \ln X_{\sigma}^{-S^{\nu\sigma}} = \\
&= \underbrace{RT J_{2\Delta} \ln \frac{k_{-\Delta}}{k_{\Delta}} \quad RT J_{5\Delta'} \ln \frac{k'_{-\Delta}}{k'_{\Delta}}}_{T\dot{S}_r^{\text{mech}}} \underbrace{RT J_{2\Delta} \ln \frac{[1_D^D]}{[1_D^D]} \quad RT J_{5\Delta'} \ln \frac{[1_H^H]}{[1_H^H]}}_{T\dot{d}_t S^{\text{mech}}} \\
&\quad \underbrace{RT (J_{1F^{(p)}} \quad J_{4F^{(p)}}) \ln \frac{k_{-F}^p [\text{HCl}]}{k_F^p [\text{F}]} \quad RT (J_{3F^{(d)}} \quad J_{6F^{(d)}}) \ln \frac{k_{-F}^d [\text{F}]}{k_F^d [\text{HCl}]}}_{T\dot{S}_r^{\text{chem,F}}} \\
&\quad \underbrace{RT (J_{1W^{(p)}} \quad J_{4W^{(p)}}) \ln \frac{k_{-W}^p}{k_W^p [\text{DBF}] [\text{CO}_2]} \quad RT (J_{3W^{(d)}} \quad J_{6W^{(d)}}) \ln \frac{k_{-W}^d [\text{DBF}] [\text{CO}_2]}{k_W^d}}_{T\dot{S}_r^{\text{chem,W}}} \\
&\quad \underbrace{RT (J_{1F^{(p)}} \quad J_{1W^{(p)}}) \ln \frac{[2_H]}{[1_H^D]} \quad RT (J_{3F^{(d)}} \quad J_{3W^{(d)}}) \ln \frac{[1_D^D]}{[2_D]} \quad RT (J_{4F^{(p)}} \quad J_{4W^{(p)}}) \ln \frac{[2_D]}{[1_H^D]} \quad RT (J_{6F^{(d)}} \quad J_{6W^{(d)}}) \ln \frac{[1_H^H]}{[2_H]}}_{T\dot{d}_t S^{\text{chem}}} \geq 0, \tag{44}
\end{aligned}$$

where terms

$$T\dot{S}_r^{\text{mech}} = J_{2\Delta} (\mu_{1_D^D}^{\circ} - \mu_{1_D^D}^{\circ}) \quad J_{5\Delta'} (\mu_{1_H^H}^{\circ} - \mu_{1_H^H}^{\circ}) \tag{45a}$$

$$\begin{aligned}
T\dot{S}_r^{\text{chem,F}} \quad T\dot{S}_r^{\text{chem,W}} &= \overbrace{I_F (\mu_F - \mu_W)}^{\mathcal{W}_{\text{fuel}}} (J_{1F^{(p)}} \quad J_{1W^{(p)}}) (\mu_{2_H}^{\circ} - \mu_{1_D^D}^{\circ} - \mu_{\text{DBF}} - \mu_{\text{CO}_2}) \\
&\quad (J_{4F^{(p)}} \quad J_{4W^{(p)}}) (\mu_{2_D}^{\circ} - \mu_{1_D^D}^{\circ} - \mu_{\text{DBF}} - \mu_{\text{CO}_2}) \quad (J_{3F^{(d)}} \quad J_{3W^{(d)}}) (\mu_{1_D^D}^{\circ} - \mu_{2_D}^{\circ} \quad \mu_{\text{DBF}} \quad \mu_{\text{CO}_2}) \\
&\quad (J_{6F^{(d)}} \quad J_{6W^{(d)}}) (\mu_{1_H^H}^{\circ} - \mu_{2_H}^{\circ} \quad \mu_{\text{DBF}} \quad \mu_{\text{CO}_2}) \tag{45b}
\end{aligned}$$

keep track of the free energy exchanged with the reservoirs by each of the two sets of degrees of freedom. Note that fueling work by the chemostats is done on chemical degrees of freedom only. Terms $\dot{d}_t S^{\text{mech}}$ and $\dot{d}_t S^{\text{chem}}$ can be rewritten by introducing cumulative concentrations for the chemical subsystem: $[2] = [2_H] \quad [2_D]$, $[1^H] = [1_D^H] \quad [1_H^H]$, $[1^D] = [1_D^D] \quad [1_H^D]$; and for the mechanical one: $[H^{\text{mech}}] = [2_H] \quad [1_H^H] \quad [1_D^D]$, $[D^{\text{mech}}] = [2_D] \quad [1_D^D] \quad [1_H^D]$. We thus have:

$$\begin{aligned}
\dot{d}_t S^{\text{mech}}/R &= \underbrace{(J_{2\Delta} - J_{5\Delta'}) \ln \frac{[H^{\text{mech}}]}{[D^{\text{mech}}]}}_{\dot{d}_t S^{\text{mech}}/R} \underbrace{- J_{2\Delta} \ln \frac{[H^{\text{mech}}][1_D^D]}{[D^{\text{mech}}][1_H^D]} - J_{5\Delta'} \ln \frac{[D^{\text{mech}}][1_H^H]}{[H^{\text{mech}}][1_D^D]}}_{-\dot{j}^{\text{mech}}} \\
\dot{d}_t S^{\text{chem}}/R &= \underbrace{(J_{1F^{(p)}} \quad J_{1W^{(p)}} - J_{3F^{(d)}} - J_{3W^{(d)}}) \ln \frac{[2]}{[1^D]} \quad (J_{4F^{(p)}} \quad J_{4W^{(p)}} - J_{6F^{(d)}} - J_{6W^{(d)}}) \ln \frac{[2]}{[1^H]}}_{\dot{d}_t S^{\text{chem}}/R} \\
&\quad \underbrace{-(J_{1F^{(p)}} \quad J_{1W^{(p)}}) \ln \frac{[2][1_H^D]}{[1^D][2_H]} - (J_{3F^{(d)}} \quad J_{3W^{(d)}}) \ln \frac{[1^D][2_D]}{[2][1_D^D]} - (J_{4F^{(p)}} \quad J_{4W^{(p)}}) \ln \frac{[2][1_H^H]}{[1^H][2_D]} - (J_{6F^{(d)}} \quad J_{6W^{(d)}}) \ln \frac{[1^H][2_H]}{[2][1_H^H]}}_{-\dot{j}^{\text{chem}}}. \tag{46}
\end{aligned}$$

Now the interpretation of $\dot{d}_t S^{\text{mech}}$ and $\dot{d}_t S^{\text{chem}}$ is clear: they are the time derivatives of the subsystem's Shannon-like entropies^{8,23}. If we could measure the two subsystems separately, the entropy production rates we would assign to them

would be:

$$\sigma^{\text{mech}} := \dot{S}_r^{\text{mech}} \quad d_t S^{\text{mech}} \quad (47)$$

$$\sigma^{\text{chem}} := \dot{S}_r^{\text{chem,F}} \quad \dot{S}_r^{\text{chem,W}} \quad d_t S^{\text{chem}} . \quad (48)$$

Concerning the other terms, we now justify the names \dot{I}^{mech} and \dot{I}^{chem} by showing that their sum gives the time variation of the mutual information between the two subsystems. The reader who is not interested in technical details can directly go to equation (51).

If we focus on the stochastic behavior of a single motor unit, the joint probability to find it in a certain mechanical and chemical state evolves according to the following master equation:

$$d_t p(\text{chem}, \text{mech}) = \sum_{\text{mech}', \text{chem}'} [\mathbb{W}_{(\text{chem}, \text{mech}), (\text{chem}', \text{mech}')} p(\text{chem}', \text{mech}')] . \quad (49^*)$$

For instance, it is analogous to the one employed in reference 24 to study another example of a molecular motor. Note how the structure of equation (49*) is analogous to the equation (13*) evolving the concentrations of the various motor species. Indeed, when the number N of single motor units in the system is large, the probabilities $p(\text{mech}, \text{chem})$ for each chemical species become sharply peaked on their average values. Thus, we can perform the substitution $p(\text{mech}, \text{chem}) = \frac{N(\text{mech}, \text{chem})}{N} = \frac{[X]}{L_M}$, which shows that equation (13*) is the macroscopic limit of equation (49*). This is crucial, because it implies that for such a case taking the macroscopic limit of stochastic quantities just means to switch from the probability of a certain state to its concentrations. In this case, the stochastic mutual information between mechanical and chemical subsystems reads:

$$I_M = \sum_{\text{mech}, \text{chem}} p(\text{mech}, \text{chem}) \ln \frac{p(\text{mech}, \text{chem})}{p(\text{mech})p(\text{chem})} \geq 0, \quad (50^*)$$

which is a measure of the correlation between the two sets of degrees of freedom. In the above equation, $p(\text{mech}) = \sum_{\text{chem}} p(\text{mech}, \text{chem})$ and $p(\text{chem}) = \sum_{\text{mech}} p(\text{mech}, \text{chem})$. Thus, the mutual information in terms of macroscopic concentrations can be defined as:

$$I := L_M I_M = \left([2_H] \ln \frac{[2_H]}{[H^{\text{mech}}][2]} \quad [2_D] \ln \frac{[2_D]}{[D^{\text{mech}}][2]} \quad [1_H^D] \ln \frac{[1_H^D]}{[H^{\text{mech}}][1^D]} \right. \\ \left. [1_D^D] \ln \frac{[1_D^D]}{[D^{\text{mech}}][1^D]} \quad [1_D^H] \ln \frac{[1_D^H]}{[D^{\text{mech}}][1^H]} \quad [1_H^H] \ln \frac{[1_H^H]}{[H^{\text{mech}}][1^H]} \right) L_M \ln(L_M), \quad (51)$$

and by straightforward computation and equation (18) we get

$$d_t I = \dot{I}^{\text{mech}} \quad \dot{I}^{\text{chem}} . \quad (52)$$

To summarize, we showed that the entropy production of the full system can be written as a sum of two individually positive terms:

$$T \dot{\Sigma} = T \dot{\Sigma}^{\text{mech}} \quad T \dot{\Sigma}^{\text{chem}} = \underbrace{T \sigma^{\text{mech}} - RT \dot{I}^{\text{mech}}}_{\geq 0} \quad \underbrace{T \sigma^{\text{chem}} - RT \dot{I}^{\text{chem}}}_{\geq 0} . \quad (53)$$

Indeed, by substituting into equation (53) expressions (48), (46) and (45), we find

$$T \dot{\Sigma}^{\text{mech}} = T \dot{S}_r^{\text{mech}} \quad T d_t S^{\text{mech}} = RT \sum_{\rho^{\text{mech}}} J_{\rho^{\text{mech}}} \ln \frac{J_{\rho^{\text{mech}}}}{J_{-\rho^{\text{mech}}}} \geq 0 \\ T \dot{\Sigma}^{\text{chem}} = T \dot{S}_r^{\text{chem}} \quad T d_t S^{\text{chem}} = RT \sum_{\rho^{\text{chem}}} J_{\rho^{\text{chem}}} \ln \frac{J_{\rho^{\text{chem}}}}{J_{-\rho^{\text{chem}}}} \geq 0 \quad (54)$$

where $\rho^{\text{mech}} = \{2\Delta, 5\Delta'\}$ and $\rho^{\text{chem}} = \{1F^{(p)}, 1W^{(p)}, 3F^{(d)}, 3W^{(d)}, 4F^{(p)}, 4W^{(p)}, 6F^{(d)}, 6W^{(d)}\}$ denote the reactions changing the mechanical and the chemical state of the motor, respectively.

At stationary state, we have that $d_t I = 0 = \dot{I}_{ss}^{\text{mech}} \quad \dot{I}_{ss}^{\text{chem}}$ (equation (51)), and also $d_t S^{\text{mech}}$ and $d_t S^{\text{chem}}$ vanish, as they are state functions. By specializing equation (53) to the stationary state (where relations (31) are employed to simplify the final expression) and defining $\dot{I}_{ss}^{\text{mech}} = -\dot{I}_{ss}^{\text{chem}} = -\dot{I}$ and $T \dot{S}_{r,ss}^{\text{mech}} = \dot{\mathcal{E}}$, we finally get the equation reported in the main text (where the ss symbol has been dropped since there we unambiguously refer to the stationary state):

$$T \dot{\Sigma}^{\text{ss}} = T \dot{\Sigma}_{ss}^{\text{mech}} \quad T \dot{\Sigma}_{ss}^{\text{chem}} = \underbrace{I_F^{\text{ss}}(\mu_F - \mu_W) - \dot{\mathcal{E}} - RT \dot{I}}_{\geq 0} \quad \underbrace{\dot{\mathcal{E}} \quad RT \dot{I}}_{\geq 0} \geq 0 . \quad (55)$$

V. THERMODYNAMIC CONSTRAINTS ON THE STATIONARY STATE DYNAMICS AND CONNECTION WITH KINETIC ASYMMETRY: THE ROTARY MOTOR AS A ROSETTA STONE FOR KINETIC AND THERMODYNAMIC ANALYSIS

In this section, we examine in detail how equation (55) implies previous results obtained in the literature, namely the kinetic asymmetry rule^{4,5}. In general, we can distinguish three possible stationary states for this system. The equilibrium (i) is always reached when the chemical potential gradient acting on the system is null ($\mu_F = \mu_W$), no matter the values of the rate constants (provided they fulfill Wegscheider's conditions in equations (36)). At equilibrium each reaction current vanishes ($J_\rho^{\text{eq}} = J_\rho^{\text{eq}} - J_{-\rho}^{\text{eq}} = 0$) and no average net displacement of the macrocycle with respect to the track can be observed. Whenever the chemical potentials of fuel and waste species are kept different (e.g., $\mu_F > \mu_W$), a nonequilibrium stationary state with net currents is reached in the long time limit (see § II B). This situation can be "symmetric" (ii) – always when the position of the macrocycle does not affect the rates of the fueling and waste-forming reactions – or "asymmetric" (iii) – the general case for the experiment we are considering. The latter is the interesting one from the point of view of directionality, because it corresponds to an average directional motion of the macrocycle along the track characterized by a non-null stationary state current J^{ss} (equation (31)).

A. Condition for directional J^{ss}

By considering the decomposition of the dissipation in equation (55):

$$T\dot{\Sigma}^{\text{ss}} = \underbrace{I_F^{\text{ss}}(\mu_F - \mu_W) - \dot{\mathcal{E}} - RT\dot{J}}_{T\dot{\Sigma}_{\text{ss}}^{\text{chem}}} \underbrace{J^{\text{ss}}RT \ln \frac{k_{\Delta}k'_{\Delta}}{k_{-\Delta}k'_{-\Delta}}}_{\dot{\mathcal{E}}} \underbrace{J^{\text{ss}}RT \ln \frac{[1_{\text{H}}^{\text{D}}]_{\text{ss}}[1_{\text{D}}^{\text{H}}]_{\text{ss}}}{[1_{\text{D}}^{\text{D}}]_{\text{ss}}[1_{\text{H}}^{\text{H}}]_{\text{ss}}}}_{RT\dot{J}}, \quad (56)$$

the following inequalities hold by virtue of the second law of thermodynamics for bipartite systems:

$$T\dot{\Sigma}_{\text{ss}}^{\text{chem}} = I_F^{\text{ss}}(\mu_F - \mu_W) - T\dot{\Sigma}_{\text{ss}}^{\text{mech}} \geq 0 \quad (57a)$$

$$T\dot{\Sigma}_{\text{ss}}^{\text{mech}} = \dot{\mathcal{E}} \quad RT\dot{J} = J^{\text{ss}}RT \left(\ln \frac{k_{\Delta}k'_{\Delta}}{k_{-\Delta}k'_{-\Delta}} \quad \ln \frac{[1_{\text{H}}^{\text{D}}]_{\text{ss}}[1_{\text{D}}^{\text{H}}]_{\text{ss}}}{[1_{\text{D}}^{\text{D}}]_{\text{ss}}[1_{\text{H}}^{\text{H}}]_{\text{ss}}} \right) \geq 0. \quad (57b)$$

As highlighted in the main text, net directional displacement of the macrocycle in the stationary state can be observed only in the presence of mechanical dissipation ($T\dot{\Sigma}_{\text{ss}}^{\text{mech}} > 0$), as directed motion is a nonequilibrium behaviour. This implies that in order to have a non null J^{ss} , a positive fueling work is required ($I_F^{\text{ss}}(\mu_F - \mu_W) > 0$ in equation (57a), otherwise the stationary state would be an equilibrium one). This is not sufficient, since the sum of energy and information flow in equation (57b) must be positive too, namely a net free energy transfer from chemical to mechanical processes must be present.

The condition for a net directional displacement of the macrocycle in the stationary state can then be written as:

$$\ln \frac{k_{\Delta}k'_{\Delta}}{k_{-\Delta}k'_{-\Delta}} \ln \frac{[1_{\text{H}}^{\text{D}}]_{\text{ss}}[1_{\text{D}}^{\text{H}}]_{\text{ss}}}{[1_{\text{D}}^{\text{D}}]_{\text{ss}}[1_{\text{H}}^{\text{H}}]_{\text{ss}}} = \ln \frac{J_{2\Delta}^{\text{ss}}J_{5\Delta'}^{\text{ss}}}{J_{-2\Delta}^{\text{ss}}J_{-5\Delta'}^{\text{ss}}} \neq 0. \quad (58)$$

Thanks to equation (31), it's easy to realize that whenever the quantity in equation (58) is positive, J^{ss} is positive too, and vice-versa, thus guaranteeing equation (57b) to always hold. We will now show how the condition (58) (obtained with nonequilibrium thermodynamic arguments) is equivalent to previous results obtained in the literature (based on kinetic arguments).

B. Connection to kinetic asymmetry and K_r

According to references 25 and 6, the condition for having directional currents in nonequilibrium chemical systems ($J^{\text{ss}} \neq 0$) can be expressed in terms of the ratcheting constant, which for the motor under study reads

$$K_r = \frac{k_{\Delta}k'_{\Delta}(k_{-F}^{\text{p}}[\text{HCl}] \quad k_{-W}^{\text{p}})^2([F]k_{-F}^{\text{d}} \quad [\text{CO}_2][\text{DBF}]k_{-W}^{\text{d}})^2}{k_{-\Delta}k'_{-\Delta}([F]k_{-F}^{\text{p}} \quad [\text{CO}_2][\text{DBF}]k_{-W}^{\text{p}})^2(k_{-F}^{\text{d}}[\text{HCl}] \quad k_{-W}^{\text{d}})^2} \neq 1, \quad (59)$$

with $K_r > 1$ implying a positive forward current, and vice-versa. By using Wegscheider's conditions (36), K_r can be rewritten as^{4,5,25,26}:

$$K_r = \frac{k_{\Delta} k'_{\Delta} (k_{-F}^p [\text{HCl}])^2 ([\text{F}] k_F^d)^2 \left(1 - \frac{k_{-W}^p}{k_{-F}^p [\text{HCl}]}\right)^2 \left(1 - \frac{[\text{CO}_2] [\text{DBF}] k_W^d}{[\text{F}] k_F^d}\right)^2}{k_{-\Delta} k'_{-\Delta} ([\text{F}] k_F^p)^2 (k_{-F}^d [\text{HCl}])^2 \left(1 - \frac{[\text{CO}_2] [\text{DBF}] k_W^p}{[\text{F}] k_F^p}\right)^2 \left(1 - \frac{k_{-W}^d}{k_{-F}^d [\text{HCl}]}\right)^2} = \frac{\left(1 - \frac{k_{-W}^p}{k_{-F}^p [\text{HCl}]}\right)^2 \left(1 - \frac{[\text{CO}_2] [\text{DBF}] k_W^d}{[\text{F}] k_F^d}\right)^2}{\left(1 - \frac{[\text{CO}_2] [\text{DBF}] k_W^p}{[\text{F}] k_F^p}\right)^2 \left(1 - \frac{k_{-W}^d}{k_{-F}^d [\text{HCl}]}\right)^2}. \quad (60)$$

=1, see equation (36a)

In the following, we show how the condition $K_r \neq 1$ for a directional current $J^{\text{ss}} \neq 0$ is implied by equations (58).

1. Equilibrium condition

The expression of the ratcheting constant in equation (60) can be further rearranged thanks to equations (32) and (33) to get (see also Eq. (5) in 4):

$$K_r = \frac{\left(1 - \frac{k_{-W}^p}{k_{-F}^p [\text{HCl}]}\right)^2 \left(1 - \frac{k_{-W}^d}{k_{-F}^d [\text{HCl}]} e^{-\frac{\mu_F - \mu_W}{RT}}\right)^2}{\left(1 - \frac{k_{-W}^p}{k_{-F}^p [\text{HCl}]} e^{-\frac{\mu_F - \mu_W}{RT}}\right)^2 \left(1 - \frac{k_{-W}^d}{k_{-F}^d [\text{HCl}]}\right)^2}, \quad (61)$$

which nicely shows that $K_r = 1$ (and then $J^{\text{ss}} = 0$) whenever $\mu_F = \mu_W$, in analogy with equation (57a). Regardless of the values of the rate constants, if there is no thermodynamic force ($\mu_F - \mu_W = 0$), the motor will relax to an equilibrium stationary state with no directional currents, and the quantity in equation (58) will be zero.

2. Proof of the implication $\Sigma^{\text{mech}} > 0 \rightarrow K_r \neq 1$

We now prove that the condition on K_r in equation (59) directly follows from equation (58). We do so by relying on some results and techniques from graph theory (see appendixes of reference 3 for an introduction). The reader who is not familiar with this kind of techniques can directly go to the next paragraph, where a simpler proof valid for the experimental case (where equations 1 hold) is presented.

First of all, we define K_r in terms of the pseudo rate constants introduced in equation (13*):

$$K_r = \frac{\tilde{k}_{\Delta} \tilde{k}'_{\Delta} (\tilde{k}_{\Delta}^p \tilde{k}_{\Delta}^d)^2}{\tilde{k}_{-\Delta} \tilde{k}'_{-\Delta} (\tilde{k}_{-\Delta}^p \tilde{k}_{-\Delta}^d)^2} = \frac{\mathcal{P}_{\text{cw}}}{\mathcal{P}_{\text{ccw}}}, \quad (62)$$

where \mathcal{P}_{cw} and \mathcal{P}_{ccw} are the product of all forward clockwise and counterclockwise pseudo reaction constants, respectively. By plugging the analytical expression (27*) for the steady state concentrations into the condition (58), the latter can be rewritten as

$$\left(\tilde{k}_{\Delta} \sum_{t \in \mathcal{T}_{\text{H}}^{\text{D}}} \prod_{\rho \in t} \tilde{k}_{\rho} \right) \cdot \left(\tilde{k}'_{\Delta} \sum_{t \in \mathcal{T}_{\text{H}}^{\text{D}}} \prod_{\rho \in t} \tilde{k}_{\rho} \right) - \left(\tilde{k}_{-\Delta} \sum_{t \in \mathcal{T}_{\text{H}}^{\text{D}}} \prod_{\rho \in t} \tilde{k}_{\rho} \right) \cdot \left(\tilde{k}'_{-\Delta} \sum_{t \in \mathcal{T}_{\text{H}}^{\text{D}}} \prod_{\rho \in t} \tilde{k}_{\rho} \right) \neq 0. \quad (63^*)$$

We recall that, for instance, the symbol $\mathcal{T}_{\text{H}}^{\text{D}}$ denotes the set of spanning trees rooted in vertex 1_{H}^{D} of the chemical reaction network representing the motor. Therefore, each parenthesis in the above expression contains a sum of 6 terms that are

each the product of 6 pseudo reaction constants. We can manipulate expression (63*) to get

$$\begin{aligned}
& \left(\tilde{k}_{-\Delta} \sum_{t \in \mathcal{T}_{\text{D}}^{\text{H}} \wedge \tilde{k}_{-\Delta} \in t} \prod_{\rho \in t} \tilde{k}_{\rho} \mathcal{P}_{\text{cw}} \right) \cdot \left(\tilde{k}'_{-\Delta} \sum_{t \in \mathcal{T}_{\text{D}}^{\text{H}} \wedge \tilde{k}'_{-\Delta} \in t} \prod_{\rho \in t} \tilde{k}_{\rho} \mathcal{P}_{\text{cw}} \right) \\
& - \left(\tilde{k}_{-\Delta} \sum_{t \in \mathcal{T}_{\text{D}}^{\text{D}} \wedge \tilde{k}_{-\Delta} \in t} \prod_{\rho \in t} \tilde{k}_{\rho} \mathcal{P}_{\text{ccw}} \right) \cdot \left(\tilde{k}'_{-\Delta} \sum_{t \in \mathcal{T}_{\text{D}}^{\text{H}} \wedge \tilde{k}'_{-\Delta} \in t} \prod_{\rho \in t} \tilde{k}_{\rho} \mathcal{P}_{\text{ccw}} \right) = \\
& 0 \quad (\mathcal{P}_{\text{cw}}^2 - \mathcal{P}_{\text{ccw}}^2) \mathcal{P}_{\text{cw}} \underbrace{\left[\tilde{k}_{-\Delta} \sum_{t \in \mathcal{T}_{\text{D}}^{\text{H}} \wedge \tilde{k}_{-\Delta} \in t} \prod_{\rho \in t} \tilde{k}_{\rho} \tilde{k}'_{-\Delta} \sum_{t \in \mathcal{T}_{\text{D}}^{\text{D}} \wedge \tilde{k}'_{-\Delta} \in t} \prod_{\rho \in t} \tilde{k}_{\rho} \right]}_{\Phi} \\
& - \mathcal{P}_{\text{ccw}} \underbrace{\left[\tilde{k}_{-\Delta} \sum_{t \in \mathcal{T}_{\text{D}}^{\text{D}} \wedge \tilde{k}_{-\Delta} \in t} \prod_{\rho \in t} \tilde{k}_{\rho} \tilde{k}'_{-\Delta} \sum_{t \in \mathcal{T}_{\text{D}}^{\text{H}} \wedge \tilde{k}'_{-\Delta} \in t} \prod_{\rho \in t} \tilde{k}_{\rho} \right]}_{\Phi} = (\mathcal{P}_{\text{cw}} - \mathcal{P}_{\text{ccw}}) \cdot (\mathcal{P}_{\text{cw}} \mathcal{P}_{\text{ccw}} \Phi) \neq 0. \quad (64^*)
\end{aligned}$$

Since \mathcal{P}_{cw} , \mathcal{P}_{ccw} , and Φ are positive quantities, the above condition is satisfied if and only if $\mathcal{P}_{\text{cw}} - \mathcal{P}_{\text{ccw}} \neq 0$ or, alternatively, $\mathcal{P}_{\text{cw}}/\mathcal{P}_{\text{ccw}} \neq 1$, thus fully proving the equivalence between condition (59) from references 4, 25, 5, and 6 and condition (58) derived from the second law (57) for bipartite systems²³.

3. Special proof for the experimental case

Here, we repeat the above proof for the special case of the experimental rotary motor, where the symmetries in equations (1) hold and the expressions for stationary state concentrations is given by equations (28). By plugging the latter in equation (58), the condition for directional current boils down to:

$$\frac{(k_{\text{-F}}^{\text{p}}[\text{HCl}] \quad k_{\text{-W}})(k_{\text{-F}}^{\text{d}}[\text{F}] \quad k_{\text{W}}[\text{CO}_2][\text{DBF}])}{(k_{\text{-F}}^{\text{d}}[\text{HCl}] \quad k_{\text{-W}})(k_{\text{-F}}^{\text{p}}[\text{F}] \quad k_{\text{W}}[\text{CO}_2][\text{DBF}])} \frac{(k_{\text{-F}}^{\text{p}}[\text{HCl}] \quad k_{\text{-F}}^{\text{d}}[\text{HCl}] \quad 2k_{\text{-W}})k_{\Delta}}{(k_{\text{-F}}^{\text{d}}[\text{HCl}] \quad k_{\text{-F}}^{\text{p}}[\text{HCl}] \quad 2k_{\text{-W}})k_{\Delta}} \neq 1, \quad (65)$$

which is equivalent to

$$(k_{\text{-F}}^{\text{p}}[\text{HCl}] \quad k_{\text{-W}})(k_{\text{-F}}^{\text{d}}[\text{F}] \quad k_{\text{W}}[\text{CO}_2][\text{DBF}]) \neq (k_{\text{-F}}^{\text{d}}[\text{HCl}] \quad k_{\text{-W}})(k_{\text{-F}}^{\text{p}}[\text{F}] \quad k_{\text{W}}[\text{CO}_2][\text{DBF}]), \quad (66)$$

and so

$$\frac{(k_{\text{-F}}^{\text{p}}[\text{HCl}] \quad k_{\text{-W}})(k_{\text{-F}}^{\text{d}}[\text{F}] \quad k_{\text{W}}[\text{CO}_2][\text{DBF}])}{(k_{\text{-F}}^{\text{d}}[\text{HCl}] \quad k_{\text{-W}})(k_{\text{-F}}^{\text{p}}[\text{F}] \quad k_{\text{W}}[\text{CO}_2][\text{DBF}])} \neq 1, \quad (67)$$

which is exactly the square root of K_{r} in equation (59) specialized for the experimental case.

4. On the magnitude of the stationary state current J^{ss}

We conclude this section by deriving the expression of the stationary state current displayed in equation (4) of the main text, which proves particularly useful to rationalize and discuss numerical simulations in the main text and in § VIB.

First of all, we rewrite the stationary state current defined in equation (31a) and equation (4) of the main text in terms of the pseudo rate constants introduced in equation (13*):

$$J^{\text{ss}} = \tilde{k}_{-\Delta} [1_{\text{H}}^{\text{D}}]_{\text{ss}} - \tilde{k}_{-\Delta} [1_{\text{D}}^{\text{D}}]_{\text{ss}} = \tilde{k}'_{-\Delta} [1_{\text{D}}^{\text{H}}]_{\text{ss}} - \tilde{k}'_{-\Delta} [1_{\text{H}}^{\text{H}}]_{\text{ss}}. \quad (68)$$

This rewriting is trivial, since $\tilde{k}_{-\Delta} = k_{-\Delta}$, $\tilde{k}_{-\Delta} = k_{-\Delta}$, $\tilde{k}'_{-\Delta} = k'_{-\Delta}$, and $\tilde{k}'_{-\Delta} = k'_{-\Delta}$ (see equation (13*)), but by plugging the analytical expression (27*) for stationary state concentrations into equation (68) we get

$$\begin{aligned}
 J^{\text{ss}} &= \frac{L_M}{\mathcal{N}} \left(\tilde{k}_{-\Delta} \sum_{t \in \mathcal{T}_{\text{ID}}} \prod_{\lambda \in t} \tilde{k}_{\lambda} - \tilde{k}_{-\Delta} \sum_{t \in \mathcal{T}_{\text{ID}}} \prod_{\lambda \in t} \tilde{k}_{\lambda} \right) = \\
 &= \frac{L_M}{\mathcal{N}} \left(\tilde{k}_{-\Delta} \sum_{t \in \mathcal{T}_{\text{ID}} \wedge \tilde{k}_{-\Delta} \in t} \prod_{\rho \in t} \tilde{k}_{\rho} \mathcal{P}_{\text{cw}} - \tilde{k}_{-\Delta} \sum_{t \in \mathcal{T}_{\text{ID}} \wedge \tilde{k}_{-\Delta} \in t} \prod_{\rho \in t} \tilde{k}_{\rho} \mathcal{P}_{\text{ccw}} \right) = \frac{L_M}{\mathcal{N}} (\mathcal{P}_{\text{cw}} - \mathcal{P}_{\text{ccw}}) = \\
 &= \frac{L_M \mathcal{P}_{\text{ccw}}}{\mathcal{N}} (K_r - 1) = \Gamma (K_r - 1), \tag{69*}
 \end{aligned}$$

where we introduced the positive quantity Γ appearing in equation (4) of the main text. All the other symbols appearing in the above equation have been defined in equations (27*) and (62). Equation (69) shows that the sign of the current J^{ss} is determined by K_r being smaller or greater than 1, with no current for $K_r = 1$, while its magnitude depends on the actual value of both K_r and Γ . Crucially, the value of Γ can be varied independently from the value of K_r , thus allowing for design modifications which can improve or stall a motor's performance (but not altering the direction of cycling) without varying its kinetic asymmetry. This kind of effects are explored in the main text and in the next section with the help of numerical simulations.

VI. NUMERICAL SIMULATIONS

In order to better understand the implications of the model for the design of molecular motors, we performed a series of numerical simulations solving numerically dynamical equations (10) while varying key rate constants and plotting the predicted effect on the stationary state current (J in equation (4) in the main text, J^{ss} in equations (31a) and (69)), the thermodynamic efficiency (η , equation (5) in the main text), and the quantity J/I_F (stationary state current divided by the rate at which fuel is consumed at the stationary state), which equates to the average number of cycles achieved per molecule of fuel (see equation (43)). As stated in § III A, this last measure is a coefficient of performance which is distinct from efficiency because it is independent of both the energy content of the fuel, instead looking at the effect of each fuel molecule, and of whether the energy is dissipated mechanically or chemically, only reflecting how well fuel use correlates with unidirectional movement. In all the simulations, thermodynamic consistency was always guaranteed by enforcing Wegscheider's conditions (36).

It should be noted that the current is unbounded and can be positive or negative, indicating forward or backward movement respectively, thermodynamic efficiency is bounded between 0 and 1, and J/I_F is bounded between 0.5 and -0.5 (the sign indicating direction). The value has a maximum of 0.5 because at minimum 2 molecules of fuel are required for a complete cycle as the rotary motor is a two-stroke design, with two barriers to move past per cycle.

A. Parameters

The parameters in Table I were adopted as reference in the numerical simulations of the model introduced in Section I. The chosen values are intended to be realistic and give a correct qualitative outcome. A precise quantitative analysis would require more experimental data and is out of the scope of the present work. Apart where explicitly said or shown in the graphs, reference parameters are kept constant in the numerical simulations.

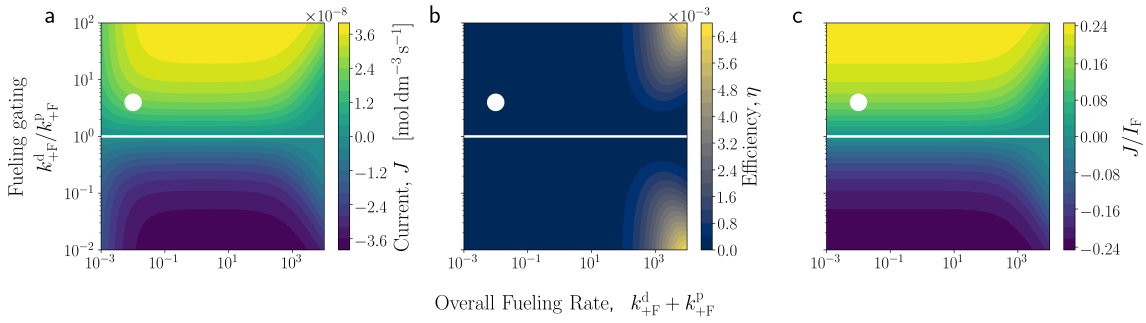
B. Supplementary analysis of the numerical simulations

1. Chemical gating of the fueling reaction

In this simulation, the sum of the rates of the fueling reactions and the chemical gating ratio of the fueling reaction were varied (Fig. 3a,b in the main text, Fig. 1). This is the only source of chemical gating as no chemical gating is introduced via the waste forming reactions in this simulation, as was the case in the experimental rotary motor (white dot). With no chemical gating (white line) there can be no current in the motor and therefore thermodynamic efficiency and the number of cycles per unit of fuel are also always zero. For intermediate rates of the fueling reaction, the current increases as the chemical gating increases. This is because an increased kinetic bias means the motor spends more time

Table I: Reference parameters employed in the numerical simulations. Rate constants k_{F}^{d} , k_{F}^{p} , k_{W} and k_{Δ} are estimated based on data collected in reference 1, where equations (1) hold, and related rotaxanes²⁷. Rate constants k_{F}^{d} and k_{W} are calculated by using equation (33) and assuming, based on known enthalpy-of-formation data, a difference of the order of 10^2 kJ mol^{-1} in the standard chemical potentials between fuel and waste species ($\mu_{\text{F}}^{\circ} - \mu_{\text{W}}^{\circ}$). The rate constant k_{F}^{p} is obtained by using equation (36a) to guarantee thermodynamic consistency. The total concentration of motor's species L_{M} is the same as in reference 1. The parameters involving chemostatted concentrations ($[\text{Fmoc-Cl}]$, $[\text{DBF}] \cdot [\text{CO}_2]$ and $[\text{HCl}]$) are estimated based on the original experiment¹ and other known data²⁸. The order of magnitude of parameters $[\text{DBF}] \cdot [\text{CO}_2]$ and $[\text{HCl}]$ does not significantly affect the results of simulations, since the microscopic reverses of fueling and waste-forming reactions (rate constants $k_{\text{F}}^{\text{p/d}}$ and k_{W}) are highly unfavorable.

k_{F}^{d}	$8 \cdot 10^{-3} \text{ mol}^{-1} \text{ dm}^3 \text{ s}^{-1}$	L_{M}	$1 \cdot 10^{-2} \text{ mol dm}^{-3}$
k_{F}^{d}	$1 \cdot 10^{-27} \text{ mol}^{-1} \text{ dm}^3 \text{ s}^{-1}$	$[\text{Fmoc-Cl}]$	$3 \cdot 10^{-2} \text{ mol dm}^{-3}$
k_{F}^{p}	$2 \cdot 10^{-3} \text{ mol}^{-1} \text{ dm}^3 \text{ s}^{-1}$	$[\text{DBF}] \cdot [\text{CO}_2]$	$1 \cdot 10^{-5} \text{ mol}^2 \text{ dm}^{-6}$
k_{F}^{p}	$2.5 \cdot 10^{-28} \text{ mol}^{-1} \text{ dm}^3 \text{ s}^{-1}$	$[\text{HCl}]$	$1 \cdot 10^{-18} \text{ mol dm}^{-3}$
k_{W}	$8 \cdot 10^{-6} \text{ s}^{-1}$		
k_{W}	$1 \cdot 10^{-26} \text{ mol}^{-2} \text{ dm}^6 \text{ s}^{-1}$		
k_{Δ}	$1 \cdot 10^1 \text{ s}^{-1}$		



Supplementary Figure 1: Graphs depicting the variation in current, efficiency, and J/I_{F} as overall fueling rate (x-axis) and fueling gating (y-axis) are changed. The white dot indicates the approximate parameters of the experimental rotary motor.

moving in the forward direction. A plateau is reached at higher values of chemical gating, indicating that other values (for instance, shuttling rate or waste formation rate) may become limiting to the current under these conditions. In this case, the waste-forming reaction is likely to be limiting, as it can be seen that the corresponding plateau in the ‘cycles per fuel molecule’ graph is at a value of approximately ± 0.25 . This is the theoretical limit (with perfect fueling gating) for a two stroke motor with no chemical gating of the waste-forming reaction (minimum two fuel molecules per cycle because there are two barriers, but half the fuel molecules will lead to no step because of a backwards waste-forming reaction).

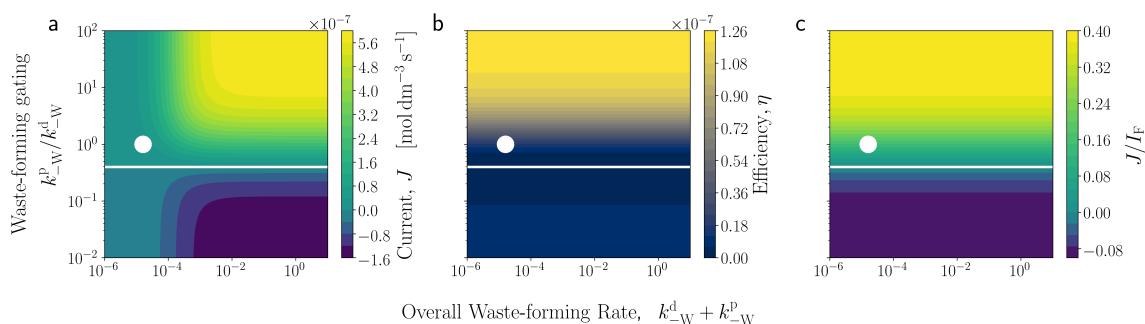
Very slow fueling rates (far left of the graphs) slow the motor to a near-stall, as the rate of reaction limits how fast the rotation can be. The graph of cycles per unit fuel shows that the motor is not actually stalled, just extremely slow as the motor still achieves the same number of cycles for each unit of fuel under these conditions. Less intuitively, very fast fueling rates (far right of the graphs) can actually stall the motor. This is because the fueling reaction is so fast that the mechanical distribution is effectively trapped as there is not sufficient time for mechanical equilibration before both co-conformers react. This promotes futile cycles, so fuel is still used despite less directional motion being achieved and therefore fewer cycles are performed for each fuel molecule used. Directional motion is still possible, but a greater chemical gating ratio is needed to achieve it. Under these conditions with a relatively very fast fueling reaction, a motor can be effectively stalled even if the chemomechanical cycle has kinetic asymmetry arising from the chemical gating.

The optimal thermodynamic efficiency balances the competing factors of a fast fueling rate, allowing the mechanical state to be pushed further from equilibrium, and fueling rates that are slow enough to limit futile cycles. The symmetric hotspots (thermodynamic efficiency is independent of direction) are therefore on the right of the graph, towards faster fueling rates where this balance is struck best. The current and J/I_{F} graphs are anti-symmetric around the horizontal axis with a chemical gating ratio of one (i.e., with the chemical gating ratio inverted, the same current and J/I_{F} are

achieved but in the opposite direction).

2. Chemical gating of the waste-forming reaction

In this simulation, the overall rate and chemical gating ratio of the waste-forming reaction were varied, while the chemical gating ratio of the fueling reaction was retained at a value of 4 (based on experimental measurements¹, Fig. 2, shown by the white dot in the figures). Therefore, even in the absence of waste-formation chemical gating, there is a



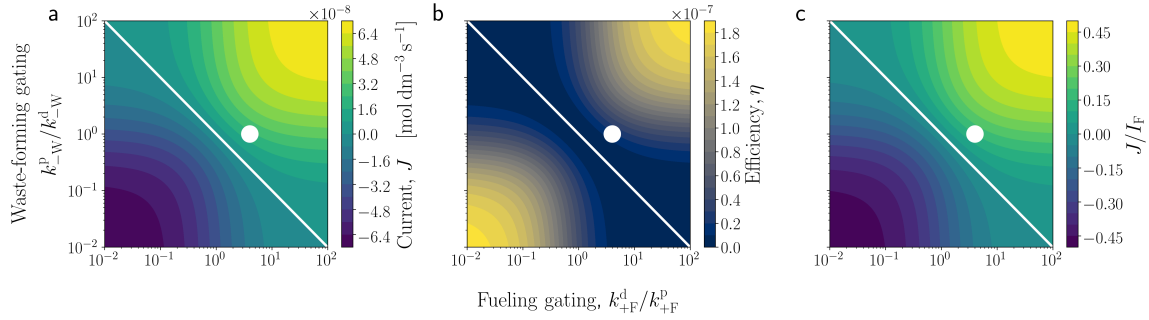
Supplementary Figure 2: Graphs depicting the variation in current, efficiency, and J/I_F as overall waste-forming rate (x-axis) and waste-forming gating (y-axis) are changed. The white dot indicates the approximate parameters of the experimental rotary motor.

driving force for the directional mechanical movement in the rotary motor from a positive information flow. Primarily this simulation shows the same as the fueling gating equivalent, while also showing that the waste-forming gating can cancel out the free energy transfer to the mechanical processes from the fueling gating and stall the motor. This gives the chemomechanical cycle no net kinetic asymmetry which is a condition synonymous with zero mechanical dissipation due to the absence of net free-energy flow between chemical and mechanical transitions. Where this direct negation occurs (white line), the current, efficiency and number of cycles per unit fuel are all inherently zero.

An interesting difference is that fast waste-forming reaction do not appear to cause a loss of current, efficiency, or cycles per fuel molecule. This is because the states preceding a waste-forming reaction (2_H or 2_D) are not able to undergo mechanical exchange and are therefore inherently always at mechanical equilibrium (precluding the mechanism by which fast fueling leads to a lower current, efficiency, and cycles per fuel). However, as the current model neglects the possibility for the rotary motor to lose both barriers, which is more likely to occur with faster waste-forming reactions, deviations from this behavior are expected. If this detail were included, it would be expected that faster waste-forming reactions would also lead to a loss of current, efficiency, and cycles per fuel.

3. Further exploration of chemical gating

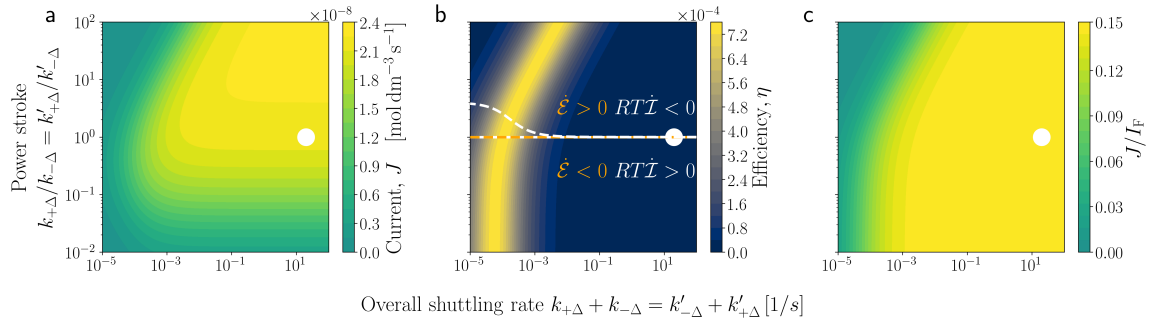
Better currents and efficiencies are theoretically obtainable with double-gated machines, in which both the fueling and waste-forming processes are chemically gated²⁹. A simulation in which the gating ratios of both reactions could be independently varied (Fig. 3) supports this idea, showing that the highest currents and efficiencies are predicted for motors in which both chemical processes are gated (top right and bottom left). Furthermore, no matter the magnitude of gating for one process, the motor would be stalled if the other process is biased in the opposite direction (white line, bottom right to top left of the graphs), leading to zero total kinetic asymmetry and precluding mechanical dissipation. Although the overall rates of the fueling and waste-forming reactions are different within the model, it appears that the contributions of fueling and waste-forming gating are symmetric and of equal importance for current, efficiency, and J/I_F . Additionally, this was the only simulation to approach the theoretical maximum limit of 0.5 cycles per fuel molecule, demonstrating that this requires large chemical gating values for both chemical processes. Under the conditions in which this limit is reached (top right and bottom left), it follows that further improvements to current and efficiency would have to come from other parts of the chemomechanical cycle. The limiting factor in this simulation is most probably the overall rates of the fueling and waste-forming reactions, though could plausibly be shuttling rate under other conditions.



Supplementary Figure 3: Graphs depicting the variation in current, efficiency, and J/I_F as chemical gating of the fueling reactions (x-axis) and the waste-forming reactions (y-axis) are changed. Shuttling rates and overall rates of the fueling and waste-forming reactions are kept constant at experimental values throughout. The white dot indicates the approximate parameters of the experimental rotary motor.

4. Introducing energy flow with power strokes without varying kinetic asymmetry

In this simulation, the shuttling rate and ratio of clockwise and anticlockwise shuttling rates were varied (Fig. 3c,d in the main text, Fig. 4), introducing energy flow via two power strokes. The simulation treats both halves of the rotary



Supplementary Figure 4: Graphs depicting the variation in current, efficiency, and J/I_F as overall shuttling rate (x-axis) and power stroke magnitude (y-axis) are changed. The power strokes are added in such a way as to keep the transition state energies and hence the kinetic asymmetry constant by compensating for changes to the mechanical transitions with the rates of the chemical transitions (see the Methods section of the main text). The white dot indicates the approximate parameters of the experimental rotary motor.

motor cycle (Fig. 2) as identical ($k_{\Delta} = k'_{\Delta}$, $k_{-\Delta} = k'_{-\Delta}$). As explained in the Methods section of the main text, in this simulation Wegscheider's consistency conditions (equations (36)) were imposed by varying rate constants k_W^p and k_F^p according to variations in shuttling rate constants, such that the kinetic asymmetry (K_r) of the chemomechanical cycle is not changed by the introduction of power strokes (effectively keeping the absolute transition state energies of the chemical transitions constant), which can nevertheless vary the positive factor Γ in equation (69). This allows us to explore changes caused only by power strokes to identify if, and how, power strokes have any effect on the behavior of the molecular motor.

For relatively slow shuttling rates (far left of the graphs) the motor is effectively stalled, with no significant current, thermodynamic efficiency, or cycles per unit fuel. As well as slow shuttling leading directly to low current, this situation is synonymous with the fast fueling reaction condition described above, in which the mechanical equilibration is slow enough that the mechanical distribution is trapped out by the relatively fast fueling reaction. Fast shuttling rates (far right of the graphs) do not appear detrimental to either the current or the number of cycles per fuel molecule, but thermodynamic efficiency with respect to how free energy is dissipated is strongly adversely affected. This is because fast shuttling hinders the generation of a concentration bias with respect to the equilibrium distribution in the mechanical states by more rapidly enabling equilibration. The position of optimal thermodynamic efficiency is determined by a trade-off between shuttling slow enough that the mechanical states are significantly far from their equilibrium distribution in the steady state, but fast enough not to too strongly promote futile cycles. It is particularly noticeable from the graphs that the optimal region of thermodynamic efficiency contours around the region in which J/I_F , the number of

cycles per fuel molecule, moves from a non-productive plateau (left of the graphs) to a productive plateau (right of the graphs). This is because slow shuttling decreases both efficiency and J/I_F , while the latter is independent of the path by which the energy from the fuel is dissipated. As a result, it is unaffected as the degree of energy dissipated through the mechanical processes decreases because of faster shuttling.

A key difference can be seen between graphs in the region describing negative energy flow from backwards power strokes (bottom, where $k_{\Delta}/k_{-\Delta} = k'_{\Delta}/k'_{-\Delta} < 1$, so clockwise shuttling requires gain of chemical potential). In this region, the thermodynamic efficiency and the number of cycles per unit fuel are independent of the magnitude of the power stroke and dependent only on the overall shuttling rate. Conversely, the current is strongly dependent on the power stroke magnitude. This arises because the current is directly reduced by the slow forward mechanical transitions associated with backward power strokes which limit the rate of the forwards chemomechanical cycles. However, efficiency and J/I_F are not affected as the decrease in the current (J) is offset by the decrease in the fuel consumption rate (I_F).

In the region with a positive energy flow arising from forwards power strokes, (top, where $k_{\Delta}/k_{-\Delta} = k'_{\Delta}/k'_{-\Delta} > 1$, so clockwise shuttling lowers standard chemical potential) all three quantities are dependent on both the overall shuttling rate and the magnitude of the power stroke. If the power stroke is too large or overall shuttling rate is too small (top and left of the graphs), futile cycles are favored, decreasing the current and thermodynamic efficiency and requiring more fuel per cycle on average. This is because the proximal species (1_D^H or 1_H^D) are sufficiently destabilized so that the unwanted proximal fueling reaction is faster than mechanical shuttling. It is especially unfavorable if a large power stroke and slow overall shuttling rate are combined.

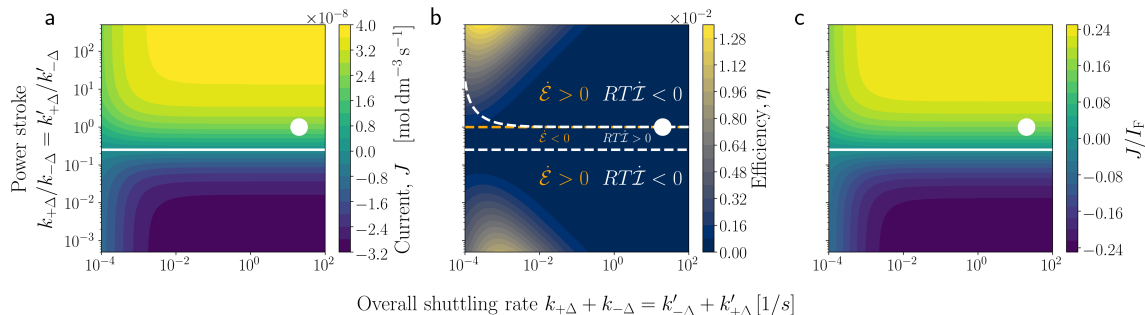
The maximum current is seen in a region with positive energy flow from forward power strokes but faster mechanical transitions (top right of the graphs). This is because, at steady state, the energy flow is increased by adding power strokes faster than information flow is decreased by biasing the mechanical equilibrium. This is a specific property of nonequilibrium regimes, where the standard chemical potentials and mutual information contributions to the total free-energy flow, which must inherently cancel at equilibrium, may differ in a nonequilibrium steady state. In mathematical terms we have, for the sum of energy and information flows:

$$\dot{\mathcal{E}} - RT\dot{I} = J(\mu_{1_D^D}^{\circ} - \mu_{1_D^H}^{\circ} - \mu_{1_H^D}^{\circ} + \mu_{1_H^H}^{\circ} - RT \log \frac{[1_D^D][1_H^H]}{[1_D^H][1_H^D]}) \geq 0. \quad (70)$$

At mechanical equilibrium, both the current J and the quantity in parenthesis are individually zero, since the concentrations are distributed accordingly to the exponential of the the difference in standard chemical potentials divided by RT . As soon as the system is driven out of equilibrium, the difference in standard chemical potentials stays constant (and positive in the region we are considering). In the presence of kinetic gating in the forward direction, the magnitude of the logarithm (negative, in the region we are considering) decreases. This is because, while the equilibrium favors 1_D^D over 1_D^H and 1_H^H over 1_H^D , the fueling kinetic gating is such to flatten this correlation, thus reducing mutual information. As a result, the mutual information contribution decreases towards more negative values slower than the standard chemical potential contribution increases towards positive values, so that the sum in the parenthesis is positive. In the presence of fast shuttling, the concentration distribution stays very close to the equilibrium one and the quantity in parenthesis is very small, thus making the efficiency very low. In other words, when the shuttling is fast, the mechanical transitions stay very close to equilibrium. This does not contradict the fact that we observe a high current in the stationary state, because ratio $k_{\Delta}[1_H^D]_{ss}/k_{-\Delta}[1_D^D]_{ss} = k'_{\Delta}[1_D^H]_{ss}/k'_{-\Delta}[1_H^H]_{ss}$ being close to 1 does not imply that the difference giving the current $J = k_{\Delta}[1_H^D]_{ss} - k_{-\Delta}[1_D^D]_{ss} = k'_{\Delta}[1_D^H]_{ss} - k'_{-\Delta}[1_H^H]_{ss}$ is small. Indeed, this is a clear example showing why the entropy production, not the current or J/I_F , is a measure of how far from equilibrium this system operates. Despite the increased current, the average number of cycles per fuel molecule remains unchanged between the high-current area (top right of the graphs) and the region with no or negative energy flow (bottom of the graphs), showing that only current is affected. In the same area, the efficiency plot shows that a very low percentage of the fueling free-energy is made available to the mechanical transitions, indicating that this regime, while optimizing current (J) and J/I_F , may not be optimal to perform work (e.g., transporting cargoes).

5. Introducing energy flow with a power stroke together with varying kinetic asymmetry

In contrast to the previous case, if Wegscheider's consistency conditions are imposed by varying rate constants k_W^P and k_F^P (and not k_W^P and k_F^P) according to variations in shuttling rate constants, the introduction of a power stroke will change the kinetic asymmetry of the chemomechanical cycle (Fig. 5). Unsurprisingly, the graphs produced for this scenario appear to display a combination of the features seen when power stroke magnitude (Fig. S.4) and kinetic asymmetry (Figs. S.1 and S.2) are altered independently. In the region with a positive power stroke and the associated larger kinetic asymmetry, the current and J/I_F increase. As with the fueling gating variation (Fig. S1) J/I_F , the number of cycles per fuel molecule, approaches the theoretical cap of ± 0.25 for a motor with an unbiased waste-forming reaction.



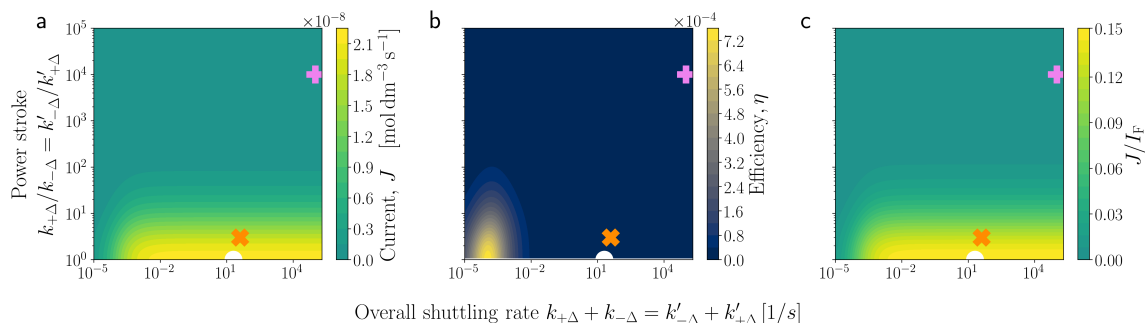
Supplementary Figure 5: Graphs depicting the variation in current, efficiency, and J/I_F as overall shuttling rate (x-axis) and power stroke magnitude (y-axis) are changed. The total power stroke magnitude is not compensated by changes in fuel addition and waste removal transition rates (k_F^P and k_W^P), therefore allowing the kinetic asymmetry to change. Instead, the transition rates of reverse processes (waste addition and fuel removal, k_W^P and k_F^P) are altered to retain Wegscheider's conditions and thermodynamic validity (see the Methods section of the main text). The white dot indicates the approximate parameters of the experimental rotary motor.

This is because the heights of the waste-forming transition states do not change (and hence remain identical to each other within the model) so any change in the power stroke is compensated by a corresponding change in the rates of the waste-forming reactions. The maximum efficiency in the region with positive power strokes is the highest seen in any simulation, appearing in the region with slower absolute shuttling for same reasons as previously discussed (see § VIB 1, VIB 2, and VIB 4), indicating that this combined approach, introducing power strokes with concomitant increase in kinetic asymmetry, may be an efficient design for molecular motors.

As with the simulations varying kinetic asymmetry by altering chemical gating of the fueling or waste-forming reactions, the direction of the rotary motor can be inverted with sufficiently negative power strokes. However, these graphs are not symmetrical and a lower maximum current and efficiency are seen in the negative current region as the chemical gating inherent to the system (*i.e.*, with no power strokes) favors a positive (clockwise) current.

6. Causing power strokes to cancel out

In this simulation (Fig. 6), shuttling rate and power stroke magnitude are, again, altered. This time, however, the



Supplementary Figure 6: Graphs depicting the variation in current, efficiency, and J/I_F as overall shuttling rate (x-axis) and power stroke magnitude (y-axis) are changed. The power strokes are introduced so that they cancel out (*i.e.*, energy release by the first mechanical transition is always mirrored by an identical energy gain in the second) and lead to no net energy flow within the chemomechanical cycle. The pink plus and orange cross indicate the properties expected from changing one fumaramide group to a succinamide. The pink plus indicates the values obtained from single-molecule experiments³⁰, while the orange cross shows the values found in a related rotaxane-based system³¹. The white dot indicates the approximate parameters of the experimental rotary motor.

direction of one power stroke is always opposite to the other while they retain the same magnitude, cancelling out each other's contribution. Therefore, no energy flow is introduced across the full chemomechanical cycle and kinetic asymmetry remains inherently unaltered. This is most easily realized by switching one binding site for a more weakly or

more strongly coordinating station. In the figures, the orange cross marks the predicted position of a motor equivalent to the experimental motor but with one station that binds the macrocycle 2.7 kJ mol^{-1} more strongly than the other, equating to a 3 : 1 bias in occupancy. This motor is predicted to operate at about 80% of the current of the original design as it appears futile cycles would be promoted and $\approx 20\%$ fewer cycles would be achieved, despite the same overall rate of fuel use. If the difference in binding strength is increased to 23 kJ mol^{-1} , the difference measured for the binding of the benzylic amide macrocycle used in the rotary motor to a fumaramide site compared to a succinamide site³⁰, the motor is effectively stalled, achieving no significant current, efficiency, or number of cycles per fuel molecule.

As before (see § VIB 1, VIB 2, VIB 4, and VIB 5), slow shuttling (left of the graphs) can be seen to stall the motor as futile cycles are favored, and once again, the area of maximum efficiency comes where this effect best balances with the degree to which the mechanical states are kept away from equilibrium in the steady state. Fast shuttling does not affect the current or the number of cycles per fuel molecule.

All graphs indicate that any change away from degenerate stations (moving up the y-axis) will always be detrimental for the motor in terms of current, efficiency, and the number of cycles per fuel molecule. This arises from a decreased capacity to support information flow in the motor as one station becomes more favored, which in turn decreases the ability of the motor to sustain a current. The full details and origins of this effect are not fully explored in this current work and warrant further experimental and theoretical investigation.

7. Conclusion

While the above by no-means represents an exhaustive analysis of all possible design feature alterations, even for the specific small-molecule motor analyzed in this paper, several important considerations have been identified. Furthermore, as future experimental systems become available to study, we envisage that similar analyses will help establish more general design considerations for molecular motors. These may become invaluable when putting molecular motors to work and may help clarify the driving forces behind the functioning of biological molecular motors.

-
- ¹ Wilson, M. R. *et al.* An autonomous chemically fuelled small-molecule motor. *Nature* **534**, 235–240 (2016).
 - ² Laidler, K. J. *Chemical Kinetics* (HarperCollins Publishers, New York, 1987).
 - ³ Wachtel, A., Rao, R. & Esposito, M. Thermodynamically consistent coarse graining of biocatalysts beyond michaelis–menten. *New J. Phys.* **20**, 042002 (2018).
 - ⁴ Astumian, R. D. How molecular motors work – insights from the molecular machinist’s toolbox: the nobel prize in chemistry 2016. *Chem. Sci.* **8**, 840–845 (2017).
 - ⁵ Astumian, R. D. Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet. *Nat. Commun.* **10**, 3837 (2019).
 - ⁶ Das, K., Gabrielli, L. & Prins, L. J. Chemically-fueled self-assembly in biology and chemistry. *Angew. Chem. Int. Ed.* **60**, 20120 (2021).
 - ⁷ Avanzini, F., Falasco, G. & Esposito, M. Thermodynamics of non-elementary chemical reaction networks. *New J. Phys.* **22**, 093040 (2020).
 - ⁸ Rao, R. & Esposito, M. Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics. *Phys. Rev. X* **6**, 041064 (2016).
 - ⁹ Feinberg, M. Complex balancing in general kinetic systems. *Arch. Ration. Mech. Anal.* **49**, 187–194 (1972).
 - ¹⁰ King, E. L. & Altman, C. A schematic method of deriving the rate laws for enzyme-catalyzed reactions. *The Journal of Physical Chemistry* **60**, 1375–1378 (1956).
 - ¹¹ Hill, T. L. *Free energy transduction in biology* (Academic Press, New York, 1977).
 - ¹² Mirzaev, I. & Gunawardena, J. Laplacian dynamics on general graphs. *Bulletin of Mathematical Biology* **75**, 2118–2149 (2013).
 - ¹³ Schnakenberg, J. Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys.* **48**, 571–585 (1976).
 - ¹⁴ Mou, C. Y., Luo, J. & Nicolis, G. Stochastic thermodynamics of nonequilibrium steady states in chemical reaction systems. *The Journal of Chemical Physics* **84**, 7011–7017 (1986).
 - ¹⁵ Lebowitz, J. L. & Bergmann, P. New approach to nonequilibrium process. *Physical Review* **99**, 578–587 (1955).
 - ¹⁶ Katz, S., Lebowitz, J. L. & Spohn, H. Phase transitions in stationary nonequilibrium states of model lattice systems. *Phys. Rev. B* **28**, 1655–1658 (1983).
 - ¹⁷ Esposito, M. Stochastic thermodynamics under coarse graining. *Phys. Rev. E* **85**, 041125 (2012).
 - ¹⁸ Maes, C. Local detailed balance. *SciPost Phys. Lect. Notes* **32**, 1–17 (2021).
 - ¹⁹ Feng, Y. *et al.* Molecular pumps and motors. *J. Am. Chem. Soc.* **143**, 5569–5591 (2021).
 - ²⁰ Falasco, G. & Esposito, M. Local detailed balance across scales: From diffusions to jump processes and beyond. *Phys. Rev. E* **103**, 042114 (2021).
 - ²¹ Gaspard, P. Comment on “validity of path thermodynamics in reactive systems”. *Phys. Rev. E* **103**, 016101 (2021).

- ²² Petersen, J., Förster, K., Turina, P. & Gräber, P. Comparison of the h⁺/atp ratios of the h⁺-atp synthases from yeast and from chloroplast. *Proc. Natl Acad. Sci. USA* **109**, 11150–11155 (2012).
- ²³ Horowitz, J. M. & Esposito, M. Thermodynamics with continuous information flow. *Phys. Rev. X* **4**, 031015 (2014).
- ²⁴ Geertsema, E. M., van der Molen, S. J., Martens, M. & Feringa, B. L. Optimizing rotary processes in synthetic molecular motors. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 16919–16924 (2009).
- ²⁵ Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotechnol.* **13**, 882–889 (2018).
- ²⁶ Astumian, R. & Bier, M. Mechanochemical coupling of the motion of molecular motors to atp hydrolysis. *Biophys. J.* **70**, 637–653 (1996).
- ²⁷ Leigh, D. A., Wong, J. K. Y., Dehez, F. & Zerbetto, F. Unidirectional rotation in a mechanically interlocked molecular rotor. *Nature* **424**, 174–179 (2003).
- ²⁸ Shirono, K., Morimatsu, T. & Takemura, F. Gas solubilities (CO₂, O₂, Ar, N₂, H₂, and He) in liquid chlorinated methanes. *J. Chem. Eng. Data* **53**, 1867–1871 (2008).
- ²⁹ Borsley, S., Leigh, D. A. & Roberts, B. M. W. A doubly kinetically-gated information ratchet autonomously driven by carbodiimide hydration. *J. Am. Chem. Soc.* **143**, 4414–4420 (2021).
- ³⁰ Lussis, P. *et al.* A single synthetic small molecule that generates force against a load. *Nat. Nanotechnol.* **6**, 553–557 (2011).
- ³¹ Alvarez-Pérez, M., Goldup, S. M., Leigh, D. A. & Slawin, A. M. Z. A chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **130**, 1836–1838 (2008).

5 | CONCLUSIONS

The majority of synthetic chemical engines realized so far can be modeled as isothermal open and bipartite chemical reaction networks. The openness arises from the exchange of matter and/or light with the environment, which provides the input power necessary for their out-of-equilibrium functioning. The bipartite structure arguably arises from the typical experimental design of such systems. Indeed, most of the systems introduced in Chapter 1 (Section 1.2) combine two *orthogonal* kinds of reactions, which result in bipartite reaction networks – it is not by chance that most of them are usually represented with square schemes. Typically, reactions of one kind serve to harvest free energy from the environment (e.g., consumption of a chemical fuel, photoisomerizations, etc.) by mediating flows between two or more reservoirs. Reactions of the other kind are usually not coupled to any reservoir except the thermal bath, and correspond to mechanical movement in space or self-assembly. The fact that the latter kind of reactions can be kept out-of-equilibrium by just performing work on the former kind suggests that a sort of internal free energy transduction among the two groups of reactions happens, which allows free energy to be dissipated by degrees of freedom not directly coupled to the external force.

In this thesis, we substantiated the general idea given above, thus recovering a thermodynamic description of chemical engines which is similar in scope to the one given by classical thermodynamics for steam engines described in Chapter 1 (see Section 1.1). The resulting picture can be sketched as in Figure 11.

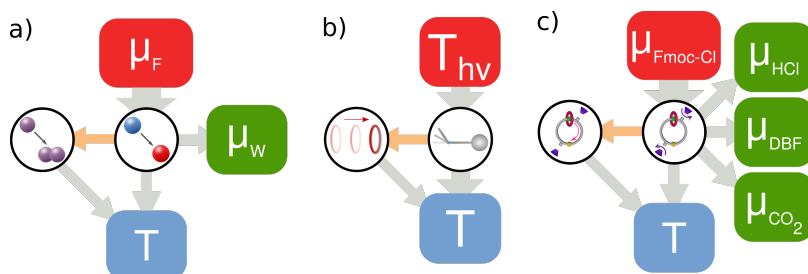


Figure 11: A thermodynamic viewpoint on chemical engines. Representation of the three chemical engines analyzed in this thesis as bipartite systems in contact with two or more reservoirs. Grey arrows show free energy or heat flows between reservoirs and the system, yellow arrows show internal free energy transduction between different sets of degrees of freedom in the system. a) Chemically-driven self assembly: chemical degrees of freedom are coupled with a fuel-to-waste conversion, self-assembly degrees of freedom are kept out-of-equilibrium by internal free energy transduction; b) Light-driven bimolecular pump: photoisomerization degrees of freedom mediate heat flow from the radiation (hot reservoir) and the thermal bath, directed motion of a macrocycle is sustained by internal free energy transduction; c) Catenane-based molecular motor: chemical degrees of freedom are coupled with fuel-to-waste conversion, directed shuttling of a macrocycle is sustained by internal free energy transduction.

To get there, we departed from chemical thermodynamics as it developed starting from the seminal contributions of Gibbs and De Donder. A simple but self-contained review of the main theoretical concepts leveraged by the entire thesis was given in Chapter 1 (Section 1.3) using a single double exchange reaction as guiding example. Despite the fact that several key ideas have historically emerged in the stochastic setup, we attempted a presentation fully grounded into mass action kinetics in order to make the content of the thesis independent – at least in principle – from previous knowledge of stochastic thermodynamics.

In Chapter 2, we established a general and systematic framework to treat the thermodynamics of arbitrary complex chemical reaction networks. Such a framework was previously developed by the group [1] based on precedent formulations [2, 3], but here we updated its presentation in view of subsequent progress specialized in the non-homogeneous and stochastic setups [4–6]. Furthermore, we extended the range of applicability of the framework to nonideal systems [7] and we used it – in its ideal version – to introduce different notions of thermodynamic efficiency for a minimalist model of chemically driven self-assembly [8] frequently employed to discuss general aspects of the phenomenon [9–12]. The extension to nonideal systems is a milestone towards the treatment of electrically driven systems, a frontier in the field of artificial chemical engines [13]. The analysis of chemically driven self-assembly demonstrated how our thermodynamic framework can be applied to chemical engines – regardless of nonlinearity in their dynamics – to get qualitative and quantitative insights on their functioning from an energetic viewpoint. Crucially, we could study the performance of the system in far-from-equilibrium regions unreachable using conventional linear regime thermodynamics.

In Chapter 3, we further extended the theory to include incoherent light as a source of free energy [14]. We did so by formulating the Marcelin-De Donder equation (Eq. (46)) for light-induced elementary photophysical processes (Eq. (26) on page 85) and by finding the proper thermodynamic potential for systems interacting with radiation. This allowed us to put light-induced reactions on the same ground of standard chemical reactions from a thermodynamic viewpoint. Then, by applying a thermodynamically consistent coarse-graining procedure specifically developed for chemical reaction schemes, we could find expressions for the currents emerging in photochemical reaction schemes which are directly comparable with widespread experimental models, whose thermodynamic properties are now clear. The usefulness of this connection has been sealed by the collaboration with the Credi group for the experimental and theoretical analysis of the second prototype of their light-driven bimolecular pump [15]. Indeed, by specializing our results to their experimental setup, we quantitatively probed that the degree of deviation from equilibrium (both in terms of entropy production and free energy stored) of the system when operated autonomously under continuous irradiation is correlated to the amount of power provided in form of light. Such kind of analysis goes beyond previous theoretical modelling of light-driven molecular machines and motors [16–18], which represent the majority of artificial chemical engines reported so far.

In Chapter 4, we formulated information thermodynamics for deterministic bipartite chemical reaction networks [19]. Since many artificial chemical engines are bipartite systems, this allowed us to refine our approach to their study by including an internal level of description. With this, we mean that the thermodynamics of chemical engines can be described in terms of

subnetworks concerning different kinds of reactions, which nevertheless interact by exchanging free energy via information and/or energy flows. The key steps to extend our framework in such direction have been introducing the notion of mutual information expressed in terms of the concentrations of bipartite species (Eq. (41) on page 134), showing that it can be expressed in terms of the Shannon-like contributions to the total entropy of the network (Eq. (42) on page 134), and finally specializing the second law for each subnetwork (Eq. (52) on page 135). Availing of the information thermodynamic level of description, we deepened the analysis of chemically driven self-assembly in various regimes, and we scrutinized the first experimental prototype of the light-driven bimolecular pump from the Credi group. Furthermore, we joined efforts with the Leigh group to perform a comprehensive analysis of their catenane based molecular motor that is general, complements previous analyses based on kinetics, and has practical implications for designing and improving synthetic molecular machines, regardless of the particular type of machine or chemical structure [20].

The approach developed in this thesis encompasses chemically driven and light driven reaction networks with possibly nonlinear and nonideal dynamics arbitrary far from equilibrium. It allowed us to characterize with unprecedented detail the thermodynamic aspects of three epitomes in the field of artificial chemical engines, uncovering a significant role for energy flow and information flow in the mechanism for the transduction of free energy from one set of reactions to another. We could also provide quantitative measures to compare the performance of different chemical engines at steady state while they perform no appreciable output work, as it is the case for most of the synthetic chemical engines made to date. As an example of interesting comparison we could draw among our case studies, we mention that the free energy gradient powering the light-driven pump analysed in Chapter 3 (Figure 11b) resulted comparable in magnitude to the one powering the catenane-based motor analyzed in Chapter 4 (Figure 11c), and one order of magnitude higher than that released by ATP to ADP conversion in typical physiological conditions. This suggests that there could be room for improving the efficiency of artificial chemical engines by powering them with smaller forces. As an example of practical implication for optimizing chemical engines' design, we mention that the presence of a positive energy flow (namely, the presence of power strokes in the mechanism of a chemical engine) can improve some features such as the efficiency. This may also help explain the role of power strokes in biomolecular machine mechanisms and thus contributes significantly to a fierce ongoing debate in this area [21, 22]. Among the most relevant theoretical results, we clarified that the information flow is precisely the thermodynamic counterpart of an information ratchet mechanism when no energy flow is involved – a pure Maxwell's demon regime – but one cannot draw a general one-to-one correspondence. We therefore concluded that the two concepts leverage two different notions of information, thus pertaining to distinct layers of description (a thermodynamic level and a kinetic one). Furthermore, in the specific case of the catenane-based rotary motor, we proved that the results about kinetic asymmetry of molecular motors [10, 23] obtained in the framework of kinetic models can be re-derived in our framework as a consequence of the second law for bipartite chemical reaction networks. Generally speaking, our findings provide a significant practical outcome that connects supramolecular chemistry to the thermodynamics of out-of-equilibrium systems and

illustrates the significance and impact that can arise from making such connections.

We conclude by mentioning some perspectives of the research work. After chemically driven and light driven systems, a complete treatment of electrically driven systems is needed to cover all the possible sources of nonequilibrium drivings for chemical engines. As already mentioned, a key ingredient towards this goal will be nonideality, which has to be combined with a spatial non-homogeneous framework [4], as electrical forces are generated by spatially separated electrodes. Two restrictions throughout the thesis have been focusing on autonomous systems operating at steady state and without performing any work on the environment. Concerning the first, a natural extension will therefore be the application of our approach to non-autonomously driven and oscillating systems which are of interest in supramolecular chemistry [24] and photochemistry [25, 26]. Concerning the second, computing the efficiency of those systems working against an external force/load [27–29] is within reach of our approach. The latter extension should bring us closer to an understanding of biochemical engines on the same grounds of synthetic ones, thus opening the possibility of direct comparisons between the two. Overall, the long standing goal is building a more and more complete framework able to guide experiments, quantify performance and possibly predict unforeseen behaviors of chemical engines.

REFERENCES FOR CHAPTER 5

- [1] R. RAO and M. ESPOSITO, “Nonequilibrium Thermodynamics of Chemical Reaction Networks: Wisdom from Stochastic Thermodynamics”, *Phys. Rev. X* **6**. (2016), 041064.
- [2] H. QIAN and D. A. BEARD, “Thermodynamics of stoichiometric biochemical networks in living systems far from equilibrium”, *Biophys. Chem.* **114.2** (2005), 213–220.
- [3] M. POLETTINI and M. ESPOSITO, “Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws”, *J. Chem. Phys.* **141.2** (2014), 024117.
- [4] G. FALASCO, R. RAO and M. ESPOSITO, “Information Thermodynamics of Turing Patterns”, *Phys. Rev. Lett.* **121**. (2018), 108301.
- [5] R. RAO and M. ESPOSITO, “Conservation laws shape dissipation”, *New J. Phys.* **20.2** (2018), 023007.
- [6] R. RAO and M. ESPOSITO, “Conservation laws and work fluctuation relations in chemical reaction networks”, *J. Chem. Phys.* **149.24** (2018), 245101.
- [7] F. AVANZINI et al., “Nonequilibrium thermodynamics of non-ideal chemical reaction networks”, *J. Chem. Phys.* **154**. (2021), 094114.
- [8] E. PENOCCHIO, R. RAO and M. ESPOSITO, “Thermodynamic efficiency in dissipative chemistry”, *Nat. Commun.* **10**. (2019), 3865.
- [9] G. RAGAZZON and L. J. PRINS, “Energy consumption in chemical fuel-driven self-assembly”, *Nat. Nanotechnol.* **13.10** (2018), 882–889.
- [10] R. D. ASTUMIAN, “Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet”, *Nat. Commun.* **10.1** (2019), 3837.

- [11] G. FALASCO et al., "Negative differential response in chemical reactions", *New J. Phys.* **21**. (2019), 073005.
- [12] K. DAS, L. GABRIELLI and L. J. PRINS, "Chemically Fueled Self-Assembly in Biology and Chemistry", *Angew. Chem. Int. Ed.* **60**.37 (2021), 20120–20143.
- [13] G. RAGAZZON et al., "Autonomous use of electrical energy by an artificial molecular machine", *ChemRxiv* **2021-16399**. (2021).
- [14] E. PENOCCHIO, R. RAO and M. ESPOSITO, "Nonequilibrium thermodynamics of light-induced reactions", *J. Chem. Phys.* **155**. (2021), 114101.
- [15] S. CORRA et al., "Kinetic and energetic insights into the dissipative non-equilibrium operation of an autonomous light-powered supramolecular pump", *Nature Nanotechnology* (2022).
- [16] E. M. GEERTSEMA et al., "Optimizing rotary processes in synthetic molecular motors", *Proc. Natl. Acad. Sci. U.S.A.* **106**.40 (2009), 16919–16924.
- [17] R. D. ASTUMIAN, "Optical vs. chemical driving for molecular machines", *Faraday Discuss.* **195**. (2016), 583–597.
- [18] A. SABATINO et al., "Individual-Molecule Perspective Analysis of Chemical Reaction Networks: The Case of a Light-Driven Supramolecular Pump", *Angew. Chem. Int. Ed.* **58**. (2019), 14341.
- [19] E. PENOCCHIO, F. AVANZINI and M. ESPOSITO, "Information Flows in Deterministic Chemical Reaction Networks", *arXiv* **2204.02815**. (2022).
- [20] S. AMANO et al., "Insights from an information thermodynamics analysis of a synthetic molecular motor", *Nat. Chem.* **14**. (2022), 530–537.
- [21] R. D. ASTUMIAN, "Irrelevance of the Power Stroke for the Directionality, Stopping Force, and Optimal Efficiency of Chemically Driven Molecular Machines", *Biophys. J.* **108**. (2015), 291–303.
- [22] W. HWANG and M. KARPLUS, "Structural basis for power stroke vs. Brownian ratchet mechanisms of motor proteins", *Proc. Natl. Acad. Sci. U.S.A.* **116**.40 (2019), 19777–19785.
- [23] R. ASTUMIAN and M. BIER, "Mechanochemical coupling of the motion of molecular motors to ATP hydrolysis", *Biophys. J.* **70**.2 (1996), 637–653.
- [24] J. LEIRA-IGLESIAS et al., "Oscillations, travelling fronts and patterns in a supramolecular system", *Nat. Nanotechnol.* **13**.11 (2018), 1021–1027.
- [25] B. BORDERIE et al., "Nonlinear dynamics, multiple steady states, and oscillations in photochemistry", *J. Phys. Chem.* **96**.7 (1992), 2953–2961.
- [26] P. L. GENTILI, L. BALDINELLI and B. BARTOLOMEI, "Design of a new photochromic oscillator: towards dynamical models of pacemaker neurons", *Reaction Kinetics, Mechanisms and Catalysis* (2022).
- [27] Q. LI et al., "Macroscopic contraction of a gel induced by the integrated motion of light-driven molecular motors", *Nat. Nanotechnol.* **10**.2 (2015), 161–165.
- [28] J. T. FOY et al., "Dual-light control of nanomachines that integrate motor and modulator subunits", *Nat. Nanotechnol.* **12**.6 (2017), 540–545.
- [29] M. KATHAN et al., "A light-fuelled nanoratchet shifts a coupled chemical equilibrium", *Nat. Nanotechnol.* (2021).