PhD-FSTC-2019-28
The Faculty of Sciences, Technology and Communication

DISSERTATION

Defence held on 07/05/2019 in Esch-sur-Alzette

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN BIOLOGIE

by

Sébastien DE LANDTSHEER
Born on 06 August 1980 in Belgium

OPTIMIZATION OF LOGICAL NETWORKS FOR THE MODELING OF CANCER SIGNALING PATHWAYS

Dissertation defence committee

Dr Ing Thomas Sauter, dissertation supervisor
Professor, Université du Luxembourg

Dr Iris Berhmann, Chairperson
Professor, Université du Luxembourg

Dr Markus Morrison
Professor, Universität Stuttgart

Dr Bernhard Steiert
Researcher, Roche

Optimization of Logical Networks for the Modeling of Cancer Signaling Pathways

UNIVERSITY OF LUXEMBOURG

by Sébastien De Landtsheer

Supervisor:
Prof. Dr. Ing. Thomas Sauter
Thesis Committee Members:
Prof. Dr. Iris Behrmann
External Jury Members:
Prof. Dr. Markus Morrison
Dr. Bernhard Steiert

This doctoral thesis has been performed at the Systems Biology Group, Life Sciences Research Unit, University of Luxembourg, Luxembourg

under the guidance of

Prof. Dr. Ing. Thomas Sauter, Systems Biology Group, Life Sciences Research Unit, University of Luxembourg, Luxembourg

Affidavit

I hereby confirm that the PhD thesis entitled Optimization of Logical Networks for the Modeling of Cancer Signaling Pathways has been written independently and without any other sources than cited.

Luxembourg, _____

_____

Sébastien De Landtsheer

"It is the time you have wasted for your rose that makes your rose so important"

'Le Petit Prince', Antoine de Saint-Exupery

General Abstract

Cancer is one of the main causes of death throughout the world. The survival of patients diagnosed with various cancer types remains low despite the numerous progresses of the last decades. Some of the reasons for this unmet clinical need are the high heterogeneity between patients, the differentiation of cancer cells within a single tumor, the persistence of cancer stem cells, and the high number of possible clinical phenotypes arising from the combination of the genetic and epigenetic insults that confer to cells the functional characteristics enabling them to proliferate, evade the immune system and programmed cell death, and give rise to neoplasms. To identify new therapeutic options, a better understanding of the mechanisms that generate and maintain these functional characteristics is needed. As many of the alterations that characterize cancerous lesions relate to the signaling pathways that ensure the adequacy of cellular behavior in a specific micro-environment and in response to molecular cues, it is likely that increased knowledge about these signaling pathways will result in the identification of new pharmacological targets towards which new drugs can be designed.

As such, the modeling of the cellular regulatory networks can play a prominent role in this understanding, as computational modeling allows the integration of large quantities of data and the simulation of large systems. Logical modeling is well adapted to the large-scale modeling of regulatory networks. Different types of logical network modeling have been used successfully to study cancer signaling pathways and investigate specific hypotheses. In this work we propose a Dynamic Bayesian Network framework to contextualize network models of signaling pathways. We implemented FALCON, a Matlab toolbox to formulate the parametrization of a prior-knowledge interaction network given a set of biological measurements under different experimental conditions. The FALCON toolbox allows a systems-level analysis of the model with the aim of identifying the most sensitive nodes and interactions of the inferred regulatory network and point to possible ways to modify its functional properties. The resulting hypotheses can be tested in the form of virtual knock-out experiments. We also propose a series of regularization schemes, materializing biological assumptions, to incorporate relevant research questions in the optimization procedure. These questions include the detection of the active signaling pathways in a specific context, the identification of the most important differences within a group of cell lines, or the time-frame of network rewiring.

We used the toolbox and its extensions on a series of toy models and biological examples. We showed that our pipeline is able to identify cell type-specific parameters that are predictive of drug sensitivity, using a regularization scheme based on local parameter densities in the parameter space. We applied FALCON to the analysis of the resistance mechanism in A375 melanoma cells adapted to low doses of a TNFR agonist, and we accurately predict the re-sensitization and successful induction of apoptosis in the adapted cells via the silencing of XIAP and the down-regulation of NFκB. We further point to specific drug combinations that could be applied in the clinics. Overall, we demonstrate that our approach is able to identify the most relevant changes between sensitive and resistant cancer clones.

Acknowledgements

Firstly, I would like to thank my supervisor and mentor Prof. Dr. Ing. Thomas Sauter, for his continuous support during my PhD project and his contagious positivity. Many times his advice and directions helped guide this research project towards success. I would also like to thank the members of the thesis evaluation committee and the jury who accepted to assess the quality of this work.

This work was financed for the most part by the European Union's H2020 program, and for a lesser part by the Luxembourgish national research fund. I am deeply grateful for these funding opportunities and for having had the chance to participate in the projects.

A number of people had a significant influence on this work over the years, either by offering technical expertise, helping out with the implementation, exploring ideas, providing personal support, critical feedback on the advancement of the tasks, or just making my time more enjoyable. At the University of Luxembourg, I would like to thank Dr. Philippe Lucarelli for all the help and the many ideas, Dr. Panuwat Trairatphisan, Dr. Thomas Pfau, Dr. Luana Presta, Dr. Maria Pacheco, Dr. Phuong Nguyen, Dr. Lasse Sinkkonen and all the members of the Systems Biology group who provided valuable feedback. At the University College Dublin, I would like to thank Dr. Walter Kolch, Dr. Boris Kholodenko, and all members of the Systems Biology Ireland group. At Merrimack Pharmaceuticals in Boston, I would like to thank Dr. Birgit Schoeberl, Dr. Andreas Raue, Dr. Wendy Qiao and Dr. Marco Muda.

I would also like to thank a handful of people have helped me going through the ups and downs of a doctoral thesis. Axel, Vincent, my father Alain, my sister Amélia: this would not have been possible without you!

I was lucky enough to be part of the MELPLEX training network, and I made there a number of incredible friends and colleagues. Alba, Anna, Biswajit, Cristiano, Estefanìa, Francesca, Greta, Isabela, Jan, Marco, Neta, Nicole, Romina, Valérie, Vesna, Živa: thanks for all the precious fun!

I am especially grateful to my former math teacher Willem de Bruijn for insisting that I start a PhD in the first place.

Lastly, I would like to express my gratitude to my wife Sara and my two children Alessandro and Leonardo, for challenging me, cheering me up, and reminding me the important things of life when I needed it the most.

Contents

List of Figures

List of Tables

Abbreviations

A-D	Anderson-Darling test
AIC	Akaike Information Criterion
AKT	Protein kinase B
AMP	Adenosine Monophosphate
APAF-1	Apoptotic Protease Activating Factor 1
ATF	AMP-dependent Transcription Factor 1
ATP	Adenosine Triphosphate
BC	Before Current era
BCL-2	B-Cell Lymphoma 2
BCR	Breakpoint Cluster Region protein
BIC	Bayesian Information Criterion
BN	Boolean Network
BRAF	v-Raf Murine Sarcoma Viral Oncogene Homolog B
CD	Cluster of Differentiation
CPU	Central Processing Unit
CTLA4	Cytotoxic T-Lymphocyte Associated Protein 4
DBN	Dynamic Bayesian Network
DNA	Desoxyribonucleic Acid
DPK	Dendritic Cell-Derived Protein Kinase
EGFR	Epidermal Growth Factor Receptor

ER	Estrogen Receptor
ERK	Extracellular Signal-Regulated Kinase
FDA	Food and Drugs Agency
FLIP	FLICE-Like Inhibitory Protein
GABA	Gamma-Aminobutyric Acid
GB	Gigabyte
GHz	Gigahertz
GTPase	Guanosine Triphosphate Hydroxylase
GUI	Graphical User Interface
HSF-1	Heat Shock Factor 1
I/O	Input/Output
IκBα	NF-Kappa-B Inhibitor Alpha
IκBSR	IκB Super-Repressor
JAK	Janus Kinase
JNK	c-Jun N-terminal Kinase
JUN	Jun Proto-Oncogene
K-S	KolmogorovSmirnov test
LASSO	Least Absolute Shrinkage and Selection Operator
MAPK	Mitogen-Activated Protein Kinase
MEK	Mitogen-Activated Protein Kinase Kinase 1
mRNA	Messenger RNA
MSE	Mean Squared Error
mTOR	Mammalian Target of Rapamycin
mTORC	Mammalian Target of Rapamycin Complex
NFκB	Nuclear Factor Kappa B
ODE	Ordinary Differential Equation

PARP	Poly-ADP-Ribose Polymerase
PBN	Probabilistic Boolean Network
PD-1	Programmed Cell Death Protein 1
PD-L1	Programmed Cell Death Ligand 1
PDGF	Platelet-Derived Growth Factor
PDGFR	Platelet-Derived Growth Factor Receptor
PI3K	Phosphatidyl-Inositol-Triphosphate Kinase
RAF	Rapidly Accelerated Fibrosarcoma Kinase
RAM	Random-Access Memory
RNA	Ribonucleic Acid
RIPK3	Receptor Interacting Serine/Threonine Kinase 3
RSK	Ribosomal S6 Kinase
SBML	Systems Biology Markup Language
SMAC	Second Mitochondria-Derived Activator Of Caspase
STAT	Signal Transducer And Activator Of Transcription
T-LGL	T-Cell Large Granular Lymphocyte Leukemia
TNFR	Tumor Necrosis Factor Receptor
TRAIL	TNF-Related Apoptosis Inducing Ligand
XIAP	X-Linked Inhibitor Of Apoptosis

Chapter 1

General Introduction

1.1 Cancer

1.1.1 Epidemiology

Cancer is not a single disease but a large group of diseases having some common features and usually considered together. Neoplasia, the phenomenon of uncoordinated cell growth in a tissue, gives rise to neoplasms, which usually present as solid masses of de-differentiated cells. Neoplasms can be benign, growing slowly and without invading the surrounding tissue, in which case they are usually not life-threatening and can be surgically removed. But they can also be malign, growing aggressively, invading the surrounding tissues and spreading to other organs, in which case they are referred to as *cancer* (Alberts *et al.*, 2002). Malign tumors can arise from virtually every cell type, but are more frequent in tissues with a high rate of cell division (Tomasetti & Vogelstein, 2015). Depending their cell type of origin, cancers are classified as carcinomas (from epithelial cells), leukemia and lymphoma (hematopoietic cells), sarcomas (mesenchymal cells), or blastomas (precursor cells and embryonic tissue). Additional cancer types are given specific names, for example seminoma and dysgerminoma (pluripotent germ cells), and melanoma (skin melanocytes). In 2016, more than 17 million people globally were diagnosed with a form of cancer, and nearly 9 million people died from it, costing nearly 210 million disease-adjusted life-years (Fitzmaurice & The Global Burden of Disease Cancer Collaboration, 2017). Figure 1 shows the annual number of deaths globally by cause, and compares the evolution of five-year survival in USA for different types of cancer over four decades, highlighting the poor prognosis for many types of cancer.

As normal cells evolve towards the neoplastic state, they acquire a succession of specific capabilities which materialize the need of cancer cells to acquire the traits that enable them to become tumorigenic and ultimately malignant (Hanahan & Weinberg, 2011). These capabilities allow them to sustain a proliferative behavior, evade growth-restricting signals, resist cell death, replicate endlessly, induce angiogenesis, and activate metastasis.

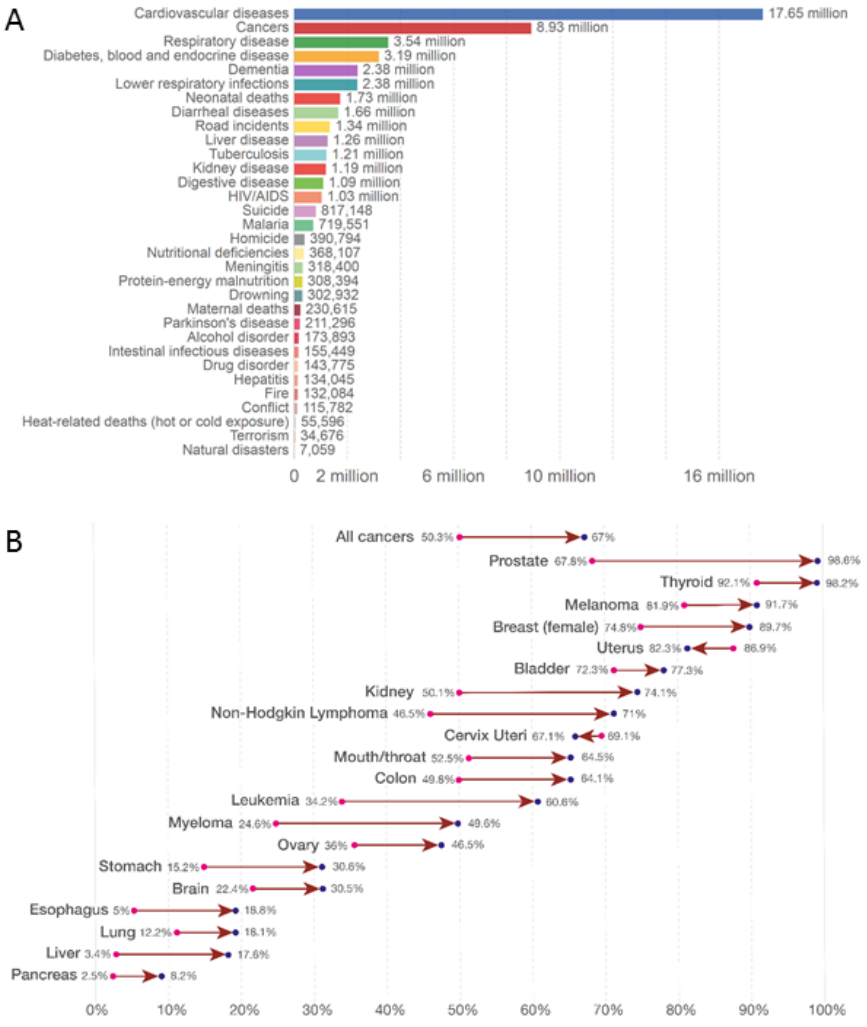


FIGURE 1: Overall cancer statistics. A) Worldwide annual number of deaths by cause. Cancer is second to cardiovascular disease only, claiming nearly 9 million lives. Data refers to the specific cause of death, which is distinguished from risk factors for death, such as pollution, diet and other lifestyle factors. B) Five-year cancer survival rates in the USA. Average five-year survival rates from common cancer types in the United States, shown as the rate over the period 1970-77 (●) and over the period 2007-2013 (●). (Roser & Ritchie, 2019)

1.1.2 Causes

The vast majority of cancers is due to environmental factors, while less than one-tenth of all cancers are due to genetics alone (Anand *et al.*, 2008). Some of these factors are difficult to control, however most are related to lifestyle. Diet is the largest cancer risk factor: 70 % of colorectal cancer deaths are linked to diet, and about a third of cancers overall. It is believed that the compounds synthesized during the cooking of red meat, nitrates and nitrites used for meat conservation, but also pesticides, dioxins, bisphenol and other food additives play the largest role in diet-related cancers (Chao *et al.*, 2005). Tobacco use is a long-recognized risk factor: smoking tobacco is associated with increased risk of developing at least 14 types of cancer, particularly of the respiratory tract. Chronic alcohol consumption is a risk factor for cancers of the upper aero-digestive tract, including cancers of the mouth, lips, pharynx,

hypopharynx, larynx, and esophagus. Other risk factors include obesity, infection with carcinogeneous micro-organisms (mainly HBV, HCV, HIV, Helicobacter pylorii, herpesviruses), radiations (including UV from the sun and tanning beds), and environmental pollutants. Correspondingly, there are many ways to decrease one's risk to develop cancer. Adequate diet and level of physical activity, smoking cessation, low consumption of alcohol, and avoidance of damaging radiations and infections together decrease one's lifetime risk of developing cancer.

1.1.3 Genetics

Cancer results from the accumulation of genetic insults in a replication-competent cell, which acquires a selective advantage compared to its neighbors. Cancer cells will typically display different types of genetic changes: mutations, deletions, duplications and other copy-number variations, whole chromosomal duplications, and other epigenetic changes. The genes affected by mutations fall into two broad classes: oncogenes and tumor suppressors.

Oncogenes are genes that are involved in growth, proliferation, or inhibition of apoptosis (Croce, 2008). When upregulated or constitutively active, these genes provide proliferative signals and increase the chances of a tumor arising. Different classes of genes can act as oncogenes: growth factors (for example *PGDF*), growth factor receptors (*EGFR*, ...), signal transducers (*ABL*, *SRC*, *RAF1*, ...), apoptosis inhibitors (*BCL2*, ...), transcription factors (*EWS*, ...), adhesion molecules (*FAK*, ...). Nearly all of targeted therapeutic drugs are inhibitors of oncogenes.

Tumor suppressors are genes involved in negative regulation of proliferation, control of the cell cycle, or initiation of apoptosis. Their existence was formulated following statistical studies indicating a second, recessive mechanism of tumorigenicity in retinoblastoma (Knudson, 1971), and the first tumor suppressor to be cloned was the *RB1* gene (Friend *et al.*, 1986). Tumor suppressors broadly fall into two classes: caretakers, who are implicated in DNA repair pathways, telomere metabolism and cell-cycle checkpoints, and gatekeepers, who frequently control apoptosis. In contrast with oncogenes, mutations of tumor suppressor genes reduce their activity, thereby reducing the function they serve (Morris & Chan, 2015).

Because mutations in oncogenes are activating, the presence of only one mutated allele is sufficient to induce the corresponding phenotype. Mutations in tumor suppressor genes, in contrast, need to be present on both alleles, and cancer cells frequently delete or inactivate whole genomic regions resulting in the absence of a second allele for the mutated tumor suppressor. Some genes, however, are extremely dosage-dependent, and therefore can manifest as tumor suppressors in the presence of a single mutated copy. Although mutation patterns are cancer type-specific, the most frequently mutated genes in cancer genomes are tumor suppressors, for example *TP53*, *CDKN2A*, *BRCA1/2*, or *PTEN* (Vogelstein & Kinzler, 2004; Kandoth *et al.*, 2013).

1.1.4 Cancer signaling pathways

All cells must be aware of their environment and respond to external changes in a coordinated manner. This is especially true in multi-cellular organisms. A number of signal transduction pathways have evolved to fill this function. Their mechanism is different depending of the case, for example the neurotransmitter GABA can bind to an ion channel which is able to directly modify the membrane potential, and the intracellular part of transmembrane receptors of the NOTCH family produces its effects via its action as transcription factor. Most of the information processing in eukaryotic cells, however, takes place via an intracellular network of proteins activating and inhibiting each other in an intricate network in which signals from different pathways is integrated.

Cancer cells frequently derive their proliferative phenotype from the over-activation of signal transduction pathways that control growth and proliferation. There is a number of mechanisms by which mutations can cause these pathways to be over-active. The most frequent one is activating mutations in the enzymatic domain of growth factor receptor tyrosine kinases (for example EGFR). Other mechanisms to affect signaling pathways are mutations in small GTPases like the Ras family of proteins, protein kinases like Src and Abl, lipid kinases like PI3K, or nuclear receptors like ER (Sever & Brugge, 2015).

1.1.4.1 Proliferation

The mitogen-activated protein kinases (MAPK) pathways are a family of signal transduction pathways found in all eukaryotes which regulate many cellular functions, including proliferation, tissue growth, gene expression, differentiation, survival and apoptosis. They consist of a hierarchy of serine/threonine-specific protein kinases. External stimuli activate the kinases at the top of the hierarchy (MAP3K), usually at the membrane via members of the Ras family of small GTPases, and the activating signal is transfered to MAP2K then to MAPK in the form of successive phosphorylations. Several 'classical' MAPK pathways exist, notably the RAF-MEK-ERK pathway, the P38 pathway, and the JNK pathway, as well as a number of less-studied, 'atypical' pathways. Activated MAPKs (ERK1/2, p38α, JNK) relocalize to the nucleus and activate a number of transcription factors. Among the transcription factors known to be activated by MAPKs are c-Fos, c-Myc, STAT3, RSK kinases, c-Jun, ATF1/2/6, p53, NFκB, HSF-1, and many more controlling a vast array of cellular functions (Cargnello & Roux, 2011). It must be noted that some MAPK targets are cytoplasmic (paxillin, DPK, RSK), and that it is likely that many targets remain to be identified.

The PI3K/AKT/mTOR pathway is an important pathway regulating the cell cycle. The mTOR protein (mammalian target of rapamycin) is the core component of two protein complexes, mTORC1 and mTORC2, which regulates cell growth, cell proliferation, cell motility, survival, and autophagy (Sabers *et al.*, 1995; Wullschleger *et al.*, 2006). These two complexes

integrate various signals, notably the presence of growth factors like IGF1, nutrients like amino acids, p53 activation, hypoxia, or the intracellular AMP/ATP ratio. The PI3K/AKT/mTOR pathway controls translation regulators S6K1 and 4E-BP1, interacts with all three RNA polymerases via the URI protein (unconventional prefoldin RPB5 interactor), and influences autophagy (Wullschleger *et al.*, 2006; Lum *et al.*, 2005). Importantly, the ERK1/2 pathway and the PI3K-mTOR pathway compensate each other and cross-talk in multiple points, which poses particular challenges for cancer therapy (Mendoza *et al.*, 2011).

The JAK/STAT pathway is a central signaling pathway, very conserved in all vertebrates, controlling many processes that are important for cancer development, like stem cell maintenance, embryogenesis, and the inflammatory response. Different types of transmembrane receptors (notably G-CSF, and IL-6R) are associated with inactive janus kinases (JAKs), which upon ligand binding undergo conformational changes resulting in the recruitment and phosphorylation of STAT proteins, which then dimerize and translocate to the nucleus and modulate the transcription of many genes related to differentiation, proliferation, and apoptosis (Rawlings, 2004). The role of mutations in JAK genes (notably JAK2) in many types of chronic and acute leukemias has been long recognized (Chen *et al.*, 2012). In solid tumors, STAT activation has been linked to both improved and reduced overall survival, and the multiple ways in which the JAK-STAT pathway cross-talks and interacts with the other main proliferative signaling pathways makes this pathway a target for the development of new therapies. However, while JAK inhibitors (e.g. ruxolitinib and tofacitinib) have shown clinical success in myelofibrosis and rheumatoid arthritis, the inhibition of JAKs and STATs in patients with solid tumors has not shown success beyond pre-clinical models (Thomas *et al.*, 2015). Figure 2 schematizes the above-mentioned signaling pathways and shows their inter-connections.

1.1.4.2 Programmed cell death

Apoptosis is the main form of programmed cell death, and results in a coordinated cell shrinkage, nuclear and DNA fragmentation, and mRNA decay (Alberts *et al.*, 2002). Apoptosis is an essential process for all metazoans, in that it maintains an healthy balance between cell survival and death. Apoptosis plays an important role in embryogenesis and organ morphogenesis. Insufficient apoptosis plays a role in cancer and autoimmunity, while enhanced apoptosis is implicated in acute and chronic degenerative diseases, immunodeficiency, and infertility. There are two activation pathways, both highly regulated, depending the origin of the apoptotic signal. The intrinsic pathway integrates different stress signals (heat, radiations, hypoxia, viral infections, and others) at the level of mitochondria (via the release of cytochrome c, which binds to APAF-1), while the extrinsic pathway is triggered by extracellular signals via receptors of the TNFR family. They converge in the activation of caspases, a family of cross-activating proteases of which there exist 11 or 12 in humans (functional CASP12 is only expressed in some populations (Saleh *et al.*, 2004)) which leads to degradation of cellular components.

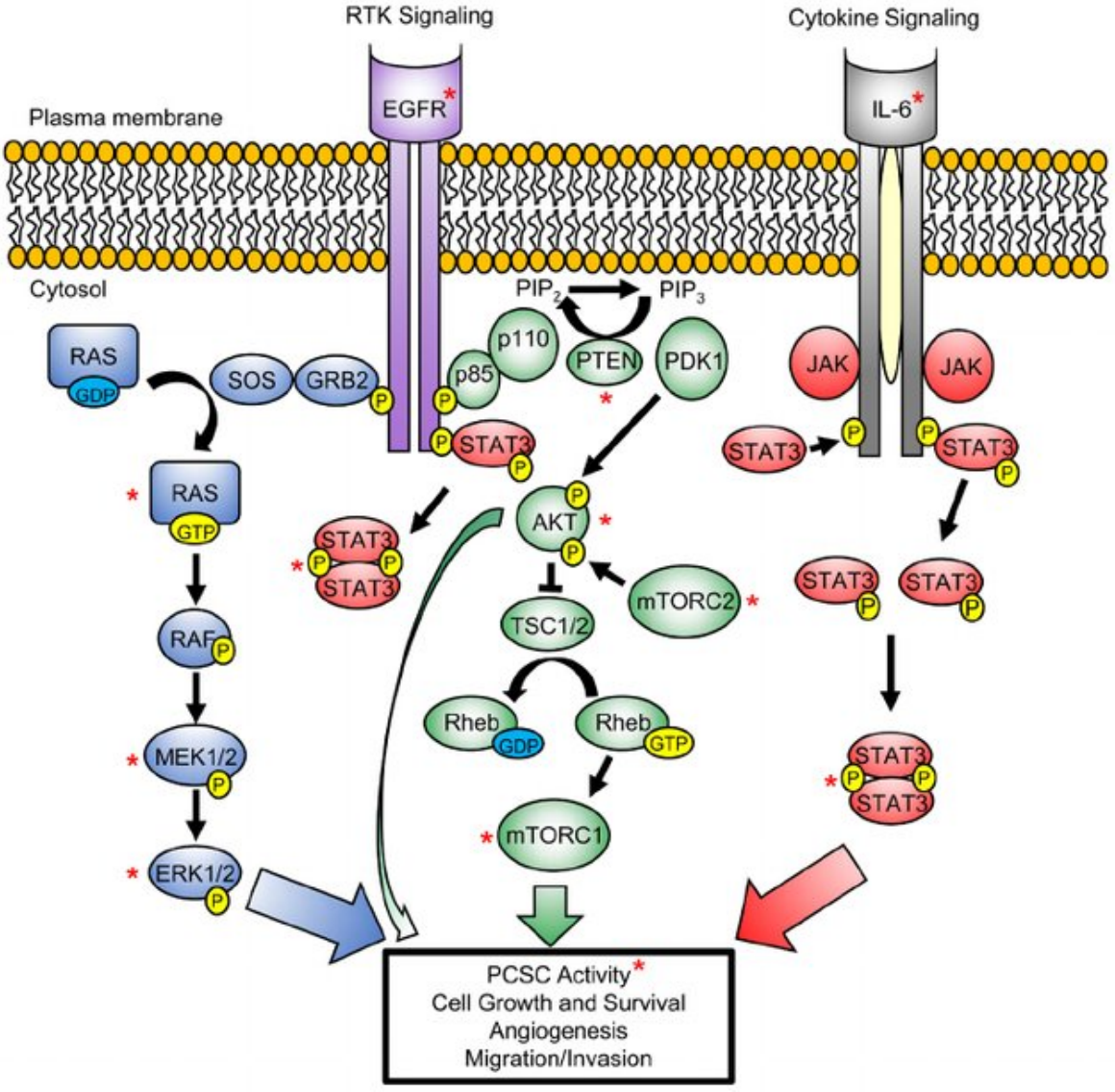


FIGURE 2: PI3K/AKT, RAS/MAPK and STAT3 signaling pathways converge to regulate and promote tumorigenesis. Activation of PI3K/AKT (green), RAS/MAPK (blue) and STAT3 (red) signaling pathways, mediated by the activation of growth factor-driven receptor tyrosine kinase (RTK) (e.g. Epidermal growth factor receptor (EGFR)) or cytokine (e.g. IL-6) signaling, promote cancer stem cells self-renewal activity and the various hallmarks of cancer development. These signaling pathways act directly, or through cross-talk activation, to mediate tumorigenesis. P denotes phosphorylation of protein at specific residue(s), which is required for its activation (yellow). Red asterisks (*) marks key proteins within these signaling pathways that have been implicated in cancer cell activity. Modified from Rybak *et al.* (2015).

Of note, there are several other forms of cellular death that are specifically mediated by an intracellular program. The most relevant to human health is probably autophagy, a cellular process of adaptive response to stress in which cells break down and recycle part of themselves, if there are excess or dysfunctional components or if there is severe nutrient shortage (Mizushima & Komatsu, 2011; Takeshige *et al.*, 1992). Recently, other forms of programmed

cell death were discovered: necroptosis (RIPK3-driven necrosis) (Linkermann, 2014), paraptosis (Sperandio *et al.*, 2000), or pyroptosis (infection-related suicide of immune cells) (Fink & Cookson, 2006).

1.1.5 Treatments

Depending on the nature and the site of the tumor, different options exist to treat patients with cancer. Surgical removal of the solid tumor, when possible, is usually the best option. Radiation therapy, in which high-energy X-rays beams are used on patients, works by damaging DNA of cancer cells. Indeed, cancer cells usually have a deficient DNA damage repair system, while normal cells have the ability to repair themselves after moderate exposure to radiation.

Surgery and radiation therapy rarely suffice to completely cure a cancer. Cytotoxic chemotherapy is the treatment of cancer with drugs. Different types of chemotherapeutic drugs exist, but in contrast with targeted therapies (see below) they all act in a somewhat unspecific way, and so also impact healthy cells. Alkylating agents (for example cisplatin, procarbazide, cyclophosphamide, ...) were the first anti-neoplastic agents to be developed (Scott, 1970). They act by binding covalently to DNA via their alkyl group, and in doing so induce breaks in the DNA strands that cannot be repaired by the cells, which then undergo apoptosis. Antimetabolites (for example methotrexate, fluorouracil, gemcitabine, ...) are drugs structurally similar to nucleotides that block enzymes involved in DNA synthesis. They also sometimes incorporate into DNA, inducing DNA damage and triggering apoptosis. Mitotic inhibitors (vincristine, vinblastine, paclitaxel, ...) bind to tubulin during certain phases of the cell cycle, therefore preventing cells to divide. Topoisomerase inhibitors (irinotecan, doxorubicin) work by inhibiting topoisomerase I and II, preventing both replication and transcription. Other drugs, like bleomycin and actinomycin, act by a mix of different mechanisms of action.

Because of their non-specific mechanisms of actions, cytotoxic chemotherapies affect all fast-replicating tissues, like hair follicles, skin, bone marrow and the gastro-intestinal tract, resulting in considerable side-effects for cancer patients (Pearce *et al.*, 2017). For this reason, they are usually administered in comprehensive, dose-adjusted regimen, and often in multi-drug combinations. Different cancer types will respond to these drugs differently, with some patients not responding to any (Ashdown *et al.*, 2015).

In addition to classical chemotherapies, advances in the understanding of the mechanisms of tumor growth have given rise to targeted therapies. The concept behind these therapies is that by specifically targeting the cancer cells and sparing the non-cancerous cells, higher doses can be given to patients in a safe way. The first treatment that was targeted to the molecular characteristics of a tumor was ^{131}I therapy for thyroid cancer (Weigel & Mcdougall, 2006), which relies on the fact that thyroid cells will take up the iodine and that radioactivity will accumulate in the cancer cells and induce apoptosis. Another early success was the tyrosine

kinase inhibitor imatinib, which inhibits the BCR-Abl hybrid signaling protein associated with the Philadelphia chromosome, in chronic myeloid leukemia cells (Deininger & Druker, 2003).

Strategically targeting the mechanisms by which cancer grows and persists has been a major trend over the last two decades. During this time, a number of small-molecule inhibitors have been developed to inhibit the drivers of oncogenesis or the signaling pathways they activate. This strategy produced a large number of tyrosine kinase inhibitors like gefitinib, erlotinib, sorafenib, trametinib, and vemurafenib, proteasome inhibitors like bortezomib and ixazomib, and also a number of monoclonal antibodies like EGFR-specific cetuximab, CD20-specific rituximab, or ErbB2-specific trastuzumab.

In addition, two classes of completely novel therapies have appeared that have changed the landscape of cancer treatments: immune checkpoint inhibitors and oncolytic viruses. Immune checkpoints refer to the multiple mechanisms by which the immune system regulates itself, either in a stimulating or inhibiting manner, in order to maintain self-tolerance. These mechanisms can be hijacked by cancer cells in order to evade the body's immune response to tumors. Current approved immune checkpoint inhibitors interfere with the CTLA4, PD-1 and PD-L1 surface receptors (Pardoll, 2016). The idea of treating cancerous lesions with viruses originates in the fortuitous discovery that in rare instances, tumor regression can be observed in cases of acute infections (Jessy, 2011). Famously, William B. Coley obtained excellent results by injecting live and attenuated bacteria directly into the tumors of his patients (Coley, 1891). After decades of research into engineered oncolytic viruses, a number of viral therapies are currently in clinical trial, and recently a modified herpes virus has been approved by the FDA for the treatment of melanoma (Fukuhara *et al.*, 2016).

However, many challenges remain for the treatment of cancer patients. The recent progresses in the care of cancer patients, and the successes of cancer research for some cancer types (Hodgkin's lymphoma or testicular cancer for example) are shadowed by the high lethality of most solid tumors and the increasing incidence of cancer in most parts of the world (Siegel *et al.*, 2016). Therefore, new paradigms are needed to properly meet the needs of cancer patients. In order for these new paradigms (new pharmacological targets or drug combinations) to emerge, there needs to be a qualitative leap in our understanding of the molecular characteristics that enable cancer cells to behave the way they do. For example, it is estimated that less than 5% of cancer patients in US in 2018 benefited from molecular testing (Marquart *et al.*, 2018). The realization of the limitations of a purely reductionist approach has triggered increased interest in strategies that aim at integrate different strata of information, and formally identify high-level biological functions that emerge from integrated biological systems (Lazebnik, 2002).

Systems biology has emerged from the convergence of a number of disciplines: theoretical biology and the mathematical modeling of enzymatic reactions and population dynamics, cybernetics and control theory, systems theory, and bioinformatics. It attempts to provide a unified, holistic understanding of the complex processes that give rise to biological functions

in living organisms. The central paradigm of systems biology is to study how the interactions between the elements of a system give rise to emerging properties of the whole system (Van der leeuw, 2004). One of the first successes of systems biology is the Nobel-winning discovery that the propagation of the action potential in neurons was the emerging property of the ion channels in the membrane (Hodgkin & Huxley, 1952). In the words of Denis Noble: *"Systems biology ... is about putting together rather than taking apart, integration rather than reduction. It requires that we develop ways of thinking about integration that are as rigorous as our reductionist programmes, but different. ... It means changing our philosophy, in the full sense of the term."* (Noble, 2006).

1.2 Modeling

1.2.1 Molecular data modeling

1.2.1.1 Statistical and mechanistic modeling

There are a number of ways to form a model of a system. Models essentially provide an approximation of a system, in a way that illuminates a certain aspect of its inner workings. In the words of Prof. George E. P. Box: 'All models are wrong but some are useful' (Box, 1976). Models can be useful to understand, quantify, or predict a particular aspect of a system.

Perhaps the modeling framework most used in applied sciences is systems of differential equations. Basically, this type of equation defines the relationship of certain quantities of a system with their rate of change. While the simplest differential equations are solvable using known formulas, most are not, and their solutions (the set of functions that satisfy the system of equations) need to be approximated by numerical methods, the best known being the Taylor series expansion. By far most differential equations encountered in biology, physics and engineering are linear, ordinary, homogeneous differential equations, often termed ordinary differential equations (ODEs). These take the general form:

$$a_0(x)y + a_1(x)y' + a_2(x)y'' + \cdots + a_n(x)y^{(n)} + b(x) = 0$$

where $a_0(x)$,$\cdots$, $a_n(x)$ and $b(x)$ are differentiable functions that do not need to be linear, and $y', \ldots, y^{(n)}$ are the successive derivatives of the unknown function y of the variable x, although commonly used ODEs are usually of the first-degree form. This type of model is well suited for the study of systems that are well characterized and defined. Indeed, because it is a realistic, dynamic representation of the elements of the system, the number of parameters (decay rates, production rates, interaction strengths, initial concentrations, ...) in a large model poses problems both in terms of theoretical solvability and computational feasibility

(Bornholdt & Bornholdt, 2005). Because explaining the temporal dynamics of precise parts of some systems is important, ODE modeling is applied to small-to-medium systems and aims at inferring the behavior of sub-modules of entire systems (Kholodenko, 2006; Alon, 2007).

At the other end of the modeling spectrum, methods that make the least amount of assumptions are the ones related to machine learning. While this ill-defined term is applied to everything from linear regression to highly complex neural nets, it generally defines a statistics-based, prediction-oriented set of protocols for applying algorithms to datasets. This way, the absence of specific assumptions on the system allows the fast processing of huge amounts of data. Without task-specific modeling assumptions, general powerful protocols can be applied on data to produce 'learners', i.e. computer programs able to perform a specific task with data, improve with more experience, and generalize from the inputs of the dataset to new inputs with good performance. The tasks assigned to machine learning algorithms are typically the ones for which explicit programming would be impractical or infeasible, like targeted advertisement, computer vision or email filtering. Successful biological applications of machine-learning includes automated image analysis, and text mining for automated database management.

However, since the typical approach in applying machine learning algorithms to large datasets is prediction-driven, these algorithms have historically been seen as 'black boxes', in the sense that it is not easy to understand why they make the predictions that they make, although this view is increasingly challenged (Gilpin *et al.*, 2018), and is not true for some of the methods (like decision trees). This is because often, the best model in term of predictive power uses the input features in ways that are not intuitive for humans. For example, ResNet (He *et al.*, 2015), the best computer vision program to date, in its 50-layer deep residual convolutional architecture, uses more than 25 million parameters to achieve a better-than-human performance. However useful this can be in different automated applications, it does not inform about how humans perceive images. The same holds for biomedical applications, where decision-making needs to be as accurate and informed as possible. Therefore, large-scale machine learning techniques are not directly applicable to modern biomedical problems like cancer, mostly because they lack the power to explain mechanistically the functioning of the system under study.

Since both the detailed, mechanistic modeling approaches and the purely data-driven approaches have serious drawbacks, other modeling frameworks may be used to try to harness signal in biological data in the light of current molecular knowledge.

1.2.1.2 Network models

At the core of systems biology is the concept of network. It has been shown that the way the different elements of a biological regulatory network are arranged reveals patterns of connections that perform essential functions (Milo *et al.*, 2002). These network motifs are thought to constitute the building blocks of larger networks, giving rise to robust non-linear behaviors. Systems biologists are mostly preoccupied by the modeling of these large complex networks,

for example the distributed model of the whole cell (Ayyadurai & Dewey, 2011) or the reconstruction of the complete human metabolism (Thiele *et al.*, 2013).

Networks are representations of discrete objects and their relationships. As such, their study is part of graph theory, a field started nearly three centuries ago (Euler, 1736). Biological networks (as well as other natural networks like human collaboration networks and the Internet) are characterized by a number of features making them different from randomly constructed networks. Firstly, they display scale-free properties, meaning that both the in-degree and out-degree distributions follow a power-law (and not a normal) distribution. Secondly, they are 'small world' networks, meaning that the path length between any two randomly chosen nodes is shorter than in a random network (Albert, 2005). The study of biological networks across species and contexts reveals a number of common structural design principles responsible for basic functionalities (Milo *et al.*, 2002). Biological networks (metabolic, signaling, transcriptional) have been studied in the hope that understanding the way they function, for example by identifying hub genes and disease-related gene modules, would provide a way to identify better targets for drug development (Barabási, 2007; Barabási *et al.*, 2011). Another application is the prediction of the function of unknown proteins, as proteins performing related functions tend to cluster together in network models (Chakrabarty & Parekh, 2016).

At the interface of statistical and mechanistic models lie a series of intermediate modeling philosophies that aim at harvesting the benefits of both formalisms. These modeling frameworks rely on topology to inform the model with prior information about the system under study and specific mathematical frameworks to produce a model that is able to resume some of the functional properties of the whole system, and generate and test hypotheses. These frameworks offer an intermediate level of granularity between fully specified, mechanistic models and abstract models, and are usually referred to as *logical models*. Figure 3 schematizes the differences between abstract, statistical modeling and mechanistic modeling. Many of these frameworks have been applied with success to cancer research (Saez-Rodriguez *et al.*, 2011; Le Novère, 2015).

1.2.2 Logic and probabilities

1.2.2.1 Concepts

The study of logic is the study of truth, the general laws that govern it, and the arguments to establish it. Logicians study the fundamentally metaphysical specific relationship between causes and consequence, or between the assumptions of an inference and its conclusion. The first and by far most important body of work concerning logic is Prior Analytics, one of the six books of the *Organon* of Aristotle (384-322 BC). In this book, the three main principles governing logical reasoning are formulated: the law of identity ('Whatever is, is'), the law of non-contradiction ('Nothing can both be and not be') and the law of excluding middle

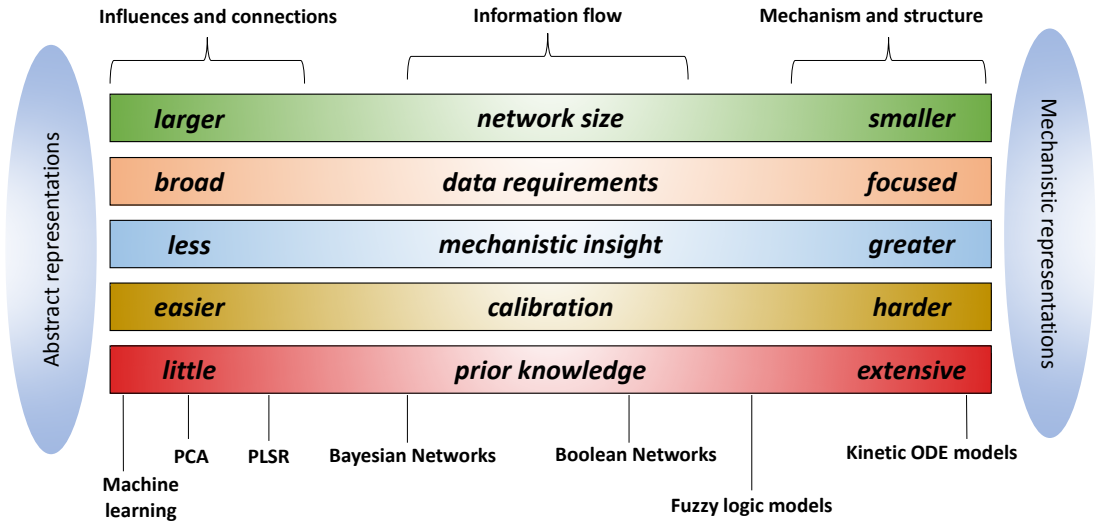


FIGURE 3: The spectrum of modeling formalisms. At the left side of the spectrum, statistical models provide broad information about large systems with low amounts of data. At the other side, mechanistic models like ODE systems are able to provide detailed knowledge, but typically require larger amounts of more focused data and are harder to calibrate. Adapted from Schivo *et al.* (2016)

('Everything must either be or not be'). These three principle are hypothesized to have laid the foundations on which knowledge accumulation becomes possible, and are at the basis of scientific method and most of Western philosophy, up to this day. However, he also posed several interesting paradoxes where a standard interpretation of the law of non-contradiction fails (Frede, 1985).

In the nineteenth century, self-taught mathematician George Boole formalized the use of logic in mathematical operations and introduced notation for logical operations (Boole, 1854). In Boolean algebra, in contrast with elementary algebra, variables and expressions do not represent numbers but logical propositions, and are assigned *truth values*, namely TRUE or FALSE, usually represented by the numbers 1 and 0 respectively. Boolean algebra defines three basic operations:

- NOT, denoted $\neg$, satisfies $\neg A = 1$ if $A = 0$ and vice-versa
- AND, denoted $\wedge$, satisfies $A \wedge B = 1$ if $A = B = 1$, and $A \wedge B = 0$ otherwise
- OR, denoted $\vee$, satisfies $A \vee B = 0$ if $A = B = 0$, and $A \vee B = 1$ otherwise

It should however be noted that either one of the AND and OR operations can be deduced from the other one following the De Morgan's laws:

- $A \wedge B = \neg(\neg A \vee \neg B)$

A	B	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}
T	T	F	F	F	F	F	F	F	F	T	T	T	T	T	T	T	T
T	F	F	F	F	F	T	T	T	T	F	F	F	F	T	T	T	T
F	T	F	F	T	T	F	F	T	T	F	F	T	T	F	F	T	T
F	F	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T

TABLE 1: The truth table for 2-input Boolean functions. F_9 corresponds to the AND function, while F_{15} is the OR function.

- $A \vee B = \neg(\neg A \wedge \neg B)$

To some extent, Boolean algebra and set theory represents the same concepts, with the AND operation $\wedge$, the intersection operator $\cap$, and the concept of conjunction equivalent in some cases. The same holds for the OR operation $\vee$, the union operator $\cup$, and the concept of disjunction. It is easy to notice that given two possible truth values for the logical statements A and B, there are four possible A - B combinations. We can construct a *truth table* containing all 16 possible Boolean operations with two inputs, as in Table 1.

Logic primarily deals with deterministic events, in the sense that if there exists a rule that A implies B, we implicitly expect this relationship to hold every time and everywhere. However, natural events do not display this type of absolute certainty, but rather some degree of stochasticity. Fundamentally, for empirical results to be certainties would imply exact knowledge of all parameters and initial conditions, however these can only be known in practice to the degree to which we can measure them. Such events are termed random, as their outcome is not certain for every instance. Random events are known since antiquity, as is attested by the existence of games of chance (dice, cards), which probably provided the incentive for the first formal studies of probability. Probability theory rests on a number of axioms, famously articulated by Andrey Kolmogorov (Kolmogorov, 1956):

1. The probability of any event is a non-negative real number
2. The probability that at least one of the elementary outcomes of a random process will occur is 1
3. The probability of sets of mutually exclusive outcomes is the sum of the probabilities of the individual events

Several other rules directly derive from these three axioms:

1. The probability of no outcome occurring is zero
2. If A is a subset of B, then the probability of A is no more than the probability of B
3. Every outcome occurs with a probability comprised between 0 and 1

1.2.2.2 Bridging logic and probabilities together

The truthfulness of some logical proposition is usually evaluated as being an all-or-nothing phenomenon, in the sense that the proposition is either completely true or completely false. However, non-classical logical systems exist, in which the truth value of propositions can take other values, for example be in-between true and false, or be undetermined. The formalization of this type of logic was started by Jan Łukasiewicz at the beginning of the twentieth century and evolved into a series of related logical formalisms. The goal of using multi-valued logical formalisms is to account for ill-posed problems and to solve paradoxes. For example, the question "Is the 10^{100th} decimal of π even?" cannot be answered with certainty at this moment, since it has not been computed yet. However, we know for sure that this digit exists, and that it is either odd or even. A truth value between 0 (completely false) and 1 (completely true) might therefore represent both our ignorance of the actual answer to the question and our confidence about the outcome of future computations. Similarly, intermediate truth values can be assigned to future events, as we do not know the future with certainty.

Many-valued logics have been proposed in the 1960s (Zadeh, 1965) and have been given the collective appellation of "fuzzy logic". Fuzziness expresses the inherent vagueness of real-world objects, as opposed to mathematical constructs. There is a relation between fuzzy theory and probability theory, in that they map degrees of uncertainty to the $[0, 1]$ interval, but they represent different types of uncertainty: whether an event occurs is random; to what degree it occurs is fuzzy (Kosko, 1990; Gottwald, 2001). Fuzzy logic has been applied to a number of problems, notably in the area of biology, for example to ontologies (Calegari & Ciucci, 2006), the analysis of microarray data (Huerta *et al.*, 2008), and brain connectivity maps (Morgan *et al.*, 2016).

1.2.3 Logical modeling of biological processes

The mechanisms by which cells perform the massive number of integrations necessary for their normal functions are not exactly known. Similarly, how these systems fail in different diseases is not known, however this information would enable to design new therapeutic strategies for conditions that are difficultly to treat currently, like cancer. These biological systems include principles that are known components of signal-processing systems, like feedback, damping, parallel execution, or redundancy. One possible approach is to model these regulatory systems and determine the parametrization of the model with measurements of the biological elements that are being modeled. This naturally demands that we choose the class of model which is appropriate for this task. Classically, the choice of model class is intimately linked to the type and amount of available data and the nature of the research question. Two extremes can be defined as follows: a detailed model, with more parameters, will be able to provide more insight into the mechanism under study, but will require large amounts of data, and failure to provide it will result in 'overfitting' of the data. On the contrary, a simpler model with less

parameters will be more robust to noise in the data but might not capture appropriately the behavior of the system. The trade-off that needs to be made is one between bias and variance, and the best model is the one which is complex enough to explain the data, but not more (Le Novère, 2015).

There are multiple dimensions by which models of biological systems can be either complex or simple, depending the assumptions they make. The main ones are detailed in Table 2.

Dimension	Complex	Simple
Time	Dynamic	Steady-state
Variables	Continuous	Discrete
Values	Unbound	Bound
Interactions	Directed	Undirected
Interactions	Signed	Unsigned
Behavior	Stochastic	Deterministic

TABLE 2: The various dimensions by which models can form simplifying assumptions

An example of an extremely simple modeling framework is the computation of correlation between variables in a series of experiments assumed to be comparable equilibria. Such analysis is able to infer steady-state, undirected, sometimes signed, sometimes continuous interactions between the variables, in a deterministic way. The level of confidence in such interactions can be computed from the distribution of the data, however the quantity of information that can be retrieved is low. An opposite example is a completely determined model, in which the precise chemical mechanisms are modeled by partial differential equations and in which a noise factor is added to account for the stochasticity of the processes. This type of model would be invaluable to understand precisely the mechanisms underlying functional responses across time and cellular compartments and make quantitative predictions, however acquiring the necessary data to constrain it might not be feasible, and the computational power necessary to simulate such a system might limit the scope of the possible analyses.

In that context, logical models have been presented as an interesting middle-ground which captures high-level systems behavior while retaining maximal flexibility and practical feasibility (Morris *et al.*, 2010). There are different types of logical model, which are introduced in the following section.

1.2.3.1 Boolean network models

Boolean networks were proposed by Kauffman half a century ago as models of biological regulatory networks (Kauffman, 1969). A Boolean network $G(V, F)$ is defined by a set of nodes $V = \{x_1, ..., x_n\}$ and a list of Boolean functions $F = (f_1, ..., f_n)$. Each node is a Boolean

variable which can be either TRUE or FALSE. A Boolean function $f_i(x_{i_1}, ..., x_{i_k})$ with k specified input nodes is assigned to each node x_i. As there are n variables, the network as a whole can be in one of 2^n possible states. The original presentation of Boolean networks considers only synchronous updating, in which all Boolean functions are computed from the values of the variables in one state at time T, and the network reaches another state at time T+1 (other update schemes are possible, notably asynchronous and semi-asynchronous). Boolean networks are example of Markov processes (a stochastic process which evolution does not depend on its history but only of its current state). By repeating the computation of node values from the values of their parent nodes, the network takes successive states until it eventually reaches a state it previously visited. The dynamics of this Boolean network, due to its inherent deterministic behavior and the finite number of possible states, indicates that the network will eventually either stay in one stable state or cycle infinitely between a number of states. Such states are called *attractors* and the set of states that lead to them form the *bassin of attraction* of such attractors. We can draw a state diagram indicating which state is reached from which, evidencing the attractors and their bassins of attraction. Consider the example Boolean network defined by the following simple interactions, taken from Wang *et al.* (2012):

- A is activated by itself AND inhibited by C
- B is activated by A AND C
- C is activated by B

These interactions can be formalized in the following rules:

- $f_1 = x_1$ OR NOT x_3
- $f_2 = x_1$ AND x_3
- $f_3 = x_2$

These rules define the following network topology:

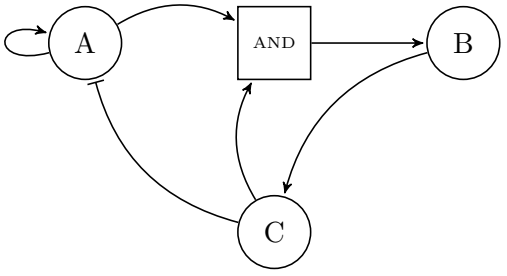


FIGURE 4: Example of a Boolean network. Each variable is completely defined by a Boolean function.

The dynamic behavior of this system can be visualized as the state diagram in Figure 5, where each of the $2^3 = 8$ states of the system is represented as a string of the truth values for the nodes A, B, C, in that order:

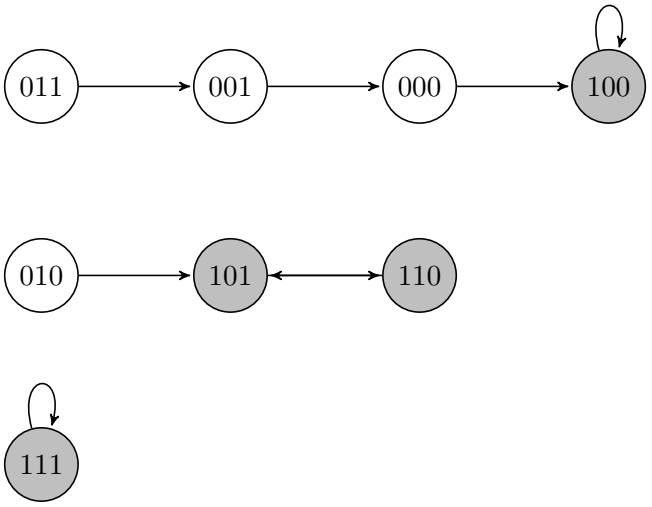


FIGURE 5: The state transition diagram of the example Boolean network. The state space forms three disjoint bassins of attraction. The shaded states are attractors.

Alternatively, the dynamics of this system can also be represented as a state transition matrix (Table 3):

ABC	000	001	010	011	100	101	110	111
000	0	1	0	0	0	0	0	0
001	0	0	0	1	0	0	0	0
010	0	0	0	0	0	0	0	0
011	0	0	0	0	0	0	0	0
100	1	0	0	0	1	0	0	0
101	0	0	1	0	0	0	1	0
110	0	0	0	0	0	1	0	0
111	0	0	0	0	0	0	0	1

TABLE 3: State transition matrix for the example Boolean network

In this case, there are three attractors ((100), (111), and (101/110)), and correspondingly three bassins of attraction. When initialized in any of the 8 possible states, this system will spontaneously evolve and reach one of the attractor states, depending in which bassin of attraction the initial state lied. The fixed points of the state space correspond to solutions to the constrain of time invariability: by posing $T + 1 = T$, we have in the above example:

- $A = A$ OR NOT C

- $B = A$ AND C
- $C = B$

By substituting the last equation into the second one, we have that $C = A$ AND C. This, in turn, can be substituted in the first equation to lead to $A = 1$. As a result, given that $C = 1$ AND $C = C$, we deduce that C (and therefore B as per the third equation) can be either 0 or 1. The two fixed points of the system are (100) and (111) (Figure 5).

It should be noted that these two states are singleton attractors, and that the third attractor is a cyclic attractor, *i.e.* a set of several states through which the network cycles indefinitely. States that only occur at the beginning of trajectories (no trajectory leads to them, like (010) and (011) in the above example) are called garden-of-Eden states.

Boolean networks have been used to model biological regulatory networks. It has been hypothesized that the binary behavior of genes, which can be expressed at a given time in a given cell or not, is appropriately represented by the ON/OFF nature of Boolean variables, and that the attractors of Boolean networks are naturally stable states representing qualitatively the expression of the different genes in the cell. In other terms, the hypothesis is that the topology of the cellular gene regulatory network defines a series of latent genetic programs (proliferation, quiescence, apoptosis), hardwired into the genome, which the cell can execute and transition from one to another depending on external signals (Huang & Ingber, 2000). Boolean networks have been successfully used to model different biological systems, notably T cell signaling (Saez-Rodriguez *et al.*, 2007), apoptosis (Schlatter *et al.*, 2009), the genetic regulation of tissue development in mammalian (Giacomantonio & Goodhill, 2010) or plant cells (Mendoza *et al.*, 1999), cell-cycle transitions in yeast (Li *et al.*, 2004), to cite a few. Boolean models are useful because the attractors predict the activity of the different components of the system in the resting and perturbed cases and therefore generate testable hypotheses (Albert & Thakar, 2014). For example, analysis of a Boolean model of the signaling network of T cell large granular lymphocyte leukemia (T-LGL) led to the identification of 19 potential therapeutic targets for the disease, more than half of which were later validated experimentally (Saadatpour *et al.*, 2011). Similarly, combined network analysis and simulation of Boolean network models of signaling in liver cells recovered the main differences between primary and transformed hepatocytes (Saez-Rodriguez *et al.*, 2011). In practice, the R package BoolNet (Müssel *et al.*, 2010) and the MATLAB toolbox CellNetAnalyzer (Klamt *et al.*, 2007) are the basic tools to study the properties of Boolean networks and perform attractor analysis, while CellNOptR (Terfve *et al.*, 2012) can be used to train networks using Boolean or other logic-based formalisms.

1.2.3.2 Extended logical models

While Boolean networks constitute the backbone of logical network theory, a number of mathematical frameworks have been devised in the following decades that generalize the strict Boolean formalism and provide additional flexibility to the modeling procedure. For example, one of the main limitations of Boolean networks is the requirement of discretization of the data from a continuous scale to ON/OFF values, which involves arbitrary thresholds and inevitable loss of information. The other main disadvantage of Boolean networks is their determinism, which is ill-suited to the study of stochastic biological phenomenon, and does not adequately account for uncertainty both in the data and the design of the logical rules.

By far the most studied of such extensions of the Boolean formalism is Probabilistic Boolean Networks (PBNs), introduced by Shmulevich *et al.* (2002). The basic idea is to extend the Boolean network to include more than one Boolean function for each node, so that we have $F_i = \{f_j^{(i)}\}$ for $j = 1, \ldots, l_i$, where each $f_j^{(i)}$ is a possible function determining the value of node x_i and l_i is the number of possible functions for node x_i. In other words, there can be more than one updating function for every node. The update function of a PBN at a given instant of time is determined by a vector of Boolean functions. If there are N possible realizations, then there are N vector functions, $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_N$ of the form $\mathbf{f}_k = (\mathbf{f}_{k_1}^{(1)}, \mathbf{f}_{k_2}^{(2)}, \ldots, \mathbf{f}_{k_n}^{(n)})$, for $k = 1, 2, \ldots, N$. k_i must satisfy $1 \leq k_i \leq l^{(i)}$, and the function $\mathbf{f}_{k_i}^{(i)} \in F_i$ with $i = 1, \ldots, n$. We can design a random vector $\mathbf{f} = (f^{(1)}, \ldots, f^{(n)})$ taking values in $F_1 \times \cdots \times F_n$. We can define the probability that predictor $f_j^{(i)}$ is used to predict node i as $c_j^{(i)} = Pr\{f^{(i)} = f_j^{(i)}\}$, which need to satisfy $\sum_{j=1}^{l(i)} c_j^{(i)} = 1$ (Shmulevich *et al.*, 2002; Shmulevich & Dougherty, 2002).

Because the update of PBNs is stochastic (determined by random variables), the dynamical behavior of PBNs can be thought of as a (non-deterministic) Markov process where the state transition matrix is completely specified by the probabilities of the constituent Boolean functions. Consider the following PBN, similar in topology to the above example, but modified in the following way: there are two possible functions to update node A: $f_1^{(A)}$ and $f_2^{(A)}$ (with respective probabilities c_1^A and c_2^A), and two possible functions to update node C: $f_1^{(C)}$ and $f_2^{(C)}$(with respective probabilities c_1^C and c_2^C). There is only one function for the update of B: f_1^B, with probability $c_1^B = 1$. These functions might have the following truth tables:

ABC	$f_1^{(A)}$	$f_2^{(A)}$	$f_1^{(B)}$	$f_1^{(C)}$	$f_2^{(C)}$
000	0	0	0	0	0
001	1	1	1	0	0
010	1	1	1	0	0
011	1	0	0	1	0
100	0	0	1	0	0
101	1	1	1	1	0
110	1	1	0	1	0
111	1	1	1	1	1

Assuming complete independence of the selection probabilities, we can construct the following matrix K summarizing the four possible ways to select a combination of update functions:

$$K = \begin{bmatrix} f_1^{(A)} & f_1^{(B)} & f_1^{(C)} \\ f_2^{(A)} & f_1^{(B)} & f_1^{(C)} \\ f_1^{(A)} & f_1^{(B)} & f_2^{(C)} \\ f_2^{(A)} & f_1^{(B)} & f_2^{(C)} \end{bmatrix}$$

The number of rows in this matrix is the number of possible network realizations N, and the number of columns is the number of nodes n. It is easy to verify that the probability that network i is selected is $P_i = \prod_{j=1}^{n} c_{K_{ij}}^{(j)}$. We can now construct the full transition matrix A for the system as follows:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ P_4 & P_3 & 0 & 0 & P_2 & P_1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & P_2+P_4 & P_1+P_3 \\ 0 & 0 & 0 & 0 & P_2+P_4 & P_1+P_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Following the theory of Markov processes (Markov, 1954) we can infer that this system will settle in a steady-state distribution of the different states of the PBN state-space when released from an arbitrary prior distribution. In the above example, if the network is released from an uniform prior distribution of $D_0 = \{\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}\}$, we can use the matrix A to compute the limiting distribution $D_\infty = \{0.15, 0, 0, 0, 0, 0, 0, 0.85\}$. We can observe that the system converges to only two states (000 and 111), which are called absorbing states. These states loosely correspond to the concept of attractors in BNs.

We can differentiate *instantaneous* PBNs, in which a new network instantiation is realized at every time-step, and *context-sensitive* PBNs, where the probability for switching to another network configuration is controlled by another, external, random variable $\xi \in \{0, 1\}$. This latter type is adapted to situations where the model is affected by latent variables outside the model (Choudhary *et al.*, 2006) and corresponds to the case $P(\xi = 1) < 1$. However, in this type of PBN the final steady-state distribution is still dependent of the initial condition. A further expansion of the model allows to overcome this limitation: by including a tiny probability for every node to randomly switch from the inactive state to the active state and vice-versa, we allow the transition of the system to the next state in ways that are not possible by any network topology. This makes the dynamics of PBNs with perturbations ergodic, in the sense that by allowing every state to be reachable from every other state, we merge all bassins of attraction and given enough time, the PBN will reach the same limiting distribution regardless of the initial state.

The central assumption of PBN modeling is that the limiting distribution of the states in a PBN model of regulatory network corresponds to stable states of the actual biological network. PBNs have been used to model several biological regulation systems (Trairatphisan *et al.*, 2013). In one example, a PBN framework was used for the inference of genetic regulatory networks from gene expression time-course data in macrophages submitted to interferon treatment and viral infection, revealing different selection probabilities for predictor functions indicating network re-wiring upon stimulation (Yu *et al.*, 2006). More recently, the ergodic sets of states that correspond to each phase of the cell cycle in a PBN model of the budding yeast were identified, indicating that the irreversibility of the process is a consequence of the topology of the network (Todd & Helikar, 2012).

Bayesian networks are another type of graphical probabilistic model used to represent the joint probability distributions of variables in a network. A Bayesian Network consists of a directed acyclic graph G and a set of random variables $X = X_1, X_2, ..., X_n$ and their conditional dependencies, which are generally understood to be causal. The graph G encodes the conditional independencies that relate the variables together. In a Bayesian Network, it is assumed that the information in graph G is complete, meaning that each variable X_i is independent of its non-descendants, given its parents, and there exists several possible representations for the parameter set θ denoting for each node X_i the conditional distributions $P(X_i|Pa(X_i))$, where $Pa(X_i)$ is the set of the parents nodes of X_i in the graph G (Pearl, 2014). Bayesian Networks can be used for inferring the state of unobserved variables and for estimating the conditional dependencies between variables when these are not known. Figure 6 shows an example Bayesian Network.

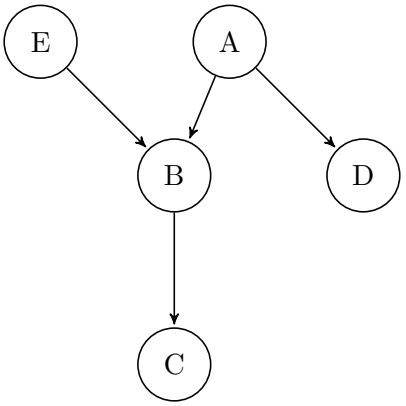


FIGURE 6: Example of a Bayesian Network. The topology of this network structure encodes a number of conditional independence statements: $I(A;E)$, $I(B;D|A,E)$, $I(C;A,D,E|B)$, $I(D;B,C,E|A)$, $I(E;A,D)$

Bayesian Networks have been applied to the study of gene regulatory networks in yeast (Friedman *et al.*, 2000) and were able to approximate the correct topology of a gene regulatory network in a simulated population in the DREAM5 challenge (Vignes *et al.*, 2011). The problem of recovering the structure of a partially observed Bayesian Network is an old challenge of machine learning (Rebane & Pearl, 2013), and is known to be NP-hard (Chickering *et al.*, 2004). Different learning methods have been devised to learn the graph G from expression data (Xing *et al.*, 2017; Fan *et al.*, 2017; Wilczynski & Dojer, 2009) with however diverging methods and results.

A natural extension of Bayesian Networks is Dynamic Bayesian Networks (DBNs), which is characterized by the presence of loops in the graph G. Figure 7 shows an example DBN. Compared with the Bayesian Network of figure 6, the presence of the additional edge cancels some of the constrains of independence (for example, the node E is no longer independent of the node A).

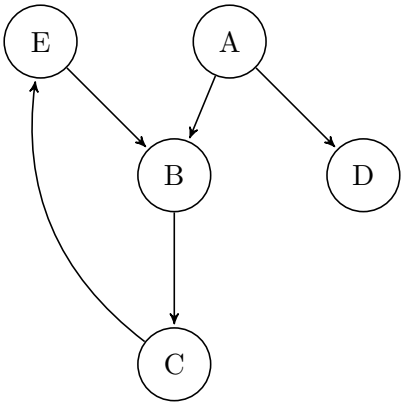


FIGURE 7: Example of a Dynamic Bayesian Network. The loop E-B-C-E defines a dynamic system.

Dynamic Bayesian network models and Probabilistic Boolean models are similar in nature and in practice, their performance for the inference of simple interaction networks is comparable (Li *et al.*, 2007). Notably, the GlobalMIT algorithm (Vinh *et al.*, 2011) was shown to compute the globally optimal structure of a DBN in polynomial time, with however poor results when the sampling rate was low. It has been demonstrated that PBNs and DBNs can represent the same joint probability distribution over their common variables (Lähdesmäki *et al.*, 2006). Usually, prior knowledge and gene expression data are combined to generate a contextualized network, informative of the active regulation paths. DBNs have been used to investigate the behavior of gene regulatory networks in the context of circadian rhythm in plants (Grzegorczyk & Husmeier, 2011), the interplay between EGFR and Hedgehog signaling in human medulloblastoma (Fröhlich *et al.*, 2015), the timeline of cell cycle regulation events (Zou & Conzen, 2005), or the prediction of recurrence of Oral Squamous Cell Carcinoma (Kourou *et al.*, 2017).

1.3 Goal of this thesis

The goal of this work is to investigate both theoretically and experimentally the possibility to efficiently model the signaling pathways that are frequently perturbed in cancer and their regulation with logical network models. Building on previous work using Probabilistic Boolean Networks, a Matlab toolbox is created to efficiently contextualize models of biological regulation and investigate them quantitatively.

In a first step, the formulation of logical rules for the interaction of molecules is conceptualized in a Bayesian context, and a platform to integrate network topologies with experimental data is created. Next, a bioinformatics toolbox is produced, which is able to optimize a set of continuous parameters for the model in a way that maximizes the fitting of the model to the experimental data. With the goal of identifying network nodes playing the largest role at the functional level, a series of systems-level complementary analyses is conceived in order to individualize the elements of the network (nodes or edges) that play a dominant role in determining the functionality of the system.

Next, the task of identifying the most crucial elements of a DBN model of signaling networks in cancer is facilitated by the design of regularization schemes, which help forming cell line-specific models of signaling pathways starting from non-specific prior-knowledge networks. The performance of the toolbox is evaluated in term of computational efficiency, and an analysis pipeline is built to investigate the resistance mechanism of melanoma cells to targeted agents in order to show the usefulness of the approach.

1.4 Outline

This thesis compiles three papers published over the course of the doctoral training, highlighting the main steps of the progression of the project. In the first paper, a bioinformatics toolbox is presented which proposes to model logical networks of signaling pathways with Dynamic Bayesian Networks. The application of the toolbox is discussed in the general context of probabilistic logical models and is demonstrated with simple examples. The second paper addresses the problem of accurate parametrization of parallel models of signaling pathways in the presence of uncertainty. The proposed method uses a measure of the density of the models parameters in the parameter space and makes the assumption that interaction strengths adopt a finite number of discrete values corresponding to the discrete genomic events leading to the specific phenotype under study. This assumption can be integrated in the optimization procedure (in the form of regularized optimization) and help characterize cancer cell lines, as demonstrated with our analysis of the phosphoproteomic data from a panel of 11 colorectal cancer cell lines. The third paper is the result of a collaboration between our department and the Technical University of Dresden. In that paper, the mechanism of resistance of melanoma cells to a pro-apoptotic drug is investigated by analyzing phosphoproteomic data from sensitive and resistant cell lines. By carefully selecting the model that balances best goodness-of-fit and model size, we are able to infer the most sensitive points of the regulatory system and correctly predict ways to re-sensitize the adapted cells. We conclude with a general discussion of our methods, our results, their limitations and possible applications.

Chapter 2

Materials and Methods

2.1 Architecture of the FALCON toolbox

The FALCON toolbox was conceived as a modeling environment for the contextualization of molecular interaction networks with quantitative biological measurements. The name FALCON stands for Fast ALgorithm for the Contextualization Of Networks. The toolbox is programmed in the proprietary language Matlab (The MathWorks, CA). There are two ways the toolbox can be used: with scripts (or console inputs), or using the Graphical User Interface. When FALCON is used on a high number of High-Performance-Computing cluster nodes, it might be useful to write custom scripts to control the I/O of the different tasks and their integration, however this step usually is platform-specific. When used on a single computer, FALCON is designed to make maximum use of the strengths of Matlab's matrix-orientated programming structure and to efficiently dispatch optimization jobs across the multiple cores of modern computers' CPUs.

The general architecture of the toolbox is presented in figure 8. During the first phase, the various input files are used to define the optimization problem in terms of positive and negative interaction matrices of the network, lists of nodes and their associated Boolean rules, and the different constraints on the parameter set stemming from the simple rules of probabilities. The correspondence between the information in the network file and the data file is ensured by a series of checks, therefore guaranteeing the mathematical coherence of the network so that for example, no node exists without either a parent node or a corresponding control input in the data file. During this phase, substantial re-arrangement of the network can also take place, with the goal of speeding up computations:

- network expansion: Boolean functions with more than two inputs are decomposed into their essential two-input functions as follows: $A \wedge B \wedge C = (A \wedge B) \wedge C$ and $A \vee B \vee C = (A \vee B) \vee C$. This creates intermediate nodes, invisible to the user but used in the computations.

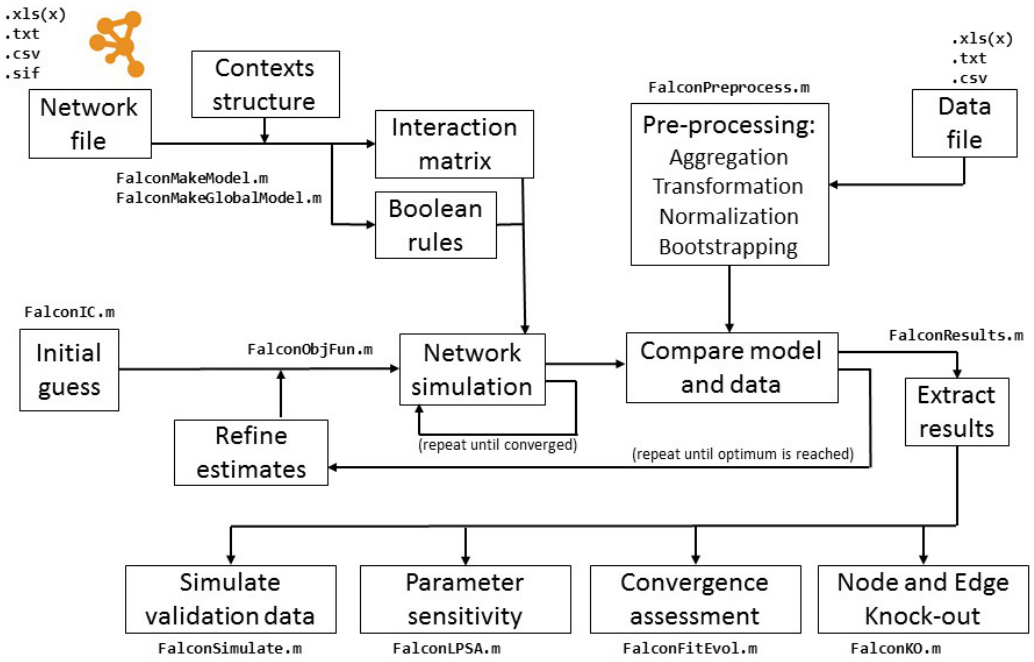


FIGURE 8: The general architecture of FALCON

- network reduction: nodes which are neither controled nor measured, and which are only connected to one input node and one output node (they have only one parent and one child) are removed from the network, and their parent and child are connected by a new edge. Indeed, if we have $A \xrightarrow{k_1} B \xrightarrow{k_2} C$, it is easy to demonstrate that this is equivalent to $A \xrightarrow{k_1 k_2} C$.

2.2 Input formats

A network can be defined by the list of its interactions. This type of format is the basis of the well-known *.sif* format, used by Cytoscape (Shannon, 2003), Pathway Commons (Cerami *et al.*, 2010), the *networkx* Python package (Hagberg *et al.*, 2008), as well as many other bioinformatics tools. It consists of a list of lines, with each line having the following structure: <*source*> *tab* <*interaction type*> *tab* <*target*>. This type of format is accepted by FALCON, with two types of edge being recognized: -> and -|, corresponding to activating and inhibiting edges, respectively. When using the *.sif* format with FALCON, the following characteristics of the edges are set to their default value:

- the name of the parameter will be automatically formed from the name of the input and output nodes.
- the type of logical interaction will be automatically set to additive.
- no other boundary will be forced on the parameter values than the ones inherent to the topology.

Alternatively, the *.csv* and *.xls(x)* formats can be used, which are compatible with Excel, the most-used tabling application. In that case, three additional pieces of information can be provided for each edge: a custom parameter name, either one of *N*, *A*, or *O* to specify the type of logical interaction (additive, Boolean AND and OR, respectively) and either one of *H* or *L* to constrain on the edge's parameter to the high or low range, respectively. *D* (default) is used when no additional constrain is present. Figure 9 shows an example of the input format for networks in FALCON.

	A	B	C	D	E	F
1	Input	Reaction	Output	parameter	gate	constrain
2	PDGFL	->	PDGFR	1	N	D
3	PDGFR	->	STAT5	1	N	D
4	PDGFR	->	cCbl	1	N	D
5	cCbl	-|	PDGFR	k1	N	D
6	bPPX	->	PPX	1	N	D
7	PDGFR	-|	PPX	1	N	D
8	PPX	-|	PDGFR	k2	N	D
9	PDGFR	->	SHP2	1	N	D
10	SHP2	-|	PDGFR	k3	N	D
11	SHP2	->	Grb2SOS	1	N	D
12	SHP2	->	GabSOS	1	N	D
13	Grb2SOS	->	GGSOS	kOR	O	D
14	GabSOS	->	GGSOS	kOR	O	D
15	GGSOS	->	Ras	kf1	N	H

FIGURE 9: The network format

Biological measurements are used by the FALCON toolbox to parametrize the DBN structure, in such a way that the node values of the simulated network reproduces as closely as possible the actual measurements. These measurements are usually present for several chemical species (proteins or nucleic acids) and under several experimental conditions (in the presence of different drugs, over time, etc.). Three tables must be provided to the FALCON toolbox, in either one of the *.csv* or *.xls(x)* formats. One table detailing the fixed input nodes values (such as the presence or absence of a drug), one table detailing the values of the nodes for which a measurement is available, and one table containing the corresponding error for these measurement. In each of the tables, the different measured species are organized in columns while the different experiments are organized on successive lines. In the case of *.csv* files, three separate files have to be provided, while for the *.xls(x)* format, the tables must be provided as three separate sheets in the same file. Figures 10, 11, and 12 show examples of the input format for biological measurements in FALCON.

Measurements need to be pre-processed in order to be exploited by the FALCON toolbox. Data pre-processing consists of several steps: transformation, normalization, and aggregation.

	A	B	C	D	E	F	G	H	I	J
1	Annotation	PDGFL	Y720F	YY731742FF	Wort	U0126	bPDK	bPTEN	bPPX	bPKC
2	WT[-]	1	0	0	0	0	1	1	1	1
3	WT[W]	1	0	0	1	0	1	1	1	1
4	WT[U]	1	0	0	0	1	1	1	1	1
5	dM[-]	1	1	0	0	0	1	1	1	1
6	dP[-]	1	0	1	0	0	1	1	1	1
7	ND	0	0	0	0	0	1	1	1	1
8										

FIGURE 10: The *inputs* part of the data format

	A	B	C	D	E	F
1	PDGFR	PLCg	STAT5	ERK12	Akt	PKC
2	0.882	0.759	1	1	0.755	0.995
3	1	1	0.822	0.609	0.117	0.876
4	0.918	0.81	0.792	0.11	1	NaN
5	0.878	0.57	0.923	0.509	0.818	NaN
6	0.907	0.793	0.912	0.88	0.224	NaN
7	0	0	0	0	0	1
8						

FIGURE 11: The *outputs* part of the data format

	A	B	C	D	E	F
1	PDGFR	PLCg	STAT5	ERK12	Akt	PKC
2	0.083	0.093	0.086	0.03	0.1	0.021
3	0.066	0.008	0.027	0.103	0.074	0.105
4	0.06	0.047	0.078	0.028	0	NaN
5	0.163	0.204	0.183	0.063	0.172	NaN
6	0.118	0.107	0.224	0.152	0.09	NaN
7	0	0	0	0	0	0.019
8						

FIGURE 12: The *errors* part of the data format

- During data transformation, a derived value is computed for every datapoint, with the goal of representing the underlying quantities in a way that is more consistent with the modeling assumptions. For example, one common transformation is the *log-transformation* of data representing concentrations in order to transform the absolute concentrations into relative ones.

- Data normalization is a very important step in which datapoints are brought to the same scale in order to be comparable when they have been measured on, or originally exist on, different scales. One other goal is to bring the probability distributions of the quantities into alignment with each other or with the normal distribution. There are many normalization procedures, most notably the Z-score or the Student's t-statistic. As the mathematical framework under which FALCON operates is probabilistic, we used *unit normalization*, also called feature scaling, in order to keep the normalized values between 0 and 1. Formally, the unit normalization of a set X of n values $X_1, X_2, \ldots X_n$

produces a set X' of values where $X'_i = \frac{X - X_{min}}{X_{max} - X_{min}}$ for all $i \leq n$, where X_{min} and X_{max} are the smaller and larger values of the set X, respectively.

- During data aggregation, multiple measurements of the same quantity are grouped in order to extract one number summarizing the best estimate of the true value, as well as one number expressing the error on this estimate. Depending the experimental design and the distribution of the replicates values, this best estimate can be either the mean or the median, and the error can be expressed either as the standard deviation, standard error of the mean, interquartile range, etc.

2.3 Algorithm

2.3.1 Network update strategy

As we use DBNs to simulate systems at the steady-state, it is needed that the node values used for the fitting to the data points represent the true value of the converged networks. Indeed, the presence of positive and negative feedback loops in DBNs induces an oscillatory dynamic behavior which may or may not represent the true dynamics of the system. The reason why the dynamics of the DBNs and the systems they represent may not be comparable, is because we use *synchronous updating*, which is not a realistic assumption for the true system, in which the different regulatory events (ligand binding, phosphorylation cascades, etc.) may happen over different timeframes. Nevertheless, if a DBN representation of a regulatory system includes all reactions and interactions happening over a short timeframe and does not include later events, then these DBNs, at their steady-state, correspond to configurations of the true system during a *quasi-steady-state*, i.e. a state of near-equilibrium in which the system moves very slowly and during which the concentrations of the constituents of the faster-moving regulatory subsystem (the ones represented in the DBN) can be considered constant (Li *et al.*, 2008).

Simulating the DBN until steady-state consists in computing repeatedly the network update function for all nodes. This operation is done in Matlab as follows:

1. initialize all nodes to a random value
2. replace the value of the *input* nodes by their true value as specified in the data file
3. compute the new value for each node X_i as a function of their activating $(j+)$and inhibiting $(j-)$parent nodes ($Pa\big(X^{(i)}\big)_{t-1}^{(j+)}$ and $Pa(X^{(i)})_{t-1}^{(j-)}$, respectively) and the relative weights $k^{(i)}$ of the parents' interaction as :

$$X_t^{(i)} = \sum_{j^+=1}^{m} k_{j+}^{(i)} Pa\big(X^{(i)}\big)_{t-1}^{(j+)} \times \Big(1 - \sum_{j^-=1}^{i} k_{j-}^{(i)} Pa(X^{(i)})_{t-1}^{(j-)}\Big)$$

4. for each node X_i defined by a Boolean function, replace the value by the one of the AND or OR continous Boolean function:

$$\text{AND: } X_t^{(i)} = \prod Pa(X_{t-1}^{(i)})$$
$$\text{OR: } X_t^{(i)} = 1 - \prod (1 - Pa(X_{t-1}^{(i)}))$$

5. compute the distance D_t between the network at time t and $t-1$ as:

$$D_t = \sum_i |X_t^{(i)} - X_{t-1}^{(i)}|$$

6. repeat until $D_t \simeq 0$

In practice, we choose a threshold value φ and we stop updating the network when $D_t \leq \varphi$. This threshold should be small enough to only be reached when the network reaches convergence, but be large enough to be resistant to numerical errors. In practice, we have used values between 10^{-14} (for small networks with a few nodes) and 10^{-4} (for networks with hundreds of nodes) with success.

2.3.2 Parameter learning

In order to learn the parameter set for which the DBN model best fits the normalized measurements, two things are necessary: a measure of the distance between the predictions of the model and the actual data, and an algorithm to minimize this measure, which is a function of the chosen parameters. For the measure, we choose to use the *mean of squared errors* (MSE), which is a frequently used measure of the fitting of a model to a dataset. The assumption for using the MSE is that each datapoint is the sum of two terms: the true value $\hat{Y}_i$ and the error ϵ_i on our knowledge of this value, and that this error, or bias, is normally distributed. More formally, for the set of predicted values Y for a regressor $f(X, \theta)$ attempting to predict the true values $\hat{Y}$ we have:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{Y}_i - f(X_i, \theta)\right)^2$$

This metric is used as the objective function for the learning of the optimal parameter set in FALCON. It should be noted that other metrics could be used, most notably the *rooted mean squared error*, or the *mean absolute error*, or various other measures normalized by the mean or range of the datapoints, with the goal of facilitating comparison between models and datasets.

We use a variant of gradient descent particularly adapted to our formulation of the model. Indeed, the presence of feedback loops forbids the straightforward computation of the learning gradients, because it would require numeric computation of the partial derivatives of each parameter by each dimension of the parameter space. While such a computation is easy to perform in the case of a regression problem, and extendable to larger models, for example feedforward neural networks by application of the *chain-rule* of derivatives (Rumelhart *et al.*, 1986), this is not possible with DBNs, because of the presence of loops. Instead, we compute empirical gradients with the interior-point method (Vanderbei & Carpenter, 1993; Waltz *et al.*, 2006): for each parameter, we approach the partial derivative of the cost function with respect to the parameter around the current estimate by assuming the linearity of this derivative for very small perturbations, and evaluating the function with parameter values slightly superior and inferior to the current value. It has been shown that under mild assumptions (Karmarkar, 1984) this quantity is in practice equivalent to the true gradient of most functions, while being fast to compute. In our case, we use the Matlab function *fmincon*, which integrates this optimization method with the use of hard constrains on the parameter values to ensure the coherence of the inferred values with the probabilistic context of our mathematical framework.

2.3.3 Initial parameter guesses

The performance of the toolbox was increased by choosing an appropriate procedure for the initial choice of the parameter values at the start of the optimization process. As the parameter values are also probabilities, one natural choice for an uninformative prior distribution might be the uniform distribution in the $[0, 1]$ interval. However, this is often not optimal. Indeed, it has been shown (Glorot & Bengio, 2010) that initializing weights in signal-processing networks far from their optimal value can result in longer training times. For this reason, one reasonable choice for the initial parameter guesses might be 0.5, or exactly the center of the range of possible values. However, since a random initialization is needed to decrease the chance for the optimization procedure to find a local minimum instead of the global one, we choose to initialize the network weights based on a bounded normal distribution, i.e. set of N parameters $K_{t_0} = [k^1_{t_0}, k^2_{t_0}, ..., k^N_{t_0}] \sim \mathcal{N}(0.5, \sigma)$ with σ chosen to be very small and subject to $K \in [0, 1]$. In practice we use $\sigma = 0.05$. Figure 13 shows the speed of convergence for a small network using different distribution for the random initial guesses.

2.4 Graphical User Interface

In order to make the use of the toolbox easier for scientists without extended knowledge of the Matlab language, we designed a Graphical User Interface, which helps users plan and perform experiments with DBN models. The interface is built with the Matlab interface design tools and can be called from the Matlab console with the command `FalconGUI`. Within the interface

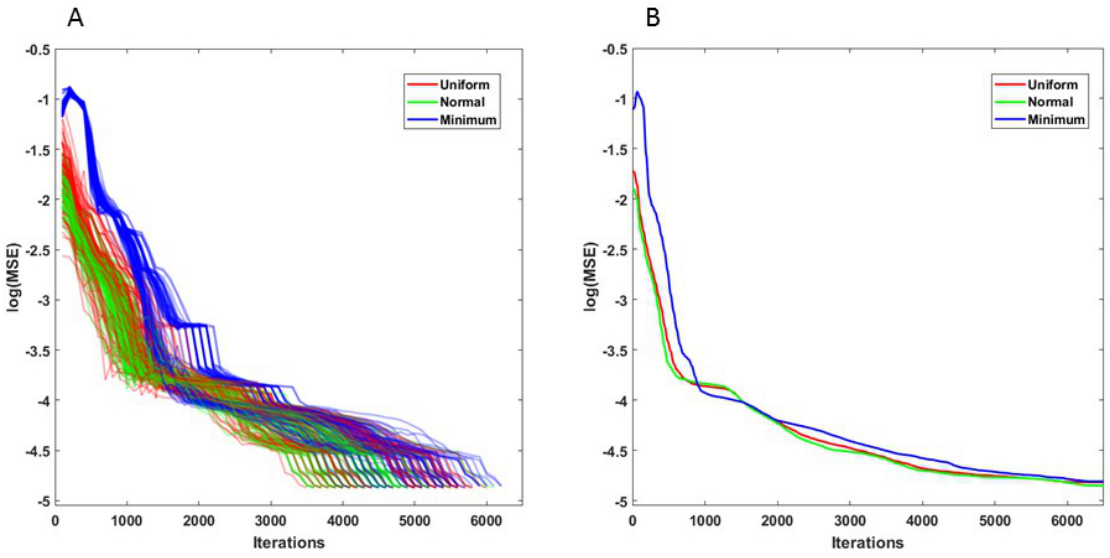


FIGURE 13: Speed of convergence of the FALCON algorithm using different stating conditions. The plots show the logarithm of the fitting cost (Mean Squared Error) as a function of the optimization instance, for the final DBN model of PDGF signaling in Trairatphisan *et al.* (2016). The weights are initialized using either in red: $K \sim \mathcal{N}(0.5, 0.05$, in green: $K \sim \mathcal{U}(0, 1)$, or in blue: $K \sim |\mathcal{N}(0, 0.01)|$. A: optimization profiles for 50 individual random runs; B: average optimization profile over the 50 runs in plot A.

environment, users can import the different input files, control the different hyperparameters of the optimization procedure, and execute a choice of custom additional analyses on the contextualized model. Figure 14 shows a screenshot of the interface.

2.5 Systems-level analyses

The main motivation for modeling cancer signaling pathways usually is the discovery of particular characteristics of the network which might be exploited pharmacologically to direct the activity of signaling pathways in one way or another. For this reason, we build an analysis pipeline to explore the functional properties of the different nodes and interactions of the network.

2.5.1 Sensitivity Analysis

Sensitivity Analysis is a central component of the standard pipeline of analysis of complex systems. Fundamentally, when establishing a model of such a system, a degree of uncertainty surrounds the parametrization of the model, and hence the model's predictions. The robustness of the model is the degree to which the model structure, parametrizations and predictions remain unchanged upon small changes in the data used to train it. The analysis of the robustness of the model can result in increased understanding of the relationships between

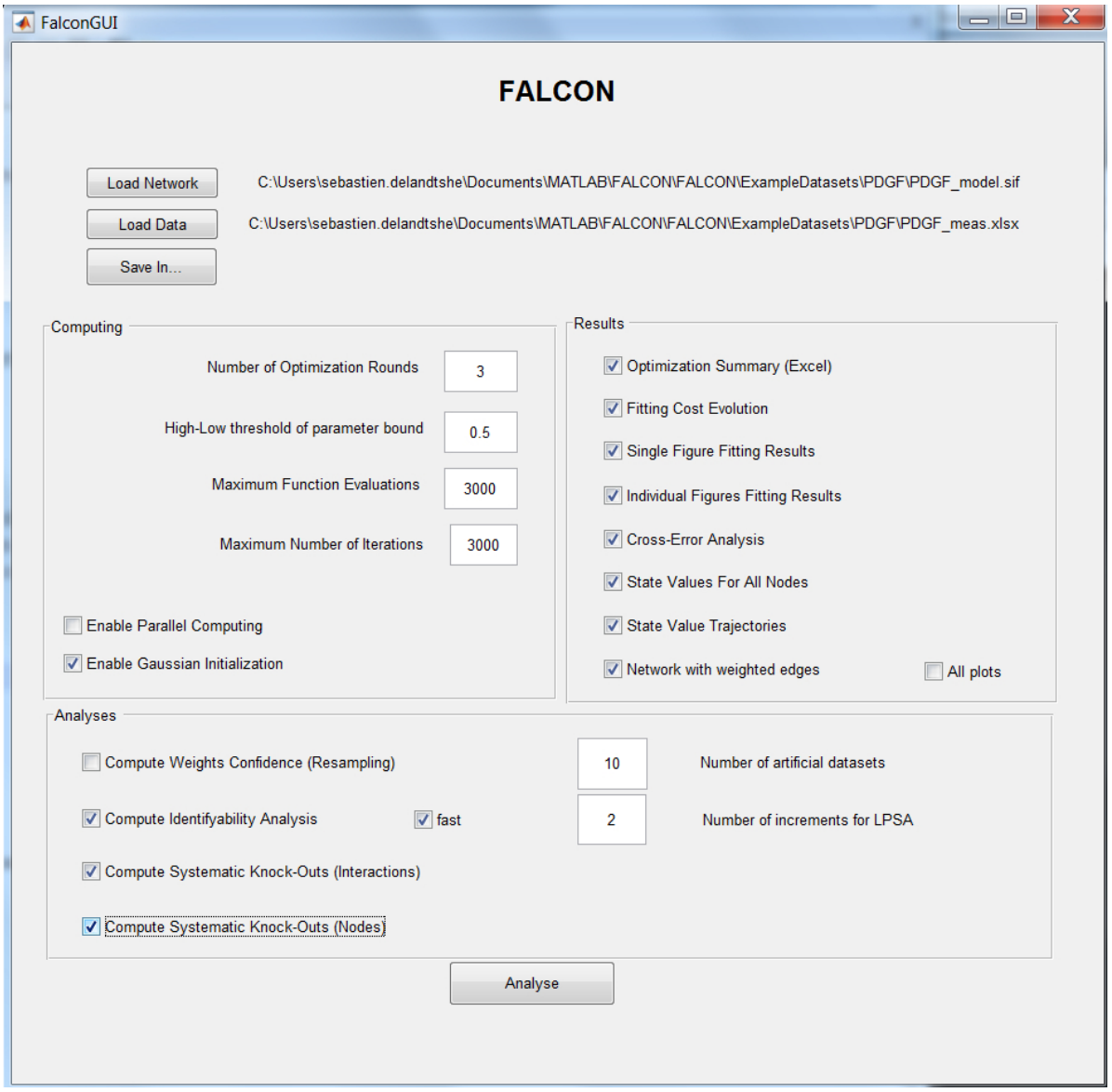


FIGURE 14: The Graphical User Interface in FALCON

the different constituents of the system and help identify the most sensitive variables. Indeed, when searching for new pharmacological targets for cancer treatment, we are interested in the nodes and interactions which modification will result in a large functional effect within the cells. The simplest form of sensitivity analysis involves the perturbation of one of the model parameters at a time while keeping all others constant. The variation in output variables can therefore be related to the amount of change in each of the model variables. Other methods exist to assess the sensitivity of the constituents of a model, notably variance-based sampling approaches (Sobol, 1990) and designed experiments (Czitrom, 1999). However, these methods are not immediately applicable in our Bayesian context where some events probabilities are necessarily correlated, and do not take into account the inherent adaptability of living systems, in which perturbations can often be attenuated by the redundancy of the signal processing pathways. In our case, we chose to implement a method of re-assessing the fitness of the model

after perturbing one of the model parameters at a time, and re-optimizing all other parameters each time. This method is similar to the determination of profile likelihood in ODE systems (Raue *et al.*, 2009). Figure 15 shows the result of the complete analysis of the PDGF signaling model from Trairatphisan *et al.* (2016). Examination of the parameters profile leads to the identification of regions of the model displaying structural non-identifiability (very flat profiles occurring when parameter perturbations can be completely compensated by another parameter or group of parameters), practical non-identifiability (asymmetric profiles occuring when the amount of observations of the model is insufficient in some dimensions), and help identifying the parameters which value is most constrained by the data, therefore helping in the design of follow-up experiments.

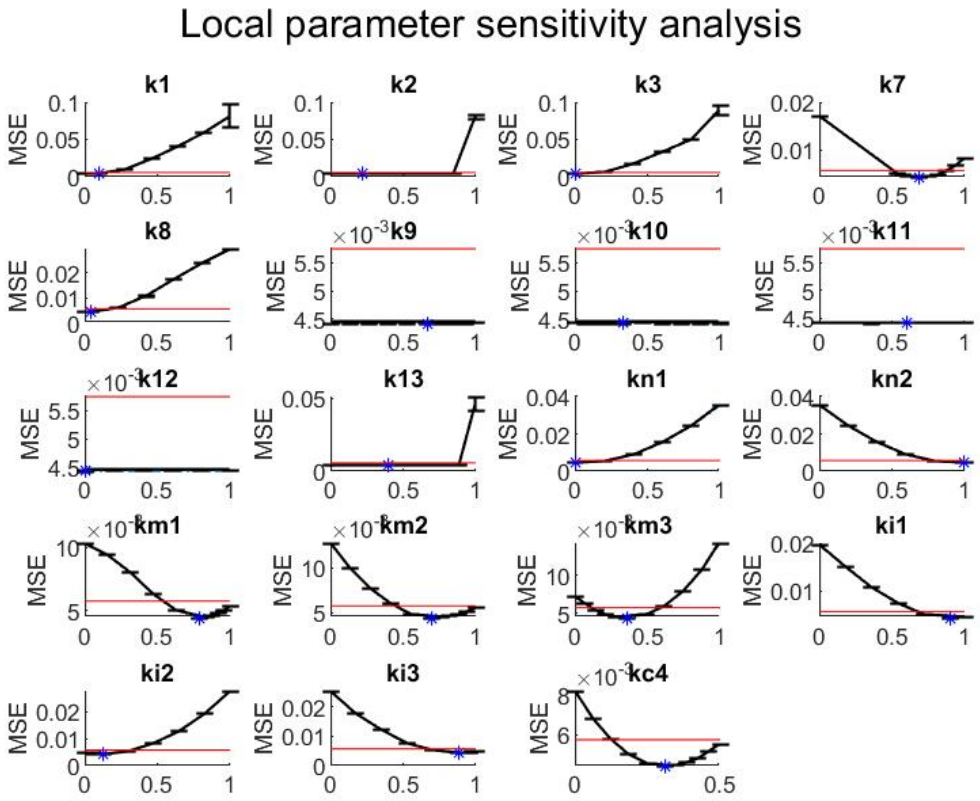


FIGURE 15: Example of Local Parameter Sensitivity Analysis in FALCON. The red lines indicate the threshold used in the *fast* mode. The blue stars indicate the optimal parameter value

2.5.2 Knock-Outs

In order to facilitate the identification of suitable pharmacological targets, we implemented an analysis pipeline to perform *in silico* knock-outs, i.e. reoptimizations of the model after removing each one of the elements at a time. By removing nodes completely from the network, it is possible to simulate the complete blocking or removal of a molecule, for example the

inactivation of a receptor by a blocking ligand. In contrast, by removing only one interaction, more specific network interventions can be simulated, for example the presence mutations modifying the specificity of a protein-protein interaction. Figures 16 and 17 shows the result of the complete analysis of the PDGF signaling model from Trairatphisan *et al.* (2016). The fitness of the re-optimized model when a node or interaction is removed from the network indicates the essentiality of this node or interaction in the original model. Non-essential nodes and interactions can by 'by-passed' by the system, in the sense that a complete blocking of this node or interaction does not result in a complete loss of the capacity of the system to reproduce its functional behavior, given the observations. In contrast, removal of essential nodes and interactions induces a change in the functional properties of the system and a larger difference between the biological measurements and the simulations of the perturbed system.

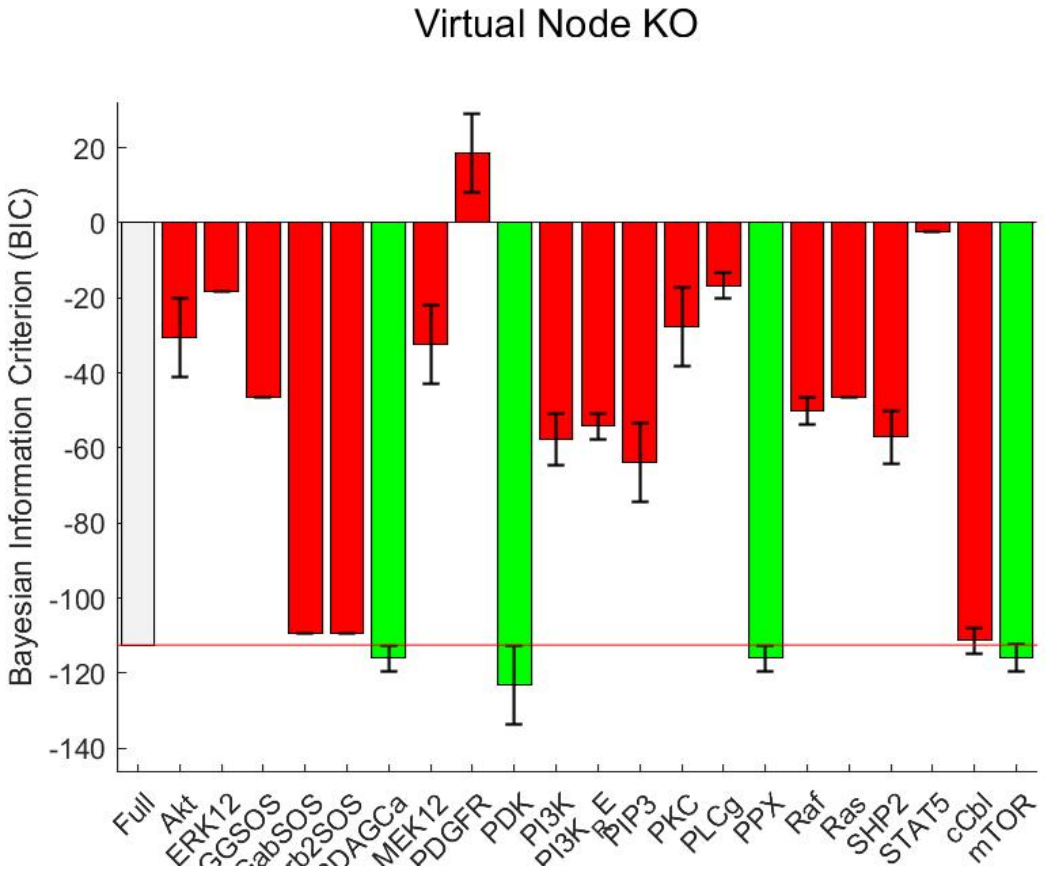


FIGURE 16: Example of virtual node knock-out screening in FALCON. The left-most bar indicates the BIC score of the unmodified optimized model. Each subsequent bar represents the BIC score for one model variant, in which one node has been removed from the network. Low BIC scores indicate better models.

2.5.3 Partial knock-outs

During pharmacological intervention with a regulatory network, the complete inhibition of a component or interaction might not always be the intended result, and partial, controlled

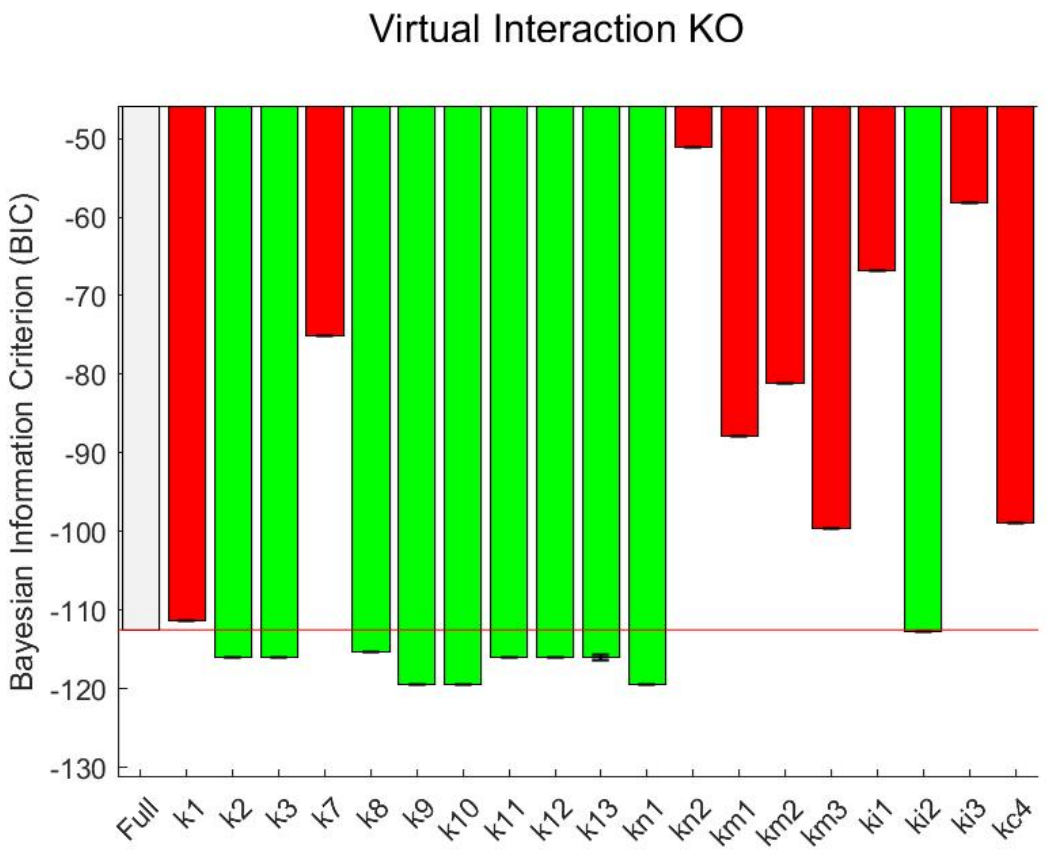


FIGURE 17: Example of virtual interaction knock-out screening in FALCON. The left-most bar indicates the BIC score of the unmodified optimized model. Each subsequent bar represents the BIC score for one model variant, in which one interaction has been removed from the network. Low BIC scores indicate better models.

inhibition might sometimes need to be achieved to attain the desired functional effect. Similarly, the technical limitations of cloning and cultivation techniques limit the degree to which a complete absence of a endogenously expressed molecule can be achieved in practice. For example, the typical experimental designs using RNA interference rarely achieve complete gene silencing (Aagaard & Rossi, 2007). For this reason, we implemented an analysis pipeline in which it is possible to perform a systematic study of the effect of incomplete silencing of each of the network nodes, one at a time. The activity profiles of each of the other components in the re-optimized model can then be examined, therefore helping in determining for each of the nodes, the degree of inhibition necessary to achieve a specified functional effect in a distant node. Figure 18 shows the result of the complete analysis of the PDGF signaling model from Trairatphisan *et al.* (2016).

2.5.4 Resampling Analysis

In addition to the sensitivity analysis, two additional ways of assessing the robustness of the model are proposed. Firstly, it is possible to built a dataset with the same structure and

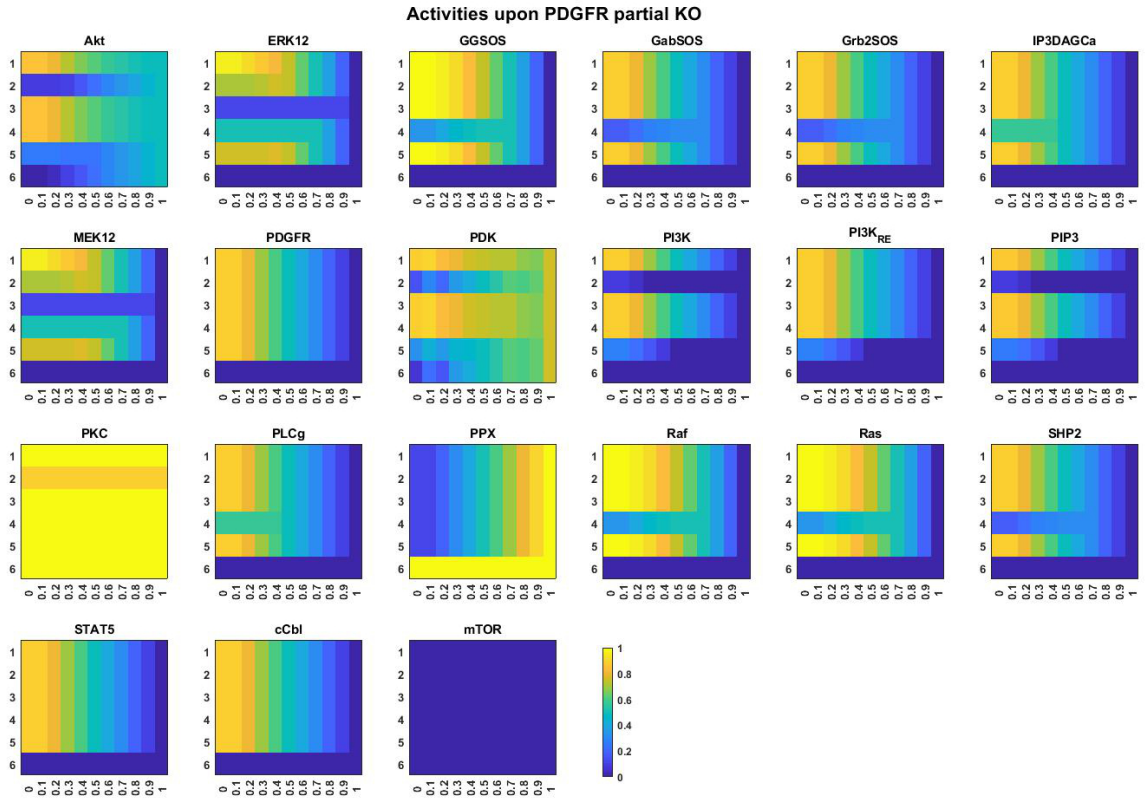


FIGURE 18: Example of partial knock-out analysis for the PDGF signaling model in Trairatphisan *et al.* (2016), showing the effect of increasing silencing of the PDGF receptor (x-axis) on the activities of the other nodes in the system in the different experimental conditions (y-axis). The scale is identical for all heatmaps.

statistical properties as the original dataset by sampling from the data distribution. This procedure, which assumes independence of the model inputs, can be used to assess the model performance with slightly perturbed data, which is a type of uncertainty analysis. In addition, the data pre-processing procedure allows the generation of bootstrapped datasets. The bootstrapping procedure consists in reconstituting an alternative dataset from the original data, by sampling with replacement. Therefore, the new dataset will consist only of instances of the original dataset, but will give a larger importance of some observations and omit some others. By combining these two modes of model exploration, a sufficient level of understanding of the dependency of the model results on the presence of some observations can be attained.

2.6 Integration of regularization schemes

Regularization is used to integrate additional constrains to the loss function of an optimization problem. From a practical perspective, regularization is used to add information to an ill-posed problem, for the specific purpose of making the model learn more generalizable functions. From

a Bayesian perspective, regularization is a way to specify a prior distribution of the parameter values of the model. Regularized objective functions have the form of a sum of the loss function V and the regularization function R: $\min_f \sum V(f(x_i), y_i, \Theta) + \lambda R(\Theta)$.

A number of regularization methods have been implemented in FALCON. The formulas for the regularization functions of the parameter set Θ are detailed in the following section.

2.6.1 Norm-based

- partial-norm for the pruning of unnecessary edges: $R_{1/2}(\Theta) = \sum_i \Theta_i^{1/2}$
- first norm for the pruning of unnecessary negative edges: $R_1(\Theta) = \sum_i \Theta_i$
- second norm for the mitigation of edges importance : $R_2(\Theta) = \sum_i \Theta_i^2$

2.6.2 Group structure-based

- homogeneity across groups: $R_{groups}(\Theta) = \sum_g \sum_i \mid \Theta_i^g - \bar{\Theta^g} \mid$ over all groups ($\bar{\Theta^g}$ is the average for group g)
- uniformity across groups: $R_{unif}(\Theta) = T(\sum_{R \in P} \mid 1_{\Theta_n^t \in R} - \prod_i b_i - a_i \mid)^{-1}$ for all rectangles $R = [a_1, b_1] \times [a_2, b_2] \times ... \times [a_T, b_T]$ such that $0 \leq a_i \leq b_i \leq 1$

2.6.3 Smoothness-based

- time-smoothing: $R_{time}(\Theta) = \sum_{T=2}^{T_{max}} \sum_i \mid \Theta_i^T - \Theta_i^{T-1} \mid$

Chapter 3

FALCON: A Toolbox for the Fast Contextualisation of Logical Networks

Sébastien De Landtsheer[1†], Panuwat Trairatphisan[1†], Philippe Lucarelli[1], and Thomas Sauter[1]

[1]Systems Biology Group, Life Sciences Research Unit, University of Luxembourg, Belvaux, Luxembourg
[†] These authors contributed equally to the manuscript

This study has been published in:

Bioinformatics, 2017 Nov 1; 33(21): 3431-3436

3.1 Introduction to the paper

Modeling biological systems with logical network models is a promising way to gain clinically valuable insights. Ideally, modeling involves maximizing the usage of available biological data to extract knowledge from the data signal, in the form of correct predictions of the systems behavior. Good models not only generate good predictions, but also get a sense of their accuracy, which is an essential property when working under uncertain assumptions. Several frameworks exist in which biological measurements are used to contextualize a network structure representing a simplified version of a non-linear regulatory system. However, they are not adapted to the large networks needed to understand cellular behavior across subsystems, or the large multidimensional datasets increasingly common in biological research.

Dynamic Bayesian Networks (DBNs) can be used to represent the same conditional probability distributions as Probabilistic Boolean Networks (PBNs), a modeling frameworks which has been successfully used to model phosphoproteomic regulatory networks, notably in cancer cells. PBNs have the advantage of representing biological interactions as simplified activating and inhibiting interactions, include logical AND and OR gates to model molecular dependencies, and are robust in the context of uncertainty. In contrast with PBNs, the simulation of DBNs is faster as DBNs converge to the long-term expected probabilities while PBNs often cycle through long cyclic attractors, requiring Monte-Carlo sampling or otherwise computationally intensive techniques.

Existing toolboxes for the optimization of DBNs are not well adapted to the task of exploring biological hypotheses. Indeed, in the case of signaling pathways, existing information about substrate specificities and binding properties (in the form of protein-protein interaction databases and pathway maps) greatly restricts the topological uncertainty to a limited number of hypotheses, reducing the need for large-scale topological optimization and network inference, known NP-hard problems on which many studies on DBNs have concentrated. Similarly, logical relationships between molecules (*e.g.* the requirement of a cofactor, competition,...) are often known or assumed, but these AND and OR logical relationships are not immediately interpretable in a Bayesian context. To be really useful, a modeling toolbox should also be usable by non-experts, account for the specificities of biological measurements and experimental settings, and make the analysis, exploration, and visual interpretation of models directly compatible with the types of research questions in the biomedical field.

In this first paper we propose an algorithm for the search of optimal parameters of DBNs, and we release a usable toolbox able to contextualize DBNs with quantitative biological measurements. We demonstrate the use of the toolbox by testing its performance on a set of existing PBN modeling problems.

3.2 Abstract

Motivation: Mathematical modeling of regulatory networks allows for the discovery of knowledge at the system level. However, existing modeling tools are often computation-heavy and do not offer intuitive ways to explore the model, to test hypotheses or to interpret the results biologically.

Results: We have developed a computational approach to contextualize logical models of regulatory networks with biological measurements based on a probabilistic description of rule-based interactions between the different molecules. Here, we propose a Matlab toolbox, FALCON, to automatically and efficiently build and contextualize networks, which includes a pipeline for conducting parameter analysis, knockouts and easy and fast model investigation. The contextualized models could then provide qualitative and quantitative information about the network and suggest hypotheses about biological processes.

Availability and implementation: FALCON is freely available for non-commercial users on GitHub under the GPLv3 license. The toolbox, installation instructions, full documentation and test datasets are available at https://github.com/sysbiolux/FALCON. FALCON runs under Matlab (MathWorks) and requires the Optimization Toolbox.

Supplementary data are available at Bioinformatics online

3.3 Introduction

The functional characteristics of eukaryotic cells are largely determined by the properties of their regulatory networks. Notwithstanding the vast amount of biological data accumulated over the past decades, a global model of the way these networks determine the phenotypes of both healthy and diseased cells remains elusive. One goal of systems biology is to understand these networks at the highest possible functional level, for example to devise therapeutic strategies for treating patients affected by diseases like cancer.

Numerous mathematical approaches exist to optimize and train regulatory network models against steady-state experimental data (Villaverde & Banga, 2014). Of these, logical models (Le Novère, 2015) are of particular interest, as they are able to capture essential features of the system being modeled and generate biological insights, while requiring less prior knowledge and experimental observations than differential equation models (Morris *et al.*, 2010). Some successful applications include the logical models of yeast cell-cycle protein network (Li *et al.*, 2004), gene regulatory networks (Mendoza *et al.*, 1999), signaling networks (Saez-Rodriguez *et al.*, 2007). In addition, logical models are in general more powerful than statistical models,

as they incorporate the relational information embedded in the network structure, while statistical models aiming at reverse-engineering biological networks from high-throughput data implicitly consider all possible topologies (Penfold & Wild, 2011).

In logical models of systems at steady-state, nodes represent the degree of activation of the constituents of the system at equilibrium and edges represent the logical functions between nodes. These functions can be either linear or non-linear functions of the parent nodes and are combinations of the fundamental AND, OR and NOT Boolean functions.

While Binary Boolean models (Kauffman, 1969) only consider full activation or complete absence, more quantitative approaches, for instance, Probabilistic Boolean Networks (PBNs) (Trairatphisan *et al.*, 2013) and Dynamic Bayesian Networks (DBNs) (Lähdesmäki *et al.*, 2006) can account for intermediate or continuous activation values and allow the integration of data uncertainty. These approaches are usually analyzed by Monte Carlo approaches (Bourdon & Roux, 2016; Trairatphisan *et al.*, 2014), which can be computationally demanding or non-intuitive to use. Here, we propose a tool called FALCON to efficiently contextualize logical regulatory networks based on steady-state experimental data. Our algorithm is based on DBNs and computes the expected value of the nodes by including an algebraic interpretation of the logical gates. The FALCON pipeline is shown in Figure 19.

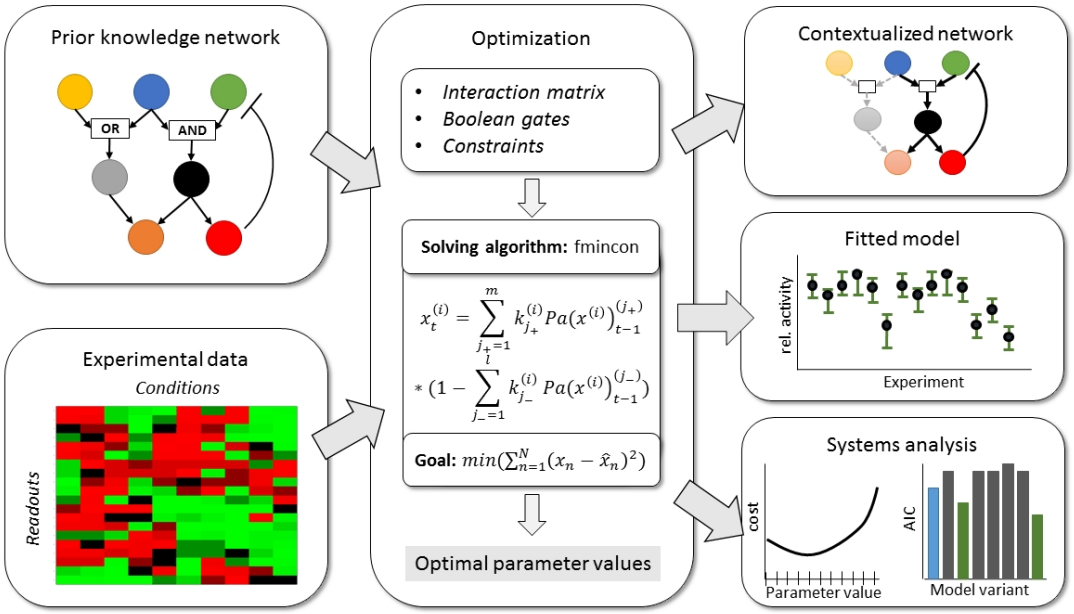


FIGURE 19: The FALCON pipeline. Prior knowledge network and experimental data are combined to generate a network optimization problem. After the optimization process, the properties of the optimal network are then analyzed

3.4 Materials and methods

3.4.1 Modelling of logical networks

FALCON models biological regulatory systems as DBNs, which are directed graphical models defined by the set of n nodes with $X = [0,1]^n$ and the probability distribution $P(X_t|X_{t-1}) = \prod_{i=1}^{n} = P(X_t^{(i)}|Pa(X_t^{(i)}))$ where $X_t^{(i)}$ denotes the ith node at time t and $Pa(X_t^{(i)})$ represents the parents of $X_t^{(i)}$. These conditional probabilities are implicitly formulated by the structure of the network. The different nodes represent the different molecules of the system, with a value corresponding to the degree to which these molecules exist in their active form (for example, phosphorylated proteins). These node values can be understood as the proportion of the molecules in the system being active, or as the probability for a randomly chosen molecule to be active at time t.

In the FALCON framework, each molecular interaction is formulated as a logical predicate associated with a weight quantifying the relative importance of that specific interaction. We model different types of biochemical interactions with two types of edges: positive and negative edges connect activators and inhibitors to their downstream targets. Hyperedges corresponding to the AND and OR logical operations link multiple nodes to an output node, and model the activity of protein complexes and competition, respectively. Each edge and hyperedge is associated to a weight $k_j^{(i)}$ representing the relative influence of the upstream node to the downstream node. Because our modeling framework is grounded in Bayesian theory, the weights need to obey the law of total probability: for each node $X^{(i)}$ having a set j^+ of m activating functions, we ensure the sum of activating weights $\sum_{j^+=1}^{m} k_{j^+}^{(i)} = 1$. Similarly, as weights of inhibiting interactions materialize the relative inhibition of upstream nodes, for nodes having a set j^- of l inhibiting functions, we ensure that $0 \leq \sum_{j^-=1}^{l} k_{j^-}^{(i)} \leq 1$.

Given a network structure established from prior knowledge, a set of parameters (weights) and a set of experimental conditions, the steady-state of the network is computed for each of the conditions and the values of the nodes corresponding to the measured species are recorded. For each one of the conditions, the nodes of the network are initialized with random values, except for the nodes considered as inputs (external to the system) for which the value is determined by the experimental condition and kept constant. The network is then updated repeatedly by computing synchronously for each node the expected value of its probability distribution, given the value of its parent nodes and the weights associated with each interaction.

$$X_t^{(i)} = \sum_{j^+=1}^{m} k_{j+}^{(i)} Pa\big(X^{(i)}\big)_{t-1}^{(j+)} * \Big(1 - \sum_{j^-=1}^{i} k_{j-}^{(i)} Pa(X^{(i)})_{t-1}^{(j-)}\Big) \tag{3.1}$$

Because all nodes at each update are considered as independent, the inputs values of AND logical gates are multiplied. The computation of OR gates follows De Morgans law, *i.e.* the

Biological equivalent	Graphical form	Algebraic computation
Activation	$A \rightarrow Z(k)$	$Z_{t+1} = A_t * k$
Inhibition	$A \dashv Z(k)$	$Z_{t+1} = 1 - (A_t * k)$
Complex formation	A AND $B \rightarrow Z(k)$	$Z_{t+1} = A_t * B_t * k$
Competitive interaction	A OR $B \rightarrow Z(k)$	$Z_{t+1} = 1 - [(1 - A_t) * (1 - B_t) * k]$
Non-competitive interaction	$A \rightarrow Z(k_1)$ $B \rightarrow Z(k_2)$	$Z_{t+1} = A_t * k_1 + B_t * k_2$ (with $k_1 + k_2 = 1$)

TABLE 4: Different types of biological interactions modelled by different Boolean functions and their algebraic representations

complement of the union of two sets is the same as the intersection of their complements. Inputs pointing to the same child node that are not members of a logical gate are summed. Table 4 summarizes the different types of interactions explicitly formulated in our framework. The algebraic formulas used for the computations can be directly derived from the conditional probability tables of the DBN formulation of the logical interactions.

The resulting dynamical system converges to a steady-state where each node value corresponds to the normalized equilibrium concentration of the activated form of the molecule in the system.

3.4.2 Contextualization algorithm

Objective function. To perform the contextualization of the model with experimental data, we extract from the network at steady-state the value of the nodes corresponding to the measurements, compare them with the normalized values from the experimental data and compute the mean squared error (MSE) between the estimated values and the measurements. We minimize this measure of the error by optimizing the value of the weights using a gradient-descent algorithm. To guarantee high efficiency while allowing for arbitrary degrees of recurrence in the networks, we use the interior-point method (Waltz *et al.*, 2006). A scheme of the FALCON workflow is presented in Figure 19.

Rapid optimization. Using the gradient-descent optimization algorithm fmincon with interior-point method, FALCON is able to rapidly estimate the set of weights that minimizes the objective function. Random initialization of the weights is done either from a uniform distribution across the $[0, 1]$ range, or from a truncated normal distribution centered on 0.5, depending on users choice. Normally distributed initial values have been shown to improve learning for deep neural networks (Glorot & Bengio, 2010) and in our hands, increase the speed of convergence of the optimization algorithm.

3.4.3 Subsequent analyses on optimized logical networks

Once a set of parameters has been inferred from a given topology and dataset, a series of additional analyses can be performed to gain more insight into the systems-level properties of the regulatory network being modeled as summarized in Figure 20.

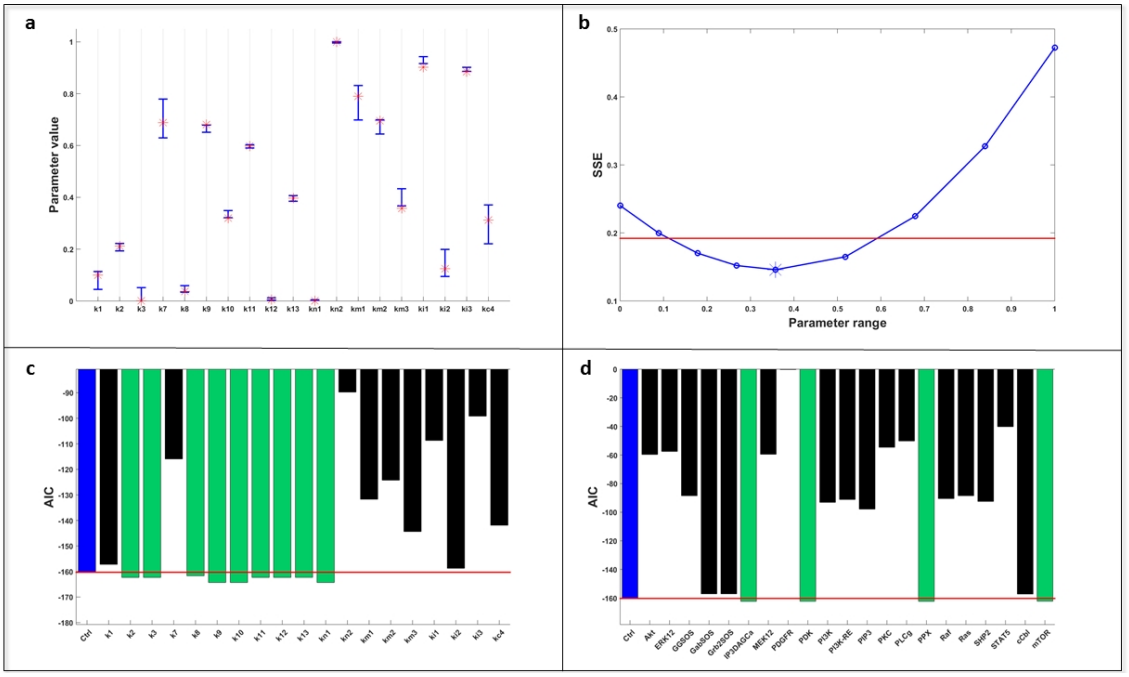


FIGURE 20: Analyses of optimized model in FALCON (PDGF model). a: Parameter robustness analysis ; red stars: optimal parameter values, blue bars: standard deviations of parameter values fitting to 10 resampling datasets. b: Parameter identifiability analysis of parameter km3 from panel a; Red line: threshold used to speed up computations in the fast mode. c: Interaction knock-out analysis. d: Node knock-out analysis. In panels c and d, the color of the bars indicates the sign of the difference with the base model (blue). Green indicate better models (AICmodel¡AICbase), black indicates worse ones. Abbreviations: MSE = mean squared error, AIC = Akaike Information Criterion

Robustness of optimized parameter values. Depending on the topology of the network, the uncertainty in the measurement of some nodes can have more impact on the parameter values of the model than others. FALCON can analyze the uncertainty on inferred parameter values by sampling a user-defined number of artificial datasets based on original experimental measurements and determining the weights of the model in the light of the new data (Fig. 20a). The artificial datasets are constructed from the average experimental measurements and their associated error, assuming normally distributed residuals.

Identifiability analysis. In order to assess the identifiability of the model parameters, an approach similar to Raue *et al.* (2009) is applied (Raue *et al.*, 2009). For each parameter, the algorithm samples the range of possible parameter values $[0, 1]$, and re-optimizes the model under the additional constraint of this parameter being fixed to each one of the sampled values. In order to obtain the most meaningful results we sample the same number of points on both sides of the optimal value. We include the option to skip the most extreme values

based on a threshold determined by the resampling analysis (red line, Fig. 20b), thereby accelerating computations. The resulting MSE profiles allow to determine which parameters are well constrained by the experimental measurements.

Interactions knockouts. FALCON allows the systematic removal of each edge in the network and provides a graphical output showing the effect on the global fitness of the model. The models are compared using the Akaike Information Criterion (Burnham & Anderson, 2004), which balances goodness-of-fit with model complexity (Fig. 20c). By using this additional analysis, it is possible to differentiate the crucial edges of the system from the ones that are dispensable, which can be pruned out.

Nodes knockouts. A frequent goal of systems biology analyses is to identify the crucial molecules of a regulatory network. Often performed via network topological properties (centrality measures), this identification is of particular interest in the case of target discovery efforts. FALCON allows the systematic evaluation of models in which each node is removed from the network. The comparison of these models using the Akaike Information Criterion allows to identify these crucial nodes not only from topological properties but from the effect their removal has on the behavior of the entire system (Fig. 20d).

Differential regulation. In many real-life modeling applications, a system is studied in different contexts. For example, during a drug screen, the same signaling pathways are studied for different cell lines, or over time. One goal of systems biology is to identify differences between the contexts in the way the system is regulated. FALCON automates such analyses by optimizing identical models in parallel for multiple series of experimental conditions. Users can discover which parts of the network are activated or shut down between cell lines/time points, and this may lead to the identification of specific interventions strategies for each context (Figure 21).

3.5 Pipeline and performance

FALCON is a highly efficient optimization tool that is capable of contextualizing small-to-large biological networks. For an easy input of model structure and experimental data, FALCON accepts different file formats (.txt, .xls, .xlsx, .csv) which are subsequently used to build logical models. Inference of network structure, interaction matrices and parameter constraints are fully automated, and the toolbox outputs a user-friendly summary comprising the optimized weights for the different interactions, both in text and graphical forms. To facilitate the use of our toolbox, we included a graphical user interface (GUI) to guide users through the different steps of the workflow. Users who are more comfortable with the MATLAB language can instead choose to use the provided driver script for full flexibility.

To showcase the performance of our toolbox, we provide four examples, including the replication of several studies, each presenting a particular challenge for the toolbox. The results of

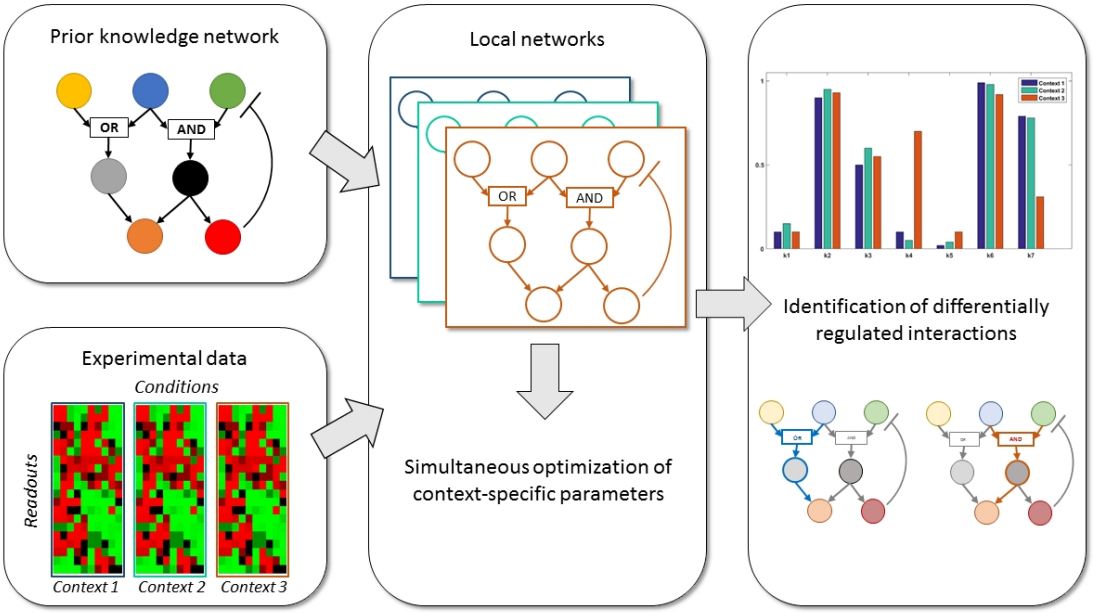


FIGURE 21: Differential analyses in FALCON. The same prior knowledge model is contextualized in parallel with different datasets corresponding to different contexts. Subsequent analysis can identify context-specific parametrizations and topologies

Example	Nodes	Edges/Parameters	Datapoints	Cost	Speed
Toy (artificial)	6	3/3	10	0	‘ 1 s
PDGF	30	19/19	36	0.004	1.3 s
Apoptosis	138	160/41	18	0.017	76 s
MAPK [FALCON]	22	32/32	175	0.036	1.1 s
MAPK [CNORfuzzy]	22	32/92	175	0.032	47.4 s

TABLE 5: Accuracy and computation times for the different examples. Note: The cost is expressed as MSE (mean squared error) and the speed is expressed in seconds (s).

our tests are shown in Table 5. All computations were performed on a desktop PC with 16GB RAM and an Intel®Xeon®CPU E3-1246 v3, 3.5GHz with Matlab 2016b.

Toy model: we demonstrate the basic functionality of FALCON on a 6-node toy model, comprising both positive and negative interactions, as well as a Boolean AND gate. The structure of this network, associated synthetic data and trained model are illustrated in Supplementary Figure S1.

PDGF: we used FALCON to optimize a platelet-derived growth factor signaling model (Trairatphisan *et al.*, 2016), comprising 30 nodes and 37 interactions (19 free parameters). The dataset was assembled from the quantification of 6 proteins by western blot analysis in HEK293 cells expressing a constitutively active form of the PDGF receptor, in the presence or absence of two types of perturbations: single-point mutations of tyrosine residues on the PDGF receptor associated with the recruitment sites of downstream signaling molecules, and kinase

inhibitors. We obtained a fitting cost (MSE=0.0041) and parameter values very similar to the original study, where the tool optPBN (Trairatphisan *et al.*, 2014) was used to perform the optimization, and in accordance with it, we are able to train the network with single perturbations and accurately predict the signaling profiles of combined perturbations experiments (see Supplementary Material).

Apoptosis: we replicated a modified model of a previous study in which a large Boolean model of apoptosis was used to investigate non-linear dose-effects of UV radiation on cultured hepatocytes (Schlatter *et al.*, 2009; Trairatphisan *et al.*, 2014). The model comprises 138 nodes and 160 interactions (41 free parameters). We correctly estimated apoptosis levels and the other associated experimental measures, and could draw the same conclusions as the original study concerning the importance of cross-talks, especially between Caspase 8 and NFκB (see Supplementary Material). While the original study used the software CellNetAnalyzer (Klamt *et al.*, 2007), which uses a multi-value Boolean formalism and concentrates on network properties, a previous replication with the optPBN toolbox (Trairatphisan *et al.*, 2014) could infer more quantitative properties, but at the expense of long computation times. Analysis of this network and data with FALCON is comparatively very fast with up to 170-fold improvement (FALCON: 76 seconds; optPBN: 4 hours 40 minutes) and we obtained a fitting cost (FALCON: MSE=0.017) comparable with the previous studies (optPBN: MSE=0.011; Schlatter *et al.* (2009): MSE=0.013). In comparison, CellNetAnalyzer, using discrete Boolean modeling and only able to consider either full activation of complete inactivity of the molecules, achieves a worse fit (MSE=0.056). The comparison of the inferred molecular states of optPBN and FALCON can be found in Supplementary Figure S6.

MAPK: we compared the performance of our tool with the software CellNOptR (MacNamara *et al.*, 2012; Terfve *et al.*, 2012) in the fuzzy logic mode (CNORfuzzy) for quantitative optimization of model states. Using the toy example provided, which is the optimized network of the DREAM4 challenge and contains 22 nodes, 36 interactions and 25 experimental conditions (Prill *et al.*, 2010), we obtained a similar fitting cost with FALCON (MSE=0.036) and with CellNOptR (MSE=0.032) but with a gain of speed of about 44 times (see Table 5).

3.6 Discussion

We present FALCON as an alternative tool for the efficient optimization and comprehensive analysis of logical models of regulatory networks. Our modeling framework, based on DBNs, is able to determine qualitative and quantitative features of the systems being modeled. Node values, being comprised in the interval $[0, 1]$, represent the probabilities for molecules to be in their active state at equilibrium. They can also be understood as the normalized average activities of the nodes. The computed parameters, or weights, also comprised in the interval $[0, 1]$ and subject to the law of total probability, represent the probabilities for the designated

interactions to influence downstream nodes. They can also be interpreted as the relative influences of the parent nodes on their children nodes and are useful in assessing the flow of the signal transduction.

FALCON, through its GUI, is easy to use for scientists without extensive modeling experience. FALCON is also very fast compared to similar tools based on PBNs, and surpassed CellNOptR in our test. The low computation costs make it possible to analyze the models at the systems level through a series of bundled additional analyses which allow to answer a number of biologically important questions: whether the parameter values are well constrained by the available data, how the experimental error influences the confidence in the parameter values, and which are the nodes and interactions most crucial to the behavior of the system versus the ones that can be pruned out. Together, our results suggest that FALCON is a very useful software for rapid model exploration, especially for large networks and large datasets.

Compared to the popular package CellNOptR, the FALCON pipeline is faster in contextualizing a small graphical model with quantitative data. The inferred parameters are also more intuitively understandable as the relative strength of the interactions, while CellNOptR combines linear and Hills equations in a way that does not encourage direct interpretation. This relative complex formulation, together with the multiple concurrent formalisms proposed and the increased computational cost suggest reserving this tool for more complex tasks, while FALCON is better adapted for exploratory studies of larger networks and datasets.

Future development of the FALCON toolbox will include full compatibility with established model representation formats (SBML-Qual, Bio-PAX), and the conversion of the toolbox to other languages, like R, Python and C++. One particular aspect that we regard as highly interesting is the use of FALCON to explore model topologies in a large-scale, systematic way to uncover previously unknown mechanisms in regulatory networks.

In terms of applications, we demonstrated that FALCON is applicable to model signal transduction networks and could easily be extended to study other biological regulatory systems. We envision that FALCON has the potential to be widely adopted by the computational biology community, including biologists with limited programming experience.

Acknowledgments

We would like to acknowledge Dr Jun Pang, Dr Andrzej Mizera, Prof. Dr Dagmar Kulms, Greta del Mistro and Dr Thomas Pfau for the fruitful discussions and suggestions on the modeling and analytical pipelines.

Funding

This project has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 642295 (MEL-PLEX) and the Luxembourg National Research Fund (FNR) within the projects MelanomSensitivity (BMBF/BM/7643621) and ALgoReCEll (INTER/ANR/15/11191283).

Conflict of Interest: none declared.

Chapter 4

Using Regularization to Infer Cell Line Specificity in Logical Network Models of Signaling Pathways

Sébastien De Landtsheer [1], Philippe Lucarelli [1] and Thomas Sauter [1,*]

[1] Systems Biology Group, Life Sciences Research Unit, University of Luxembourg, Belvaux, Luxembourg

This study has been published in:

Frontiers in Physiology, 2018 May 22; 9:550

4.1 Introduction to the paper

Models of cellular regulatory networks have the potential to increase our understanding of disease processes like cancer. One of the main difficulties of modeling is the choice of the correct size for a model. The ideal size for a model depends on the amount of data available to parametrize it, the amount of signal in said data, and the correlation between the samples. However, as these are usually difficult to determine, one frequent technique is to restrict the size of the model as the result of the optimization process itself, a process known as regularization. In optimizing a regularized model, parameters can be subjected to diverse constrains which tend to either eliminate them from the problem structure, maintain them as small as possible, or pair them with another free parameter. Only with increasing data can these constrains be lifted and the size of the model increased.

Regularization is heavily used, in different forms, when using large models. Many regularization schemes have been developed, notably the so-called L_1 and L_2 norms, parameter dropping, etc. When large models are used in conjunction with insufficient amounts of data to model a highly non-linear process, the phenomenon of overfitting, in which the performance of the model is high on the training data but low on data that was not part of the training set, can occur, with devastating consequences for the accuracy of the model predictions.

In this paper, we reason that the same mutation in different cell types will produce similar proximal effects, and that the difference in distal functional effects between cell types (e.g. the differential sensitivity to a certain drug) results from the non-linear combination of the effects of different mutations. Under this assumption, we aim to analyze the differences in the parametrizations of cell type-specific models in the light of the data and determine when to smooth these differences out. In practice, when the differences are small enough compared to the amount of evidence supporting them, it can be assumed that they are due to noise and are not biologically meaningful.

Unfortunately, such regularization scheme does not exist for multiple cell types. We therefore formalize a regularization function which quantifies the local density of parameters in the parameter space and the difference with a uniform distribution, and we implement this function as a regularization of cell type-specific parameters in our Matlab toolbox FALCON. We demonstrate the usability of this method first with a toy model, then by reanalyzing the data of a published study.

4.2 Abstract

Understanding the functional properties of cells of different origins is a long-standing challenge of personalized medicine. Especially in cancer, the high heterogeneity observed in patients slows down the development of effective cures. The molecular differences between cell types or between healthy and diseased cellular states are usually determined by the wiring of regulatory networks. Understanding these molecular and cellular differences at the systems level would improve patient stratification and facilitate the design of rational intervention strategies. Models of cellular regulatory networks frequently make weak assumptions about the distribution of model parameters across cell types or patients. These assumptions are usually expressed in the form of regularization of the objective function of the optimization problem. We propose a new method of regularization for network models of signaling pathways based on the local density of the inferred parameter values within the parameter space. Our method reduces the complexity of models by creating groups of cell line-specific parameters which can then be optimized together. We demonstrate the use of our method by recovering the correct topology and inferring accurate values of the parameters of a small synthetic model. To show the value of our method in a realistic setting, we re-analyze a recently published phosphoproteomic dataset from a panel of 14 colon cancer cell lines. We conclude that our method efficiently reduces model complexity and helps recovering context-specific regulatory information.

4.3 Introduction

One goal of Systems Biology is to understand emerging functional properties of biological systems from the interactions of their components (Van der leeuw, 2004). Such understanding would allow the design of new pharmacological strategies to treat diseases that arise when these systems do not function adequately, like cancer. One frequent approach is to map experimental measurements to the model variables of the system, and infer the most likely parametrization. To be useful, a well-parametrized model of a complex system should not only be able to predict non-obvious, non-linear behaviors, but also provide a mechanistic explanation for these behaviors and to suggest hypotheses about ways to control the system.

The most informative modeling approaches include prior information about the system (Aldridge *et al.*, 2006). Classically, dynamical systems like regulatory networks of mammalian cells are modelled with systems of ordinary differential equations, describing in detail the status of chemical species like proteins or membrane receptors over time. Alternatively, logical models (Le Novère, 2015; Morris *et al.*, 2010; Hill *et al.*, 2012) were introduced several decades ago for the modeling of regulatory networks (Kauffman, 1969). As they are simpler in their formulation, they are easier to handle computationally, scale better to large models and datasets, and are easier to interpret. The prior knowledge used to construct logical network models frequently comes from reviewing the literature of a certain mechanism, disease or signaling

pathway, and may be summarized in a database like STRING, Reactome or WikiPathways (Szklarczyk *et al.*, 2017; Joshi-Tope *et al.*, 2005; Kutmon *et al.*, 2016; Rigden *et al.*, 2016).

Logical models can be used to model stochastic processes. Probabilistic Boolean Networks (Shmulevich *et al.*, 2002) have been introduced to simulate logical models in the presence of uncertainty, as they allow combining multiple Boolean networks with the respective continuous selection probabilities in one mathematical model. They have successfully been applied to the modeling of biological regulatory networks (Trairatphisan *et al.*, 2013). This framework can be generalized to Dynamic Bayesian Networks (DBNs), a general class of models that includes Hidden Markov models and Kalman filters (Murphy, 2002), and can be used to represent the same joint probabilities between variables. In a graphical model of a DBN, the values of the different nodes represent the probabilities for randomly chosen molecules to be in an active state, while the edges represent the probabilities of the parent nodes to activate their targets. Network update is performed according to the laws of probabilities.

There is, however, a number of impediments to successful biomolecular modeling. Firstly, the prior knowledge used to build the model could be inaccurate, or more frequently, incomplete, or both. In other words, compared to the true network, databases likely contain additional edges, as well as miss others. Secondly, the information contained in databases is often generic, collected across cell types, genetic backgrounds, and experimental conditions. Given an interaction graph and a series of contexts (cell types, patients), the task of determining which interactions are context-specific and which ones are context-independent rapidly becomes intractable. This task is however essential to reduce the model complexity, as overly complex models are prone to overfitting (thus less generalizable), computationally expensive, and might be less interpretable than simpler ones. In addition, identification of the most variable model parameters between contexts has the potential to be directly informative about the mechanisms at play and help draw parallels between contexts.

Inter-patient variability is an important factor for many diseases, and in particular cancer. Intra-tumor heterogeneity has been recognized for a long time (Fidler *et al.*, 1982) and it has been established that the heterogeneity of cell lines isolated from different patients spans the genomic, epigenetic, transcriptomic, and proteomic levels, resulting in large phenotypic differences, even within the same tissue of origin (Hoadley *et al.*, 2014). Additionally, the patients' own genetic backgrounds and the tumor micro-environment also play a role in increasing the heterogeneity of clinical responses (Marusyk & Polyak, 2010; Junttila & De Sauvage, 2013; Zhou *et al.*, 2008). However, recent successes in matching a biomarker with the sensitivity to certain targeted anti-cancer therapies, notably in the case of HER2-overexpressing breast cancer (Vogel *et al.*, 2002), EGFR-mutated non-small-cell lung cancer (Lynch *et al.*, 2004), BCR-ABL fusions in chronic myelogeneous leukemia (Sherbenou & Druker, 2007), and $BRAF^{V600E}$-mutant melanoma (Bollag *et al.*, 2010) suggest that the general approach of targeting specific mechanisms in subsets of patients harboring functionally similar tumors is clinically promising.

A number of methods have been devised for the general task of variable selection. Various methods rely on the intuitive notion of comparing models comprising different subsets of the independent variables (Hocking, 1976). This strategy is however problematic for several reasons. Firstly, the number of possible subsets grows very fast with the number of variables, leading to the infeasibility of testing them all. Secondly, repeatedly optimizing a model structure using the same dataset violates the central assumptions of the F-tests or χ^2-based statistics used for comparisons, which are designed to test a single hypothesis. Strategies like forward-selection, backwards elimination, or combinations of both are consequently affected by numerous problems, notably biased parameter estimation and artificially low p-values (Burnham & Anderson, 2002; Harrell, 1995).

Fitting an overspecified model first and clustering the parameters in a second step is not a sound method to achieve sparsity, as the parameter estimates might not be stable, resulting in inaccurate clustering. Furthermore, the two objectives are not coupled, which is problematic: a small difference between the values of two parameters might or might not be supported by the data. It makes more sense to specify our assumptions about the distribution of the parameter values as part of the objective function. Regularization is a technique for adding prior information to a regression problem. It consists in adding to the loss function a function of the parameters alone. More formally, when attempting to learn the parameter set θ from dataset $X = [x_1, x_2, ..., x_n]$ with a model M, the objective function O takes the form:

$$O = f(M(X, \theta), X) + \lambda g(\theta) \tag{4.1}$$

where f is the loss function, for example the sum of squared errors. The hyperparameter λ is used to balance goodness-of-fit with the regularization objective $g(\theta)$. The most common form of regularization is the Tikhonov regularization (Tikhonov, 1963), also called *ridge regression*, which materializes the assumption that small model parameters are more probable than larger ones. Also called the L_2 norm, the Tikhonov regularization term takes the form:

$$g(\theta) = \sum_{j=1}^{T} (\theta_j)^2 \tag{4.2}$$

where T is the number of parameters of the model. The L_2 norm is used to impose a penalty on large parameter values. Its popularity is due to the fact that the function is convex, continuous and differentiable everywhere, and is therefore well adapted to gradient descent optimization. It is mostly used in predictive models to avoid overfitting and produces models that are more generalizable. Because the gradient of this function becomes very small around zero, Tikhonov regularization does not achieve sparsity under most conditions and therefore does not perform variable selection, however this can be solved by the use of thresholds.

Intuitively, the most sensible sparsity constraint should be the L_0 norm, or the cardinality of the non-zero parameter set:

$$g(\theta) = \sum_{j=1}^{T} 1_{(\theta_j \neq 0)} \tag{4.3}$$

where $1_{(C)}$ is the *indicator* function, and is equal to the number of cases where condition C is true. However, this is usually not feasible in practice, as this function is discontinuous and cannot be used in many optimization algorithms. A good approximation is the L_1 norm, which sums the absolute values of the parameters, without squaring them:

$$g(\theta) = \sum_{j=1}^{T} |\theta_j| \tag{4.4}$$

The L_1 norm, or LASSO (Tibshirani, 1996) can be used to reduce the size of a model by efficiently removing variables (*i.e.* set their coefficients to zero) which contribute the least to the model. Importantly, by screening a range of regularization parameter λ, it is possible to order the variables according to their importance. It is natural to use it then, for contextualizing models of biological systems with measurements from different contexts to point to their differences. Different approaches have used the L_1 norm to contextualize network models of signal transduction in mammalian cells. However the assumption is either that there is no relationship between the different cell lines (Eduati *et al.*, 2017; Lucarelli *et al.*, 2018), or that the differences to the mean value should be minimized (Merkle *et al.*, 2016). While the latter works in the case of only two cell lines, it does not when comparing more. The reason is that heterogeneity between cell lines is expected, and we know that different mechanisms are at play in a given experiment. By penalizing any difference, such regularization does not allow parameters to have two or more possible values. However, cancer-related perturbations to molecular interactions occur in discrete steps. Driver mutations often result in the complete loss of the function of a certain protein, for example p53, or constitutive enzymatic activity, for example the common mutation of genes of the RAS family (Kandoth *et al.*, 2013). The desired regularization should therefore penalize differences between contexts but allow for a structure in the parameter space. While a number of methodologies exist (Hill *et al.*, 2012; Dondelinger *et al.*, 2013) to regularize network models of signaling pathways for time-stamped data, in that case the structure of the prior on the parameter space is known, as time is oriented. We propose that the correct assumption for analyzing perturbation data from multiple cell lines, cell types, or across patients, is that network parameter values would form *clusters* corresponding to the most common signaling deregulations. However, methods to efficiently identify the parameters of a biological model and cluster them at the same time are missing.

The general problem of regularizing a model towards a specific, although unknown, structure has been investigated before. The vast majority of the proposed methods combine L_1 and L_2 norms in various ways. Group LASSO (Yuan & Lin, 2006) was introduced to allow the selection of entire groups of variables. This was then extended to a hierarchical selection of nested groups of variables (Zhao *et al.*, 2009), partially overlapping groups of variables (Jacob *et al.*, 2009),

and to the induction of sparsity within groups by penalizing for pairwise differences between coefficients of variables belonging to the same group, with the OSCAR algorithm (Petry, 2006) and the clustered LASSO(She, 2010). Later Simon *et. al* proposed the sparse group LASSO (Simon *et al.*, 2012), a modification of the *elastic net* criterion proposed by Zou *et al.*, which combines the L_1 and L_2 norms (Zou & Hastie, 2005). The fused LASSO (Tibshirani *et al.*, 2005) is applicable when there is a natural ordering in the model variables, like time-stamped or spatially organized data. Several groups have tried to decouple the steps of clustering and model fitting, either by considering all possible clusters (Jenatton *et al.*, 2009) or by applying first hierarchical clustering based on the measurements covariance matrix (Bühlmann *et al.*, 2013).

While these approaches have proven useful in some cases (Zhang *et al.*, 2014; Steiert *et al.*, 2016), they do not apply well to the case of regulation networks, because group zero-sparsity (removal of entire groups of variables, as opposed to within-group sparsity) is not necessarily desired, except in the case of network pruning. We therefore implemented a regularized version of the objective function of the FALCON toolbox (De Landtsheer *et al.*, 2017), to lower the degrees of freedom of the model by encouraging the grouping of model parameters across contexts, regardless of the number of groups. This can be achieved by detecting anomalies in the parameter values distribution, assigning a penalty to groups of values more alike the reference null distribution. In our case (Bayesian Networks), the uniform distribution $[0-1]$ is assumed to better represent the prior of uncorrelated parameter values, as they are usually interpreted as probabilities. Under different modeling formalisms, other distributions would be more appropriate, for example for ODE-based or constraint-based models. We show how the penality correlates with other measures, with unsupervised clustering, and we demonstrate the use of regularized fitting, first on a small synthetic network model, then with biological data.

4.4 Methods

4.4.1 Algorithm

We propose a measure of uniformity of the parameter values distribution modified from previous work in the field of quasi-random sequences (Sobol, 1976). Given a parameter space $\mathbf{P}$ and N parameter vectors with T parameters $\theta_1, \theta_2, ..., \theta_N$, with $\theta_n = \{\theta_n^1, \theta_n^2, ..., \theta_n^T\}$, we compute for each $t \in T$ the average absolute deviation from the expected local density of points D_t with:

$$D_t = \sum_{\mathbf{R} \in \mathbf{P}} |1_{(\theta_n^t \in \mathbf{R})} - Vol(\mathbf{R})| \tag{4.5}$$

for all rectangles $\mathbf{R} = [a_1, b_1] \times [a_2, b_2] \times ... \times [a_T, b_T]$ such that $0 \leq a_i \leq b_i \leq 1$, and with $Vol(\mathbf{R})$ being the volume of the T-dimensional rectangle $\mathbf{R}$.

$$Vol(\mathbf{R}) = \prod_i b_i - a_i \tag{4.6}$$

The first term in equation 4.5 represents the *observed* density of points, while the second one represents the *expected* density. These two quantities are equal in the case of perfect uniformity. We then define the *uniformity* U of the parameter vector as the inverse of the average deviation over the T parameters:

$$U_t = \frac{T}{D_t} \tag{4.7}$$

and the *uniformity* of an entire model parameter set as the average over all vectors:

$$U = \frac{1}{N} \sum_{i=1}^{N} U_i \tag{4.8}$$

In one dimension, this metric has an intuitive interpretation, as shown in figure 22: when parameter values are as different as they could be, the expected difference between any two values can be calculated from their relative rank in the set. For example, the distance between two successive observations is $\theta_n^t - \theta_{n-1}^t = 1/N$. When values cluster together, they create windows in which the local density is either higher or lower than this expected value. Note that in one dimension, the rectangles $\mathbf{R}$ are equivalent to the distance between the points, and to the convex hull of these points, while it is not true in higher dimensions.

4.4.2 Uniformity as a penality in regularized fitting

We analyze the sensitivity of our new metric to the amount of structure in sets of model parameter values by computing it for a large number of sets of uniformly, independently distributed random values. We compare uniformity with the standard deviation, with the results of the Kolmogonov-Smirnov (K-S) (Massey, 1951) and Anderson-Darling (A-D) tests (Anderson & Darling, 1954), and with the sum of pairwise distances. The two non-parametric statistical tests aim at comparing the empirical distribution of the values in the set with a reference distribution, in this case the uniform distribution. The sum of pairwise distances is used in (Petry, 2006; She, 2010), the standard deviation is examplative of measures of spread around a single value, like in (Merkle *et al.*, 2016). In addition, we compute for each set the optimal number of clusters (explaining 90% of the variance) using the k-means algorithm and the elbow method (Ketchen & Shook, 1996). Using the inferred number of clusters, we compute the sum of intra-cluster distances. We performed this comparison with 10^4 vectors. Also, to

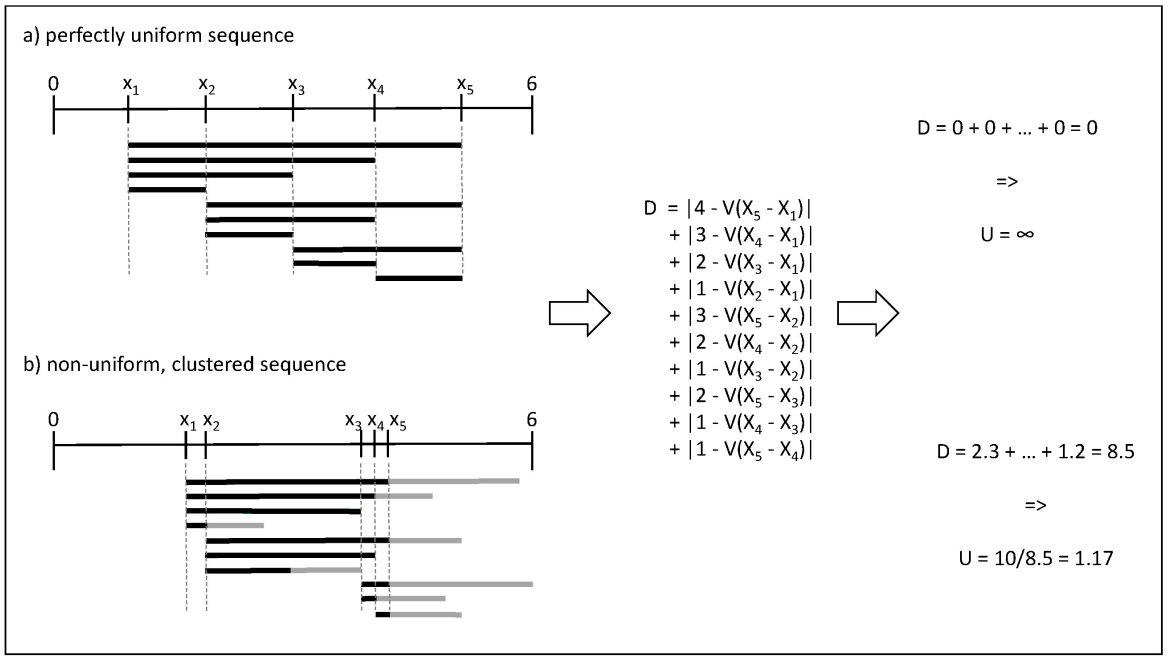


FIGURE 22: Illustration of the computation of uniformity for two sets of 5 parameter values within the range $[0, 6]$. (A) In the first case, all pairwise distances are equal to the expectation given the rank of the value in the set. (B) In the second case, the grey bars indicate the differences compared to the expected density in a given interval.

assess the usability of this metric for large-scale computations, we compare the running time of the different computations for sets of size 10, 20, and 40, simulating models with increasing number of contexts.

To illustrate that the use of uniformity as a penalization in an objective function results in the convergence of parameter values into clusters, we iterate a gradient descent process for random sets of uniformly, independently distributed random values. This is equivalent as optimizing a null model using uniformity as a regularizing function, and shows the effect of this penalization in the absence of data. We used gradient descent (using empirical gradients and the interior-point method) with a learning rate of 10^{-3}, collect the shape of the set over 100 updates, and we compare with the centroids of the k-means clustering. All computations were done using Matlab 2017a on a standard desktop computer which specifications are detailed in section 4.4.3.3.

4.4.3 Modeling experiments

Modeling experiments in this paper used the toolbox FALCON (De Landtsheer *et al.*, 2017), a Matlab-based versatile tool to contextualize logical models of regulatory networks. Briefly, FALCON uses a Dynamic Bayesian framework (Lähdesmäki *et al.*, 2006) in which Boolean operations are explicitly defined as arithmetic, continuous logical functions. FALCON emulates a Probabilistic Boolean Network with user-defined topology and uses experimental data

from perturbation assays to optimize the weights of the network, which represent the relative activating and inhibiting influences of the network components with respect to the logical functions. For the large-scale analysis of biological data, we used a custom implementation of FALCON running on a high-performance computing platform which specifications are detailed in section 4.4.3.3.

$$O = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 + \lambda U(\theta) \tag{4.9}$$

where Y is the vector of measurements for the observed nodes, $\hat{Y}$ is the vector of corresponding predictions and $U(\theta)$ is the uniformity of the parameter set θ across contexts, as defined by equations 4.5 to 4.8 above, with λ being a scalar that controls the relative contribution of the penality to the objective function. The code and data files used for both the synthetic model and the biological example are available at the address https://github.com/sysbiolux/FALCON. Additional driver scripts are provided in the Supplementary Materials.

4.4.3.1 Synthetic toy model

In order to assess the use of our regularization scheme for finding context-specific parameters, we design a simple toy model with 7 nodes and 9 edges. Two of these nodes are inputs, while two others are measured. We set the model parameters differently for four conceptual cell lines, in such a way that while most parameters are conserved, some would be different, and shared across several (but not all) cell lines. Figure 23 shows a graphical representation of the network, the values chosen for the model parameters, and the final synthetic data used for model fitting.

To realistically simulate biological data, we use our toy model to generate synthetic steady-state data for the measured nodes by simulating the network with different combinations of values for the input nodes, thereby mimicking a designed, perturbation experiment. We simulate noise in the data by adding a two-component Gaussian perturbation around the theoretical value, as explained in Supplementary Methods. The magnitude of the perturbation was chosen to reflect the signal-to-noise ratio of typical biological measurements, for example phosphoproteomics or microarray data.

4.4.3.2 Biological dataset

To show the usefulness of our approach in a biological setting, we reanalyze the dataset from (Eduati *et al.*, 2017), in which the authors measured 14 phosphoproteins under 43 different perturbed conditions (combinations of 5 stimuli and 7 inhibitors) in 14 colorectal cancer cell lines. Using CellNetOpt (Terfve *et al.*, 2012), they contextualized independent logical ODE models (Wittmann *et al.*, 2009) for each cell line, and proceed to train a statistical model using

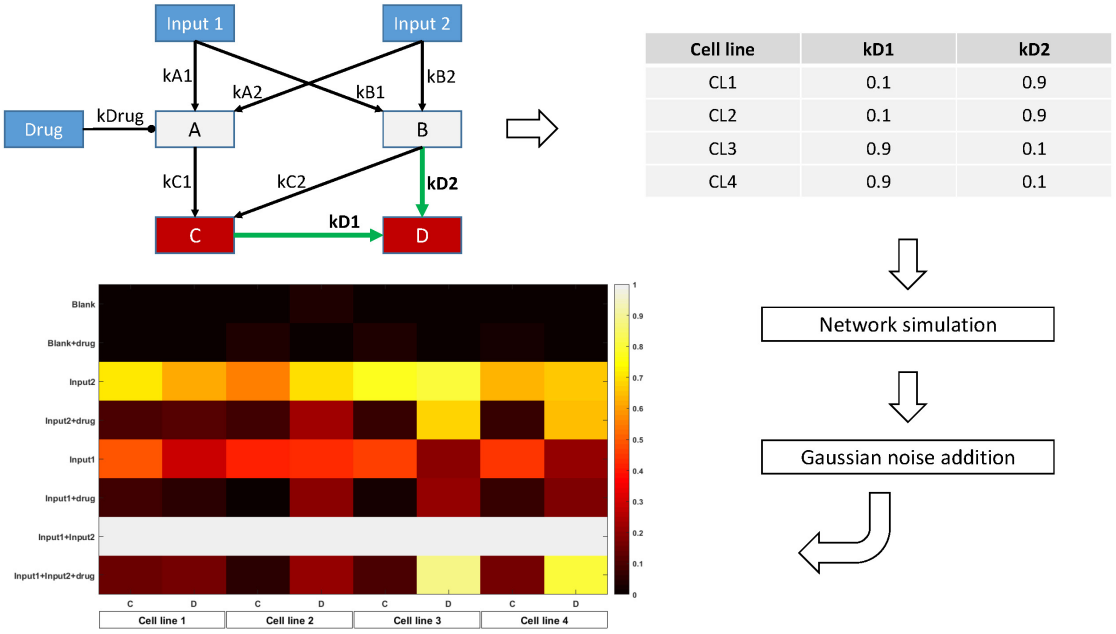


FIGURE 23: Overview of the toy model design. The topology is parametrized in order to display two-by-two similarity between cell lines. For each cell line, the Bayesian Network is simulated with the corresponding parameter values for 8 different combinations of the input nodes values. Random Gaussian noise is added to the values of the two output nodes C and D, simulating biological measurements. The heatmap shows the final node values for each condition, cell line, and node.

the cell-specific parameters to predict the responsiveness of the cell lines to a panel of drugs. This study provides an example of the use of system-level analyses to gain understanding of functional properties that cannot be inferred by genomic features alone. We normalized the data (log_2 difference compared to control) linearly to the $[0-1]$ range across cell lines.

Logical ODE models like the one used by Eduati *et al.* rely on a transformation of the discrete state-space of Boolean models into a continuous one, in such a way that Boolean behavior is preserved on the vertices of the unit cube, i.e. when the inputs are in $\{0,1\}$. While there are many such possible transformations (Wittmann *et al.*, 2009), the authors chose to use normalized Hill cubes, which are sigmoidal functions of the inputs. The strength of such an approach is the ability to take into account the non-linear 'switch-like' nature of molecular interactions, however at the expense of doubling the number of free parameters (Hill functions are defined by a threshold and a slope). In contrast, our approach uses maximum one parameter per interaction and is restricted to linear relationships, which ensures coherence with the laws of probabilities. To infer the DBN model corresponding to the logical ODE model proposed by Eduati *et al.*, we kept the original topological information, but defined the update function for each node by a multivariate linear function of its parent nodes. In our framework, if two nodes A and B are both activators of a third node X, we have for each time-step t: $X_t = k_A A_{t-1} + k_B B_{t-1}$ with probabilities $0 \leq k_A \leq 1$ and $k_B = 1 - k_A$. Similarly, if a node

X is activated by node A but inhibited by node B, we have $X_t = A_{t-1}k_B(1 - B_{t-1})$ with probability $0 \leq k_B \leq 1$.

We used the phosphoprotein data to fit the probabilities for each interaction simultaneously for all cell lines. The complete model comprised 363 nodes and 1106 parameters. The objective function included a penalty computed from the average uniformity of the parameters across cell lines, according to equations 4.5 to 4.8. We optimized 49 models, varying the hyperparameter λ from 2^{-20} to 2^5, and we recovered the optimal parametrization for each cell line in the form of regularization paths. We used the value of 0.01 as threshold for deciding if two parameters should be merged into a single one. For each value of the regularization strength λ, we computed the mean squared error (MSE) and the number of different parameters P in the regularized model, and from these calculate the Bayesian Information Criterion (BIC), which we calculate as $Nlog(MSE)+log(N)P$, with N the number of individual points in the dataset. Lower BIC values indicate models with favorable balance between goodness-of-fit and model complexity (Schwarz, 1978; Burnham & Anderson, 2004).

We selected the model with the lowest BIC for further analyses. We grouped cell line-specific parameters together using the above-mentioned threshold, and re-optimized the model using the obtained topology without the regularization term, in order to obtain unbiased parameter estimates. We performed hierarchical clustering with 1000 bootstrap resamplings on the parameter values using WPGMA and euclidian distance.

Furthermore, we investigated whether the recovered parameter values are associated with drug sensitivity. We downloaded the IC_{50} values for the 14 cell lines and 83 drugs directly targeting either one of the network's nodes or a target used in clinical practice to treat colorectal cancer from the Genomics of Drug Sensitivity in Cancer database (www.cancerrxgene.org). We computed the linear regression models between each drug and each of the 31 parameters which showed high variability between cell lines (CV$\geq$ 10%). The F-statistic was used to compute a p-value for each test, and q-values were computed from these, using the Benjamini-Hochberg procedure to control the False Discovery Rate.

4.4.3.3 Materials

- Hardware
  - Synthetic model: standard desktop computer equipped with an Intel Xeon E3-1241 CPU clocked at 3.50GHz and 16GB of RAM under Windows 7
  - Biological example: high-performance computing platform with 49 nodes running Matlab2017a, each node consisted of one core of a Xeon-L5640 clocked at 2.26GHz with 3GB RAM
- Software

- Matlab 2017a (Mathworks, Inc.)
- FALCON toolbox (https://github.com/sysbiolux/FALCON)
- Optimization Toolbox (http://nl.mathworks.com/products/optimization/)
- Parallel Computing Toolbox (http://nl.mathworks.com/help/distcomp/)
- Bioinformatics toolbox (http://nl.mathworks.com/help/bioinfo/) (optional)

4.5 Results

4.5.1 Uniformity as a measure of structure

We computed the uniformity U, the standard deviation, the sum of pairwise distances, the K-S statistic, the A-D statistic, and the optimal number of clusters using the k-means algorithm and the elbow method, for 10^4 one-dimensional sets of uniformly, independently distributed random values. The complete correlation plots are presented in Supplementary Materials. We always show uniformity U on the logarithmic scale. Figure 24A shows the relation between uniformity and the standard deviation, while Figure 24B shows the correlation between uniformity U and the p-value of the K-S test. Similar results were obtained with the A-D test. The relationship between uniformity, the standard deviation, and the K-S p-value are further explored in Figure 24C, and the computation times are compared in Figure 24D.

Firstly, $log(U)$ is positively correlated with the p-value of the K-S and A-D non-parametric tests evaluating the distance to the reference uniform distribution, showing that low uniformity is indicative of structure. Secondly, the comparison with the standard deviation shows that low-uniformity sets can have drastically different standard deviations, but that the inverse is not true. This is explained by the fact that sets with tightly clustered values will nevertheless be spread around the global average if there is more than one cluster. Figure 24C shows a 3D plot of uniformity, standard deviation, and the K-S p-value and illustrates the point that simple measures of spread are not adapted to the regularization of a set of parameter values if the ground truth is that there is more than one cluster. The figure also displays a graphical representation of the 10 values in the set for four chosen sets, to show that low-uniformity sets correspond to clustered values (with low K-S p-values) while low standard deviation is associated with single clusters.

One important argument for choosing a metric in a regularized optimization problem might be its low computational cost. Comparison of the running time for uniformity with other metrics shows that the new metric can be computed very efficiently (Figure 24D), several orders of magnitude faster than the non-parametric tests or the clustering algorithm. This low computational cost makes is well adapted to the repetitive computations characteristic of gradient-descent optimizations.

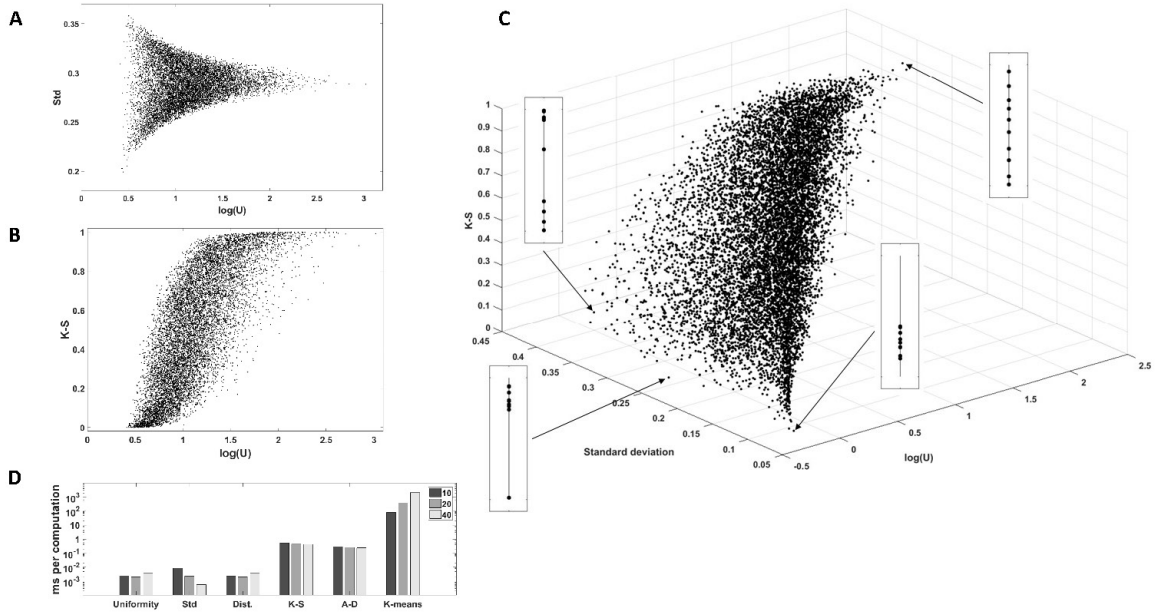


FIGURE 24: Evaluation of uniformity U as a measure of structure, for 10^4 one-dimensional sets of 10 values; (A): comparison with standard deviation. (B): comparison with the p-value of the K-S test (similar results were obtained with the A-D test). (C): 3D-scatterplot of uniformity, standard deviation and K-S p-value. (D): Computation times for the different metrics. log(U): $\log_2$(uniformity); Std: standard deviation; Dist: sum of pairwise distances; K-S: p-value of the Kolmogonov-Smirnov test; A-D: p-value of the Anderson-Darling test; K-means: k-means clustering, number of clusters determined with the elbow method.

In addition, we performed experiments using gradient descent either with the standard deviation, sum of pairwise distances, or uniformity U as an objective function on sets of randomly, uniformly distributed random values. Using the regularization objective as the objective function, without data or model to produce an error function, helps understanding the effect of regularization when signal is low in the data. The traces in Figure 25 reveal the strength and direction of the bias applied on each value in the set in the absence of cost function. Penalizing on the standard deviation results in a homogeneous pull towards the average value (Figure 25A), which does not accomplish the goal of forming clusters. Using the sum of pairwise distances, in turn (Figure 25B), results in grouping of values together, however the clusters themselves are still pulled together. In contrast, the traces in Figure 25C show that using uniformity U, the values form a number of groups, but that these groups are more stable. This is due to the fact that the computation of uniformity U measures local density both below and over the expected value, which means that not only clusters but also voids produce low-uniformity sets. As a result, once values with all clusters have merged, the average of the different clusters remain very similar in number and value to the centroids of the k-means clustering.

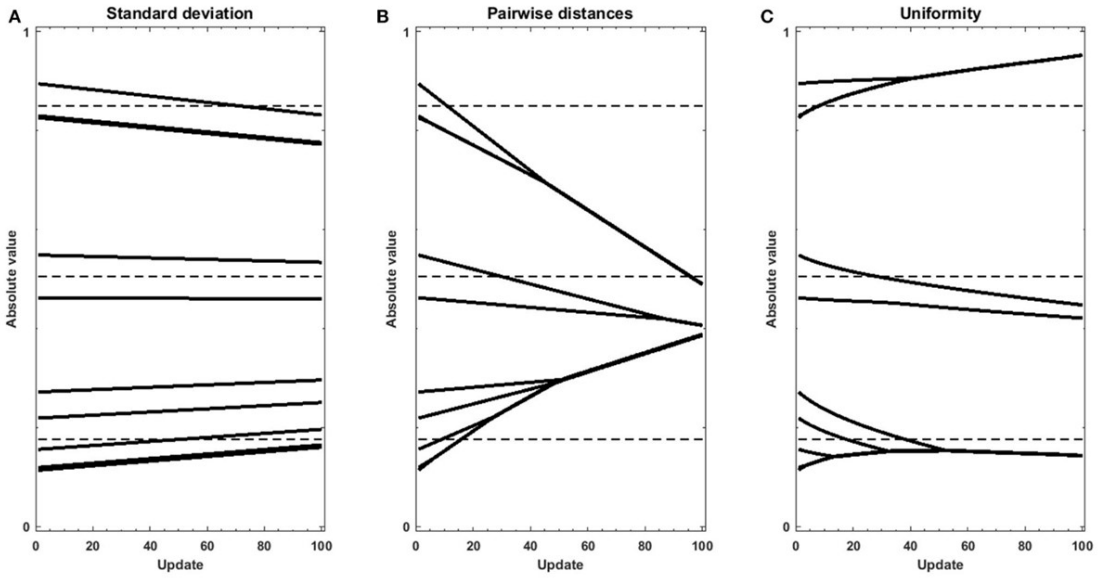


FIGURE 25: Gradient descent trajectories for a set of randomly uniformly distributed values displaying a certain level of structure, using different metrics as objective function: (A) standard deviation, (B) sum of pairwise distances and (C) Uniformity U. The dotted lines show the values of the centroids of clusters as determined by the k-means + elbow method for the original vector.

4.5.2 Toy model

To test the ability of a regularization function using uniformity U to recover context-specific parameters of a network model, we generated an example Bayesian Network which we parametrized for four different imaginary contexts. In our example, the contexts are cell lines, and their regulatory network are identically parametrized two by two. We used the network to generate measurements for two of the nodes while two other nodes were controlled. We added noise to this synthetic data to simulate background noise and normally distributed measurement errors. We used the toolbox FALCON to contextualize the network for the four cell lines, with and without regularization based on the uniformity U of the set of parameter values. We screened 41 values of the hyperparameter λ. The computations took a total of 220 seconds on a standard desktop computer. The results are presented in Figure 26. The regularization paths in Figure 26A show the optimal parameter values over a range of regularization strengths λ. The unregularized model is parametrized differently for each cell line, and the regularization induces a grouping of the parameters values across cell lines. However, this clustering occurs at different values of λ. As regularization strength increases, so does the error of the model (Figure 26B), while the number of unique parameters in the model decreases as they are merged together. We used the Bayesian Information Criterion to balance goodness-of-fit with model size and identified $\lambda = 2^{-4.5}$ as the best model configuration. Figures 26C and 26F show the fitting cost for each cell line for the unregularized model and the regularized

one, respectively. Figures 26D and 26G show the correlation of the simulated values with the measurements, for the unregularized model and the regularized one, respectively, and Figures 26E and 26H show the correlation of the inferred parameter values with the real values for the unregularized model and the regularized one, respectively. Together, these results show that while the new model displays a higher MSE, the inferred parameters are much closer to the ground truth. Regularization transfers a portion of the variance from the parameters back to the data, and so decreases the part of the error on the parameter estimates due to noise. More importantly, the grouping of the samples is easily recovered (Supplementary figure S2), which also carries information: the cell lines are identical two-by-two.

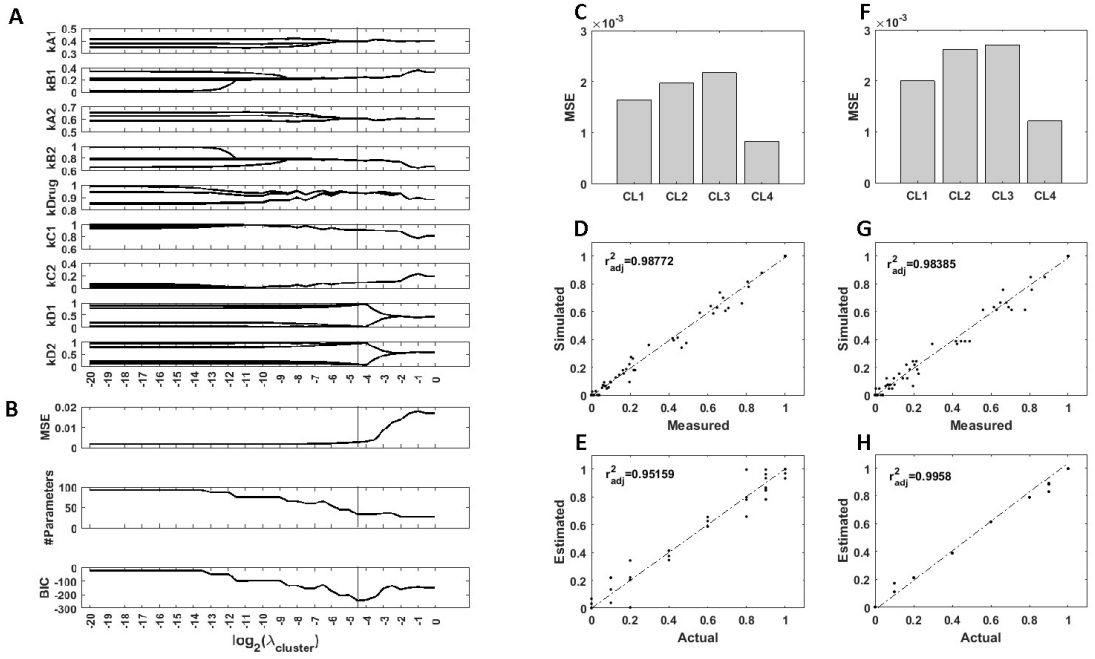


FIGURE 26: Results of the synthetic toy model analysis. (A) Regularization paths for each parameter of the network model. When regularization strength increases, values across the four contexts are encouraged to merge. (B) Mean squared error (MSE), number of different parameters of the model, and Bayesian Information Criterion (BIC) for different regularization strengths. (C, D, E) Unregularized model. (F, G, H) Sparse model. (C, F) MSE for the four contexts with both models. (D,G) Comparison of the simulated node values with the measurements for both models. (E, H) Comparison of the inferred parameter values with the ground truth for both models. r^2_{adj}: adjusted Pearson's correlation coefficient.

4.5.3 Biological dataset

In order to assess the applicability of our new method of regularization to uncover context-specificity in a realistic modeling setting, we reanalyzed the data from (Eduati *et al.*, 2017) using a Dynamic Bayesian Network adapted from the topology of the ODE model. The dataset comprised 8428 datapoints (14 phosphoproteins for 14 cell lines under 43 experimental conditions). We screened 49 values for the hyperparameter λ. The computation time was 1761 hours, or 42 hours when parallelized among 49 computing cores. The results are presented in

Figure 27. Minimum BIC was reached when $\lambda = 0.5$, which corresponds to a model in which 26 of the 79 network parameters can be parametrized identically for all cell lines, and the remaining ones can be organized in 2 to 9 groups. Overall, the most variable parameter across cell lines is the ERK-EGFR negative feedback (Figure 27A and B). Notably, interactions relating to the PI3K/Akt/mTOR axis, to the JUN pathway, and to p38 regulations showed relatively high heterogeneity compared to the crosstalks between them. A number of interactions reveal differential parametrizations for certain cell lines, for example CCK81 in the case of TGFRβ activation by EGFR (Figure 27C), or COLO320HSR in the case of RASK activation by IGF1 (Figure 27D). Figure 27E shows an example of regularization path where no cell line specificity is left in the model with the optimal topology. In addition, many interactions (narrower arrows in Figure 27A) show very low values for all cell lines, suggesting that they do not play an important role in this experiment. The complete set of 79 regularization paths is presented in the Supplementary Materials. The changes in BIC are shown in Figure 27F, displaying a marked minimum around the value 0.5. The goodness-of-fit was similar for all cell lines, with MSE values ranging from 0.018 to 0.035 (Figure 27G). While these results are in line with the ones reported in the original study, it should be noted that in our final model, the role of TAK1 is less prominent, a fact that can be explained by the difference of modeling paradigm. Indeed, while in (Eduati *et al.*, 2017) TAK1's *node responsiveness* parameter τ is extremely low for all cell lines while edges from and to TAK1 are quite variable, our modeling framework considers all nodes equally responsive, and as a consequence low TAK1 activity is represented by low edge parameter values.

Figure 27H shows a heatmap of all model parameters for all cell lines. The dendrograms show the clustering of model parameters and cell lines based on their parameter values. We chose WPGMA to perform hierarchical clustering using the euclidean distance between parameter vectors, with 1000 bootstrap replicates. The support for the nodes in the cell line dendrogram are indicated as percentages. Interestingly, cell lines HT29 and HT115 cluster strongly together, while they are highly dissimilar in their genomic alterations. In general, we noted a poor correlation between the genomic and functional pattern over this set of cell lines, a fact already noted in the original study. Cell lines COLO320HSR and CCK81 are the cell lines functionally most unlike the others. This is also visible in the raw data (see Supplementary Materials), notably in the amplitude of the Akt/PI3k/MEK activations.

Next, we explored the possible associations between the 31 most variable model parameters and sensitivity to 83 chosen drugs. The 25 most statistically significant of these linear associations are presented in the Supplementary Materials. While no parameter-drug pair shows strong significance (most likely due to the high number of hypotheses tested), we noticed a pattern in which some parameters seem to correlate with sensitivity to MEK inhibitors. Figure 27I shows that the parameters relating to PI3K activation by IRS1 and IGF1R are inversely correlated to the log(IC50) of refametinib and trametinib, two known MEK inhibitors.

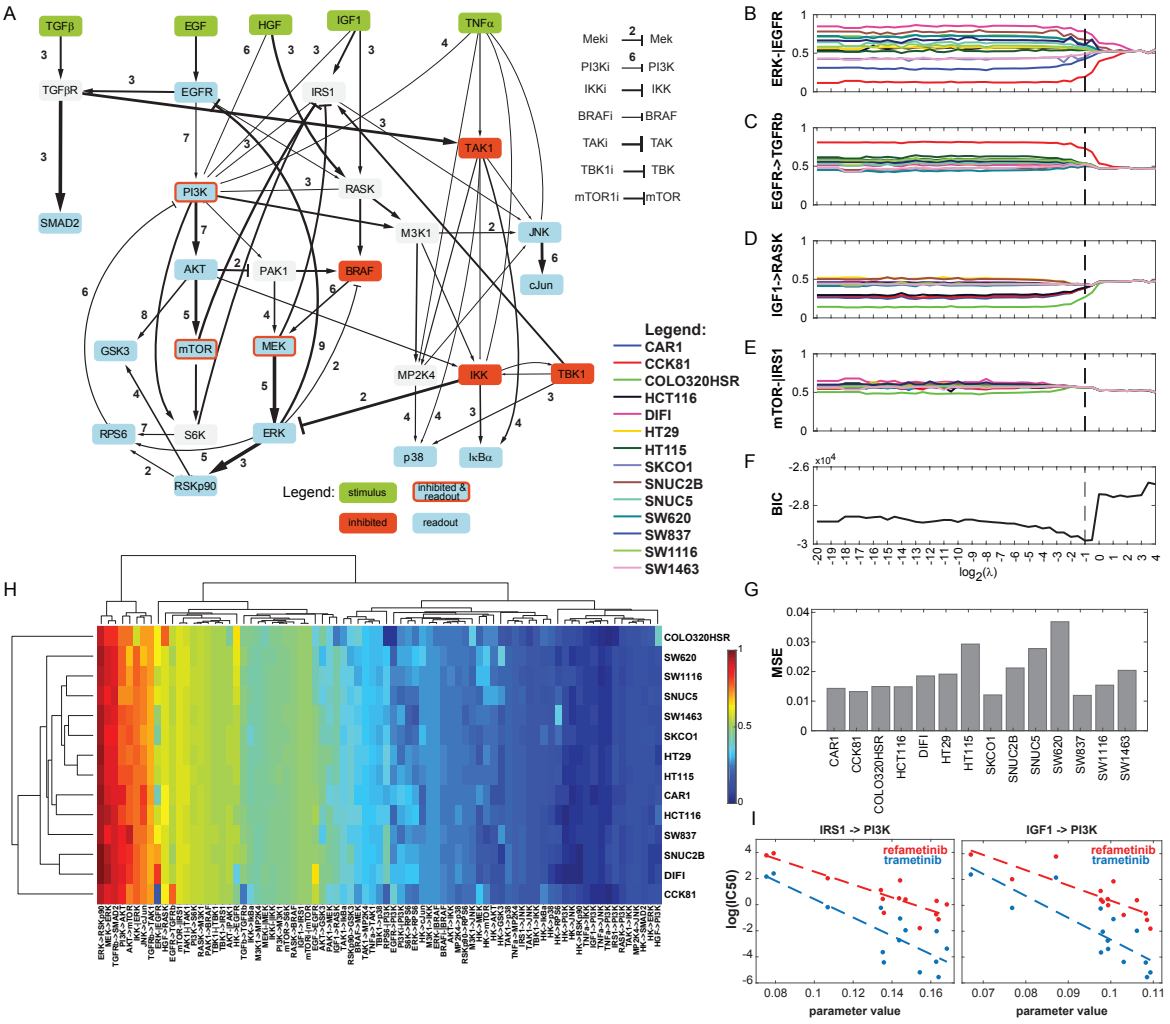


FIGURE 27: Results of the analysis of the biological dataset. (A) Optimized network topology (adapted from (Eduati *et al.*, 2017)). The width of the arrows represents the median parameter value across the 14 cell lines, with wider arrows corresponding to the most active interactions. The number next to the arrows is the number of clusters that the 14 cell lines form for the optimal regularization strength. (B), (C), (D), (E) Regularization paths for four chosen interactions, showing decreasing amounts of cell line-specificity. (F) BIC (Bayesian Information Criterion) path. (G) MSE (Mean Squared Error) for the 14 cell lines for the optimized model. (H) Heatmap of the values of the 79 parameters for the 14 cell lines. Dendrograms were produced with WPGMA using euclidian distance. (I) Correlation between two PI3K-related parameters and sensitivity to two MEK inhibitors. Left: IRS1-PI3K; refametinib: r^2=0.737, p-value=0.133; trametinib: r^2=0.671, p-value=0.176; Right: IGF1-PI3K; refametinib: r^2=0.701, p-value=0.146; trametinib: r^2=0.652, p-value=0.185.

4.6 Discussion

We propose a new measure of the degree to which sets of values are clustered around an unknown number of centers. We use this new metric, called uniformity U, as a penalization in the objective function of models of signal transduction. Previously, regularization applied to the parameters of such models have assumed either that parameter values would be mostly

identical across the different studied contexts (using measures of spread), and looked for departures from this assumption for context-specific parametrizations, or that the parameter values would change in correlation with another, known variable between samples (e.g., smoothly over time). While these assumptions make intuitive sense, they are probably not usable in the case of models of regulatory networks in a large number of cell lines. Indeed, functional relationships between molecules in cells, like enzymatic rates and binding strengths, usually exist in a small number of versions for a specific interaction. Because we do not expect these properties to change along a continuum but in a discrete way, it is natural to assume that model parameters of a regulatory network display the same type of structure. Our method efficiently reduces the complexity of network models. In our toy model example, we decrease the number of parameters from 32 to 11, and correctly recover the fact that two groups of cell lines exist and should be parametrized differentially. In our biological example, we decrease the number of parameters from 1106 to 272, without increasing the error disproportionately.

We show that this method is applicable to biological studies by re-analyzing the dataset from (Eduati *et al.*, 2017). Our analysis indicates that the most variable interactions relate to the PI3k/Akt/ERK axis, in particular the ERK/EGFR negative feedback. Interestingly, it has been shown that negative regulation of the MAPK pathway by ERK is a highly complex mechanism and comprises several components, many of which are affected by cancer mutations (Lake *et al.*, 2016).

By performing hierarchical clustering on model parameters after fitting the data to the best model topology, we recover a grouping of the cell lines that correlates poorly with the genomic alterations. We hypothesize that this means we capture a degree of functional heterogeneity that cannot easily be explained by the cell lines' genomic features. Further indication that our network approach is able to recover phenotypical information that is not obvious in the raw measurements is provided by the pattern of relatively strong correlation between a number of parameters and sensitivity to several MEK inhibitors. This observation fits into the recent developments made in integrating network modeling approaches with advanced statistical modeling, where machine-learning methods have been used to successfully predict sensitivity to single drugs and to drug combinations (El-Chaar *et al.*, 2014; Way *et al.*, 2018). Further work is needed to quantify the merits of our regularization scheme when applied in such context.

Our key contribution is the demonstration that using a simple measure of parameter coefficients density inside the parameter space, it is possible to regularize a large network model and to efficiently group together model parameters for which the difference is not well supported by the data. By *de facto* removing part of the noise in parameter estimates, we are able to decrease model complexity. Furthermore, our regularization scheme is easily adaptable to stronger or weaker priors. Equation 4.8 can be modified as follows:

$$U = \frac{1}{N} \sum_{i=1}^{N} U_i w_i \tag{4.10}$$

with w being the set of relative weights for the different parameters. When $w_i = 1 \forall i$, all parameters are regularized with the same strength. This weighted average allows the specification of additional prior information, namely that the structural assumptions might not be true everywhere, or that our confidence in these assumptions might be stronger in some cases than others.

It is likely that in the near future, single-cell proteomic studies will provide ever-larger datasets, therefore challenging modeling formalisms and requiring them to adapt to larger number of features (Spitzer & Nolan, 2016). While statistical analyses have largely benefited from regularized parametrizations in the form of more predictive models, the current regularization objectives are not well adapted to the study of signaling networks.

A natural extension of this regularization scheme is to consider subsets of M parameters, corresponding to coherent parts of the model, like known signaling pathways. In that case, regularization will act simultaneously on the different constituent parameters of the pathway, and will allow the determination on cell line-specific pathway activity, a high-level information which is usually recovered by ontology-based pathway analysis. However, in such two-step analysis, the confidence for the different parameters is lost. In addition, ontology-based analyses use pathway knowledge from databases, thus suffer from their incompleteness and inaccuracy.

Finally, although we have demonstrated the applicability of this novel method to the study of regulatory networks with logical models, it would be straightforward to extend its use to other modeling environments. For example, systems of ODEs, which are often used to model regulatory networks, might benefit from the addition of a new kind of regularization, using the same methodology presented in this paper. More generally, regularization based on the uniformity of coefficients would in principle be applicable to any type of regression problem and therefore has the potential to be integrated in many analytical frameworks, and be relevant to advanced statistical analysis.

Author contributions

SDL conceived the study, conducted experiments, and wrote the manuscript. PL proposed critical improvements, conducted experiments, and helped editing the manuscript, TS supervised the study, improved the study design, the experimental design and the manuscript.

Funding

This project has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 642295 (MEL-PLEX) and the Luxembourg National Research Fund (FNR) within the projects MelanomSensitivity (BMBF/BM/7643621) and ALgoReCEll (INTER/ANR/15/11191283).

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

The authors would like to acknowledge Dr Thomas Pfau for technical help with the computations and Dr Jun Pang for valuable comments on the manuscript.

Chapter 5

Systemic network analysis identifies XIAP and IκBα as potential drug targets in TRAIL resistant BRAF mutated melanoma

Greta Del Mistro [1,2,*], Philippe Lucarelli [3,*], Ines Müller [1,2], Sébastien De Landtsheer [3], Anna Zinoveva [1,2], Meike Hutt [4], Martin Siegemund [4], Roland E. Kontermann [4,5], Stefan Beissert [1], Thomas Sauter [3], Dagmar Kulms [1,2,#]

[1] Experimental Dermatology, Department of Dermatology, TU-Dresden, Dresden, 01307, Germany

[2] Center of Regenerative Therapies Dresden, TU-Dresden, Dresden, 01307, Germany

[3] Systems Biology, Life Science Research Unit, University of Luxembourg, Belvaux, 4367, Luxembourg

[4] Institute of Cell Biology and Immunology, University of Stuttgart, Stuttgart, 70569, Germany

[5] Stuttgart Research Center Systems Biology, University of Stuttgart, Stuttgart, 70569, Germany

[*] These authors contributed equally to this work

[#] Lead contact

This study has been published in:

Npj Systems Biology and Applications, 2018, 4(1), 39

5.1 Introduction to the paper

TRAIL is a cytokine produced in most human tissues, and particularly immune cells. A natural inducer of apoptosis, it is active *in vitro* against a variety of cancer types, and has been the basis of decades of research towards a potent anticancer drug. Unfortunately, TRAIL-related molecules have shown limited clinical activity. Several mechanisms have been proposed to explain resistance to TRAIL-related apoptosis inducers, pointing to the heterogeneity of this process across cell types.

Melanoma is a particularly aggressive form of cancer which occurrence has been increasing significantly over the last decades, and is predicted to continue to do so in the near future. Patients diagnosed with metastatic melanoma have relatively few available chemotherapeutic options, with most patients not responding to classical chemotherapies, and developing resistance to BRAF inhibitors and PD-1 or CTLA-4-targeted immunotherapies. Therefore more therapeutic options are needed to address the needs of these patients, and TRAIL-based therapy appears as a viable option.

In this paper, the resistance mechanism of the BRAF-mutated A375 cell line to a multimeric form of TRAIL is investigated at the signaling level. To model the signaling network of melanoma cells, we designed a DBN model which we simulate with the FALCON toolbox. Based on multiple phosphoproteomic measurements performed by first author Greta Del Mistro at the TU-Dresden, we contextualize the DBN model for both the parental, naive A375 cell line and a version adapted to low concentrations of IZI1551, a novel form of multimeric TRAIL molecule, in order to recover the specific network parameters for each cell line.

To increase the accuracy of the modeling, we evaluate the difference between the two cell lines parametrization in light of the signal-noise ratio of the dataset. We accomplish this by optimizing both models as one meta-model and including a multi-term regularization function in the objective function to impose a penalty on the complexity of the model, both in the sense of the total size, and the pairwise distance between the cell type-specific parameters. When goodness-of-fit and model complexity are appropriately balanced, the resulting model indicates which interactions are differentially regulated in the two contexts. Our analysis indicates that activation of NFκB and upregulation of apoptosis regulator XIAP occur during adaptation and drive resistance. Our model correctly predicts that inhibition of NFκB or XIAP re-establishes sensitivity to IZI1551 and opens up possible clinical combination therapies for patients with metastatic melanoma.

Data collection was performed by Greta Del Mistro at the TUD. I performed the network analysis, including implementation of the specific regularization scheme, the generation of the related figures and parts of the manuscript.

5.2 Abstract

Metastatic melanoma remains a life-threatening disease because most tumors develop resistance to targeted kinase inhibitors thereby regaining tumorigenic capacity. We show the 2nd generation hexavalent TRAIL receptor-targeted agonist IZI1551 to induce pronounced apoptotic cell death in *mut*BRAF melanoma cells. Aiming to identify molecular changes that may confer IZI1551 resistance we combined Dynamic Bayesian Network modeling with a sophisticated regularization strategy resulting in sparse and context-sensitive networks and show the performance of this strategy in the detection of cell line-specific deregulations of a signaling network. Comparing IZI1551-sensitive to IZI1551-resistant melanoma cells the model accurately and correctly predicted activation of NFκB in concert with upregulation of the anti-apoptotic protein XIAP as the key mediator of IZI1551 resistance. Thus, the incorporation of multiple regularization functions in logical network optimization may provide a promising avenue to assess the effects of drug combinations and to identify responders to selected combination therapies.

Keywords: Apoptosis, Drug Resistance, Dynamic Bayesian Network, Melanoma, Regularization, Systems Biology, TRAIL, XIAP

5.3 Introduction

Dysregulation of two major mitogen-activated pathways (RAS-RAF-MEK-ERK and PI3K-AKT-PTEN) are key drivers of melanoma development and progression (Khattak *et al.*, 2013), with 66% of patients expressing a constitutive active mutant of the MAP (mitogen-activated protein)-kinase BRAF (*mut*BRAF, V600D or V600E) (Paluncic *et al.*, 2016). The initial response rates of patients to first-line therapy with targeted *mut*BRAF inhibitors dabrafenib or vemurafenib is almost 100%, however about 70% of patients acquire resistance to the treatment within one year (Margolin, 2016; Griffin *et al.*, 2017). Accordingly, downstream inhibition of the MAP-kinase MEK with e.g. trametinib is used as a second-line therapy or even initially combined with *mut*BRAF-inhibitors (Luke *et al.*, 2017; Boespflug *et al.*, 2017). Still, the prognosis for patients with metastatic melanoma remains particularly poor and is mostly associated with high tumor relapse rates (Sullivan & Flaherty, 2013; Margolin, 2016; Khattak *et al.*, 2013).

Therefore, alternative treatment options are demanded as first or second line therapy to overcome acquired resistance. In this context, cell death induction by the tumor-selective death ligand TRAIL (Tumor necrosis factor-Related Apoptosis-Inducing Ligand) might serve as an alternative treatment option. Unfortunately, melanoma cells were shown to stay largely resistant against conventional TRAIL treatment (Thayaparasingham *et al.*, 2009; Hörnle *et al.*, 2011). Importantly, conventional trimeric TRAIL and receptor-agonistic antibodies as single

agents failed in clinical trials, due to the limited therapeutic activity in patients (Dimberg *et al.*, 2013). To overcome this therapeutic limitation we have developed novel second generation TRAIL receptor-targeted agonists, with increased bioactivity enhancing the cytotoxic capacity towards cancer cells. These fully human TRAIL-Fc-fusion proteins consist of two single-chain TRAIL molecules fused covalently to the Fc-part of human IgG, forming a potent hexameric TRAIL-receptor agonist (IZI1551). Systemic administration of IZI1551 in mice xenograft models resulted in a potent antitumoral activity with improved pharmacokinetic properties showing no side effects (Siegemund *et al.*, 2018; Hutt *et al.*, 2017).

However, both MAPK signaling as well as TRAIL receptor activation can lead to the activation of the transcription factor NFκB (Dimberg *et al.*, 2013; Richmond & Ueda, 2013; Smith *et al.*, 2014; Yongping *et al.*, 2016), which may impair the therapeutic outcome due to upregulation of survival genes. To take this sensitive balance between pro-and anti-apoptotic signaling into account and to explore novel treatment options, a holistic understanding of the signal transduction network within melanoma cells is a prerequisite.

To assess the relevance of individual interactions within the signal transduction network of melanoma cells, we applied Dynamic Bayesian Network (DBN) modeling, which allows to efficiently contextualize and analyze logical networks. The parameters of DBN models can be estimated using quantitative (quasi) steady-state protein data, thereby for example allowing comparisons between cell types (De Landtsheer *et al.*, 2018), here of therapy-responsive and -resistant melanoma cell lines. Large-scale DBN modeling is feasible with the recently published FALCON toolbox (De Landtsheer *et al.*, 2017), a Matlab framework designed for computational performance, and comprising a wide range of systems-level analyses. Furthermore, additional constrains on the parameter set can be included in the optimization problem in the form of biased estimators. Such regularized objective functions are frequently used to balance goodness-of-fit with existing prior assumptions (Harrell, 1995). Two desired properties of the parameter values of a meta-model encompassing the different cell types can be formulated. Firstly, it is expected that the phenotypic differences between the cell lines are due to a limited number of molecular changes, and that therefore most cellular processes are identically parametrized across both cell types. Secondly, the final model should be as sparse as possible by focusing on the most essential interactions, to increase its predictive power and to facilitate interpretation.

In order to identify the molecular changes between TRAIL-sensitive melanoma cells compared to melanoma cells that have acquired resistance to TRAIL we used a mixed regularization scheme incorporating these two assumptions within the FALCON toolbox to estimate parameter values for the two cell types and discover the most significant changes that may confer therapy resistance.

5.4 Results

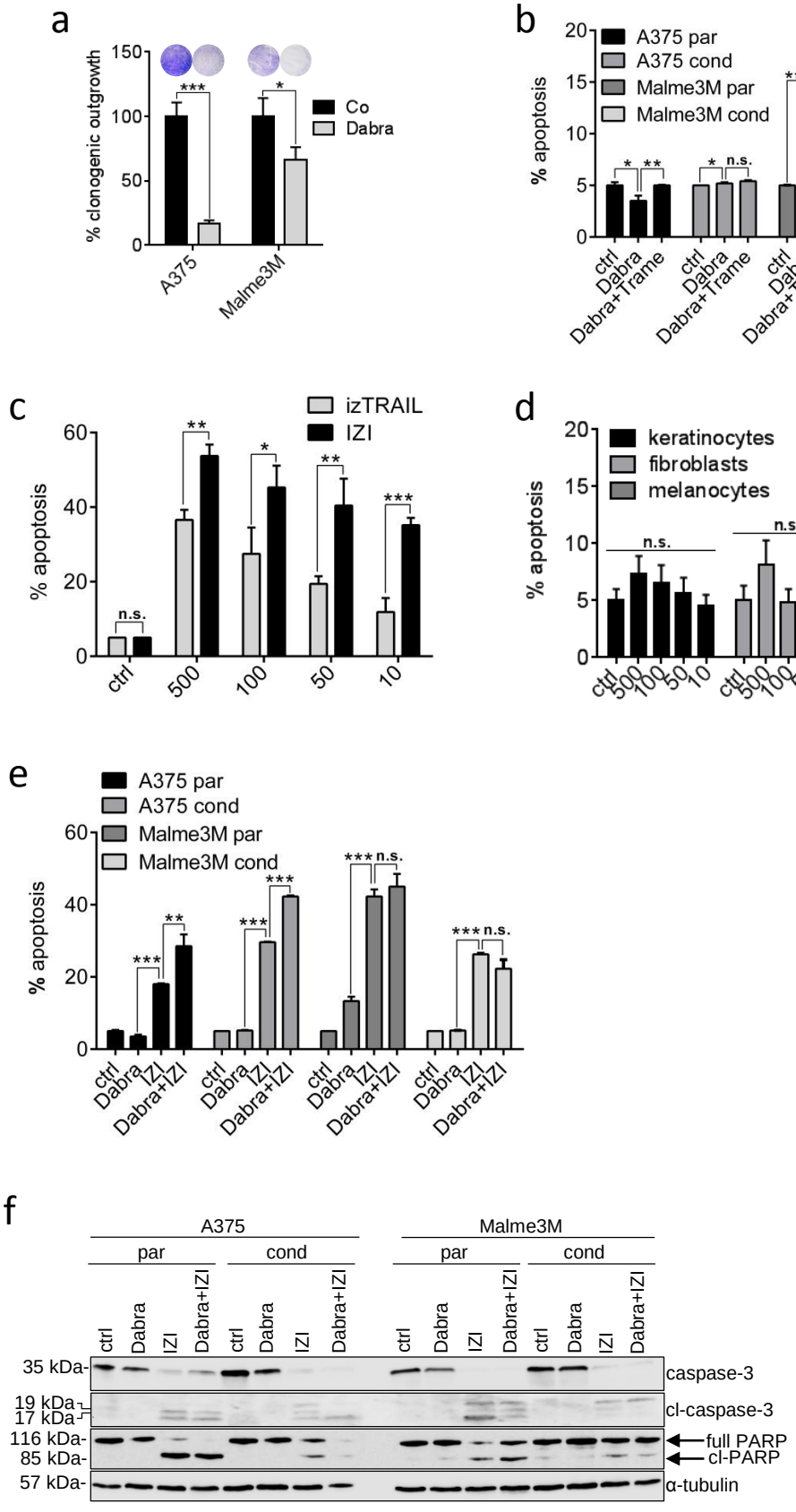
5.4.1 Hexavalent TRAIL receptor agonist IZI1551 is superior in killing *mut*BRAF melanoma cells to conventional TRAIL or specific MAP-kinase inhibitors

Once diagnosed, the first-line therapy of *mut*BRAF melanoma includes administration of specific kinase inhibitors like vemurafenib or dabrafenib. Accordingly, treatment of two *mut*BRAF melanoma cell lines A375 and Malme3M with dabrafenib (Dabra) reduced clonogenic outgrowth, indicating growth inhibition to occur in response to *mut*BRAF inhibition (Figure 28a and Figure S1). However, active cell death induction remained largely absent in response to dabrafenib alone, as well as in combination with the MEK inhibitor trametinib (Trame), as used as second line therapy for patients who have acquired resistance against *mut*BRAF inhibitors (Figure 28b). To mimic dabrafenib resistance we conditioned melanoma cells to a sub-lethal dose of dabrafenib (1 μM) over a period of six months. Neither conditioned, nor non-conditioned, parental cells responded with significant cell death induction to the combination of two downstream MAPK pathway inhibitors (Figure 28b), implying that additional cell death induction might be superior to MEK-inhibition in (re-)sensitizing *mut*BRAF melanoma.

Consequently, conventional trimeric isoleucine-zipper linked TRAIL (izTRAIL) induced moderate apoptotic cell death in A375 melanoma cells while hexavalent scTRAIL-Fc fusion protein (IZI1551) even showed increased cytotoxic activity (Figure 28c). IZI1551-induced cytotoxicity was shown to be largely tumor-selective, as it spared primary keratinocytes, fibroblasts and melanocytes of the skin from apoptotic cell death induction (Figure 28d). Moreover, IZI1551 (IZI) was shown to be significantly more potent in actually killing parental as well as in re-sensitizing dabrafenib-conditioned *mut*BRAF melanoma cells than the specific *mut*BRAF inhibitor dabrafenib (Dabra) (Figure 28e). Accordingly, apoptotic cell death induction through cleavage of the executioner caspase-3 as well as its substrate PARP was exclusively evident in both, parental and conditioned cells upon treatment with IZI alone or in combination with dabrafenib (Figure 28f).

5.4.2 Monitoring IZI1551 susceptibility using mathematical modelling

In order to investigate the potential of IZI1551 as an alternative treatment option for malignant melanoma, we aimed at identifying molecular changes and switches that might occur during acquired TRAIL resistance, and thus conditioned *mut*BRAF A375 and Malme3M melanoma cells to the EC50 IZI1551 dose (5 ng/ml) for six months (Figure 29a and Figure 29b and Table S1). Compared to parental cells (pA375; pMalme3M), conditioned cells (cA375; cMalme3M) stayed largely resistant to treatment with a lethal dose of IZI1551 (50 ng/ml, Figure 29c, compare Figure 29b and Table S1). The overall response to IZI1551 was lower in parental 3D



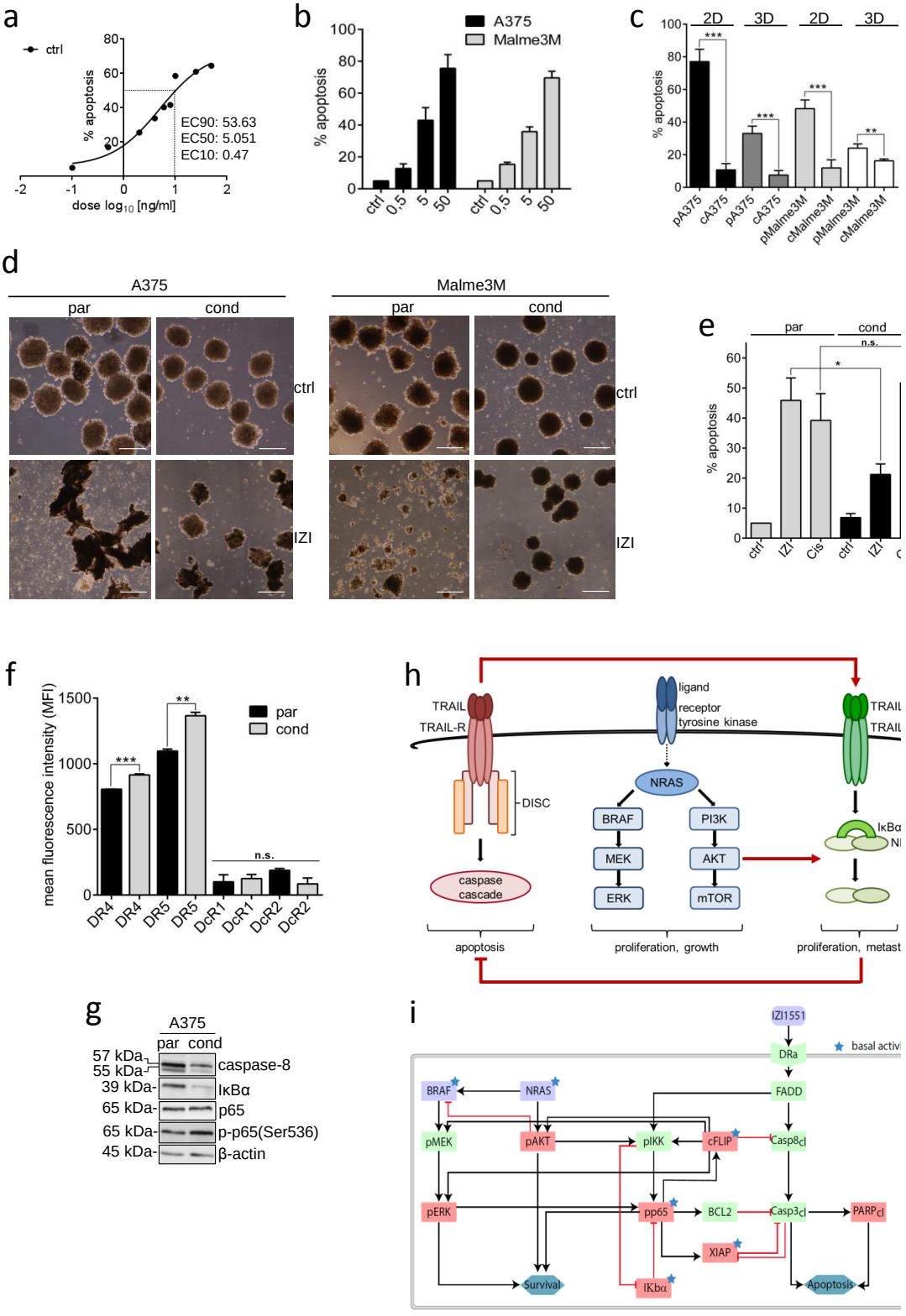
spheroid culture, mimicking the architecture of tumor metastasis in vivo (Vörsmann *et al.*, 2013; Edmondson *et al.*, 2014; Anton *et al.*, 2015), as compared to regular 2D cell culture, and remained largely absent in conditioned 3D spheroids (Figure 29c). Accordingly, only parental

FIGURE 28: IZI1551 is superior in killing melanoma cells than TRAIL or specific MAP kinase inhibitors. (a) Clonogenic outgrowth of *mut*BRAF A375 and Malme3M melanoma cells treated with dabrafenib (Dabra; 10 μM) for 8 days was compared to untreated cells (*p $\leq$ 0.05; ***p $\leq$ 0.001). (b) Parental and dabrafenib-conditioned melanoma cell lines A375 and Malme3M were treated with dabrafenib (Dabra; 10 μM) alone or in combination with trametinib (Trame; 1 μM). After 48 h apoptosis was determined using a Cell Death Detection ELISA (CDDE) (*p $\leq$ 0.05; **p $\leq$ 0.01; ***p $\leq$ 0.001; n.s. = not significant). (c) A375 melanoma cells were treated with increasing doses of izTRAIL or IZI1551 as indicated (ng/ml). After 24 h apoptosis was determined using a CDDE (*p $\leq$ 0.05; **p $\leq$ 0.01; ***p $\leq$ 0.001; n.s. = not significant). (d) The same dose kinetics of IZI1551 as in (C) was applied to primary human keratinocytes, fibroblasts and melanocytes. After 24 h apoptosis was determined using a CDDE (*p $\leq$ 0.05; ***p $\leq$ 0.001; n.s. = not significant). (e) Parental (par) and dabrafenib-conditioned (cond) melanoma cell lines A375 and Malme3M were treated with dabrafenib (Dabra; 10 μM) or IZI1551 (IZI; 50 ng/ml) alone or in combination. After 24 h of IZI1551 and 48 h of dabrafenib treatment apoptosis was determined using a CDDE (**p $\leq$ 0.01; ***p $\leq$ 0.001; n.s. = not significant) and (f) monitored by Western-blot analysis using antibodies against caspase-3 and PARP. α-tubulin served as loading control. For CDDE and clonogenic assay, the mean $\pm$ SD of three independently performed experiments is shown. WBs represent one out of three independently performed experiments.

3D spheroids were shown to be disrupted due to cell death induction 24 h after treatment with IZI1551 (Figure 29d).

Interestingly, IZI1551-resistant cA375 cells had not generally lost the capacity to induce cell death, since they could still respond to cisplatin treatment with apoptosis induction (Figure 29e). Neither downregulation of apoptosis-inducing TRAIL receptors 1 (DR4) and 2 (DR5) nor upregulation of TRAIL-decoy receptors DcR1 or DcR2 was evident to confer IZI1551 resistance in conditioned melanoma cells (Figure 29f). The major dysregulation identified at the molecular level displayed downregulation of the initiator caspase-8 (Figure 29g, Figure S2a). This is due to the conditioning process in which only those cells survive the constant exposure to 5 ng/ml TRAIL agonist which express only low levels of caspase-8. Under these conditions, caspase-8 seems to serve a non-catalytic scaffold function, leading to cytokine production via NFκB activation, instead of cell death (Henry & Martin, 2017; Hartwig *et al.*, 2017). Along this line, also the protein level of the NFκB inhibitor IκBα was shown to be reduced, accounting for constitutive activation of the transcription factor NFκB, being also evident by enhanced phosphorylation of its p65 subunit Figure 29g). It therefore appeared that melanoma cells surviving TRAIL receptor activation selectively reduced the apoptotic signal transduction by downregulating the receptor-associated initiator caspase-8 and at the same time activated NFκB, which is usually associated with upregulation of anti-apoptotic genes (Dutta *et al.*, 2006).

For the mathematical modeling based network analysis we therefore focused on the integration of three signal transduction pathways that may influence melanoma progression and treatment: MAPK signaling - as frequently dysregulated in melanoma, extrinsic death receptor-driven apoptosis, and alternative death receptor-driven anti-apoptotic NFκB activation (Figure 29h). In order to disentangle the complexity of melanoma resistance, we established a DBN model comprising the selected signal transduction pathways as well as their crosstalk



to precisely identify the most sensitive nodes within this signal transduction network that may serve as druggable targets. The network topology was assembled from literature and public databases (Metacore and Ingenuity), and comprised 19 nodes and 29 parameters. We

FIGURE 29: Monitoring IZI1551 susceptibility using mathematical modeling. (a) Dose response curve of 9 different IZI1551 concentrations to determine the EC50 concentration. (b) A375 and Malme3M melanoma cells were treated with increasing IZI1551 doses (0.5; 5; and 50 ng/ml) and apoptosis determined 24 h later in a CDDE. (c) Parental and IZI1551-conditioned A375 and Malme3M cells in 2D cell culture and in 3D spheroid culture, respectively, were treated with IZI1551 (50 ng/ml). After 24 h apoptosis was determined using a CDDE (**p $\leq$ 0.01; ***p $\leq$ 0.001), and (d) monitored by transmission microscopy of 3D spheroids. Scale bar = 250 μm. (e) Parental and IZI1551-conditioned A375 cells were treated with IZI1551 (50 ng/ml) or cisplatin (30 μM). After 24 h apoptosis was determined using a CDDE (*p $\leq$ 0.05; n.s. = not significant). (f) Surface expression level of TRAIL receptors 1 (DR4) and 2 (DR5) and decoy receptors 3 (DcR1) and 4 (DcR2) of parental (par) and conditioned (cond) A375 cells was scored by FACS analysis (**p $\leq$ 0.01; ***p $\leq$ 0.005; n.s. = not significant). (g) The expression level of caspase-8, IκBα, NFκB(p65) and phosphorylated-p65(Ser536) in untreated parental and IZI-conditioned A375 cells was monitored by Western-blot analysis. β-actin served as loading control. One representative Western-blot out of three independently performed experiments is shown (for two more replicates, see Figure S2a). (h) Schematic overview of MAPK-dependent, TRAIL-induced pro-apoptotic and NFκB-driven anti-apoptotic signal transduction pathways. (i) Topology of the Dynamic Bayesian Network (DBN) model of the signal transduction pathways. Black arrows indicate the activation, red arrows the inhibition of target proteins (purple nodes = model inputs, red nodes = measured proteins, green nodes = not measured proteins, blue nodes = functional measurements, asterisks = constitutively active proteins). For CDDE and flow cytometry analysis the mean $\pm$ SD of three independently performed experiments is shown.

calibrated models independently for each cell type (Figure 29i) with quantitative protein expression and activation data of MAPK members AKT and ERK, the pro-apoptotic protein PARP, and anti-apoptotic proteins including IκBα, NFκB (p65), FLIP, and XIAP derived from immunoblotting of unstimulated and stimulated parental versus conditioned A375 cells (Figure S2b and S2c).

We deliberately established the DBN model exclusively on data derived from parental and conditioned A375 cells, intending to utilize Malme3M cells to validate the predictive power of the model retrospectively.

5.4.3 Accurate modeling requires apoptotic proteins in parental but mostly NFκB driven anti-apoptotic proteins in conditioned cells.

Acquired TRAIL resistance during conditioning of cells to IZI1551 caused severe modifications in expression levels of anti-apoptotic proteins XIAP and FLIP, as well as in the activation status of pro-survival proteins NFκB, IκBα, AKT, ERK, and pro-apoptotic protein PARP over time (1, 2, 4, 8, 16, 24, 48 h) (Figure 30a and Figure S2a and the heatmap Figure S2b which includes the values of the untreated and treated samples). We analyzed the differences in normalized protein expression in parental and IZI1551-conditioned A375 cells for each time point. The largest overall differences between the profiles of parental and IZI1551-conditioned A375 cells were observed at the three different time points that might be referred to as: initiation phase (4 h), execution phase (16 h), and adaption phase (48 h) (Figure 30b). To analyze the network modularity within these three phases, we performed systematic in silico

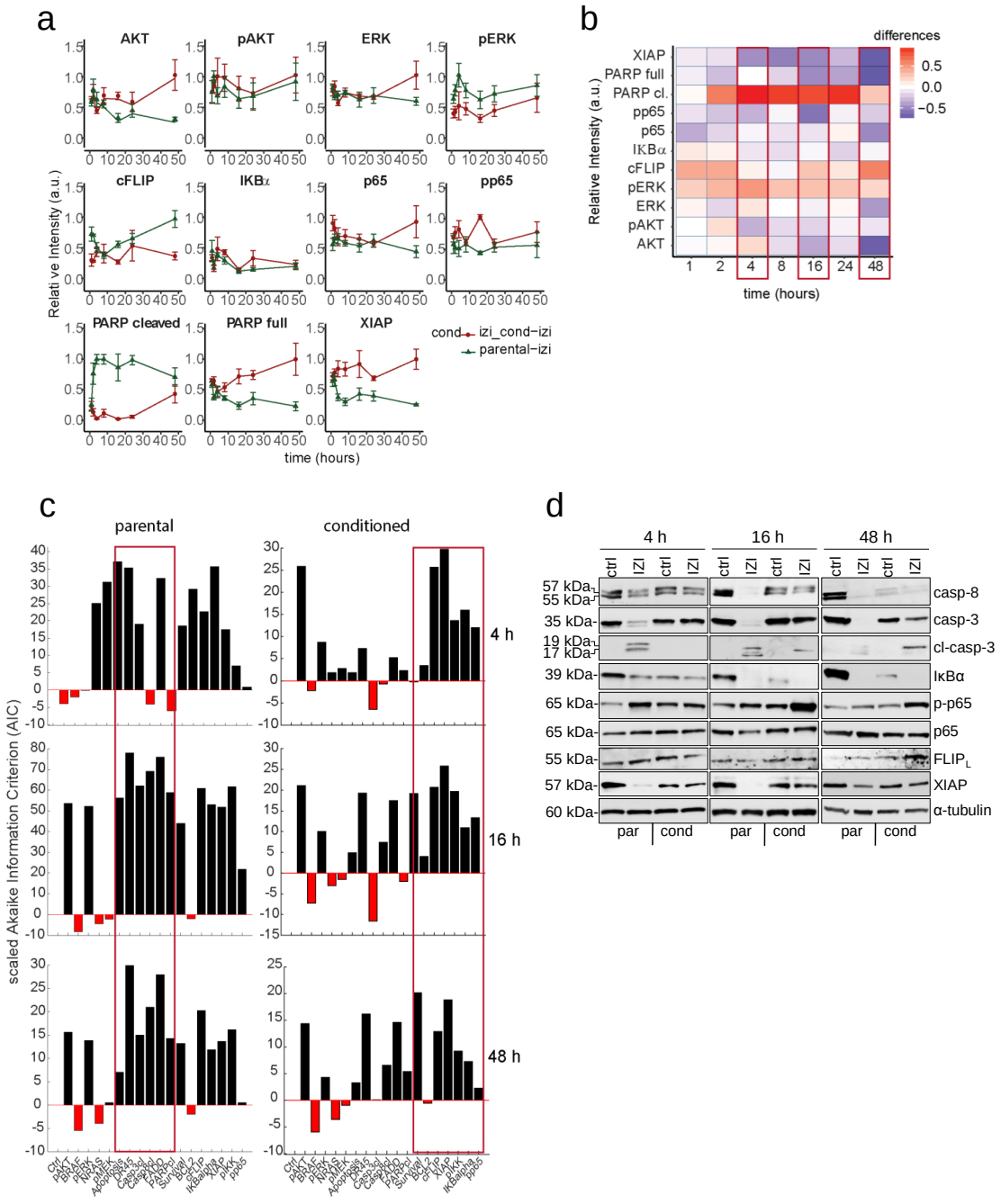
protein knock-out experiments in the parental and IZI1551-conditioned cells at 4, 16 and 48 hours. We used the Akaike Information Criterion (AIC) as selection criteria to verify if the selective removal of each individual node can be compensated by the network. Based on this analysis mostly pro-apoptotic proteins were shown to play an essential role in the execution and adaption phase of parental cells in response to IZI1551 treatment. In contrast, pro-apoptotic proteins only played a minor role in IZI1551-treated conditioned cells, because right from the initiation phase and through the execution and adaptation phases NFκB-dependent anti-apoptotic proteins were shown to be the most indispensable for accurate modeling of the system (Figure 30c).

In accordance with the results from the *in silico* knock out, Western-blot analysis confirmed lack of pro-apoptotic caspase-3 processing and concomitant NFκB activation to be key characteristics of conditioned A375 cells in response to IZI1551 treatment compared to parental cells (Figure 30d). As a consequence IZI1551-induced depletion of anti-apoptotic proteins FLIP and XIAP was fully compensated most likely due to upregulation of these NFκB-dependent genes (Figure 30d) (Thayaparasingham *et al.*, 2009; Hörnle *et al.*, 2011). Immunoblotting in concert with the mathematical model analysis strongly implied the balance at the TRAIL receptor of IZI1551 conditioned cells to switch from pro-apoptotic caspase-dependent signal transduction to NFκB-driven anti-apoptotic signaling, which finally may confer TRAIL resistance.

Considering the overall network sensitivity upon *in silico* knock out of each node, the execution phase (16 h) seems to represent the time-point of maximal vulnerability to systems perturbation between parental and conditioned A375 cells. Based on the analysis and taking into account the fact that the initiation phase (4 h) might not represent the steady-state of the signaling network, while in the adaption phase (48 h), the effects of transcriptional regulation might already be too large and might alter the wiring of the signaling machinery, we selected the 16 hour time-point for further analysis.

5.4.4 Model analysis predicts that dysregulated XIAP and IκBα drive IZI1551 resistance in melanoma.

In order to integrate the experimental data into a coherent picture and to gain a systems-level understanding of the signal transduction network affected by IZI1551 conditioning, we implemented different regularization algorithms in the FALCON toolbox to identify the cell type-specific parameters. We therefore combined two regularization methods. The partial-norm (L1/2) regularization method optimizes identical models for multiple series of experimental conditions in parallel and allows discovering those parts of the network that are active or inactive between cell lines, resulting in pruning of inactive edges within the experiments. The grouped L1 regularization for each interaction focuses on the differences between parental and conditioned cells and tends to reduce the model size by assigning the same parameter value



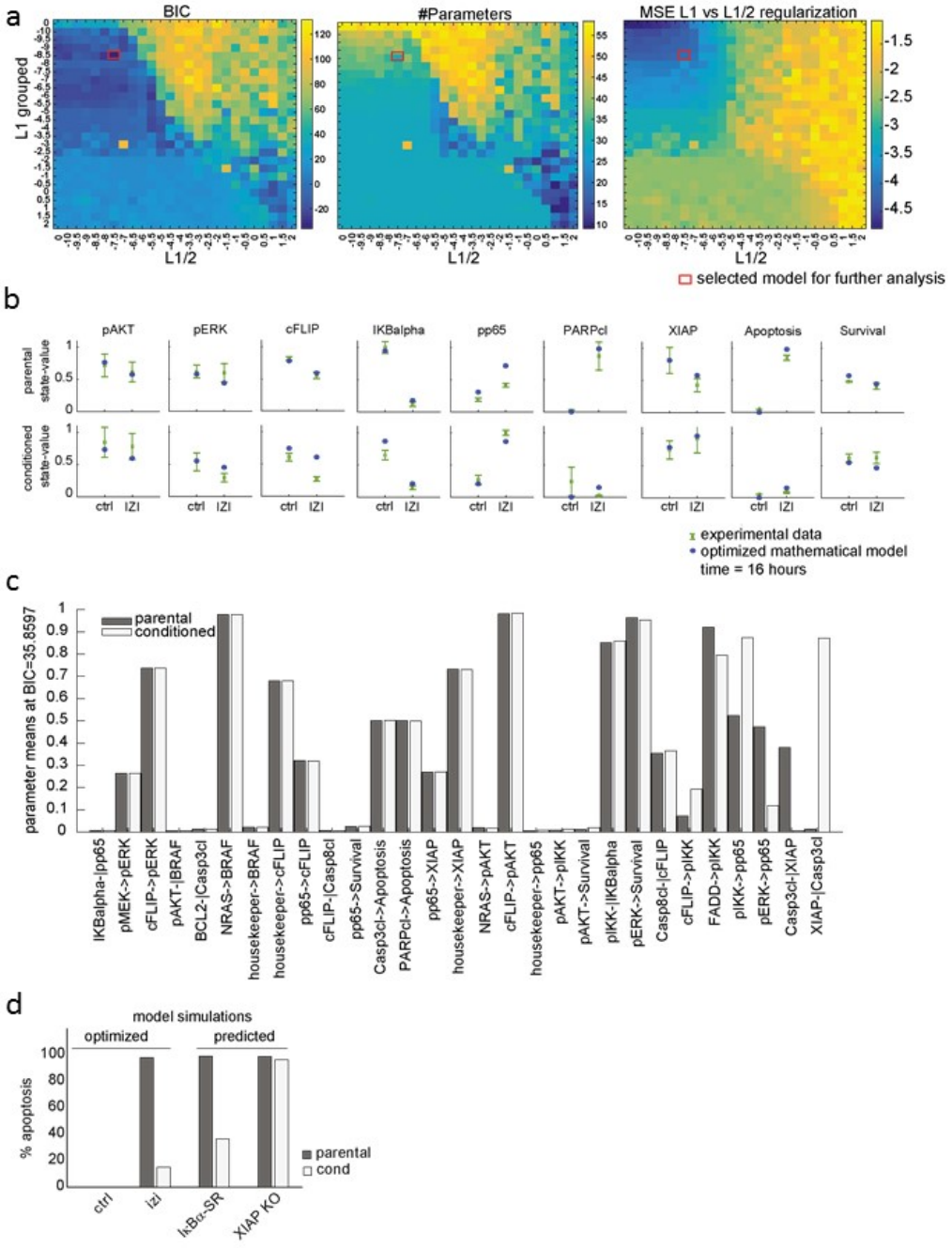
for both cell types for a given interaction. Here, we combined both methods to identify the minimal network structure and uncover cell type-specific differences.

To identify the minimal set of reactions in the network (L1/2) as well as the minimal number of parameters between cell lines (L1 groups), we screened values for the regularization strengths from 10^{-10} to 10^{2} by half-log steps, thus giving 25 different values to test, plus 0. We performed the optimization for each combination of these values, thus optimizing in total 26 L1/2 x 26 L1-grouped regularization strengths = 676 models. Out of 676 different model structures

FIGURE 30: Accurate modeling requires apoptotic proteins in parental but mostly NFB-driven anti-apoptotic proteins in conditioned cells. (a) The expression pattern of AKT, ERK, FLIP, XIAP, IκBα, NFκB(p65) and PARP proteins were analysed by quantitative immunoblotting in whole-cell lysates of parental and IZI1551-conditioned A375 cells stimulated with IZI1551 (50 ng/ml) for 1, 2, 4, 8, 16, 24, and 48 h. The corresponding protein values were normalized between 0 and 1. Error bars represent the SEM with n = 5. (b) Heatmap representing the time-dependent differences between protein expression in parental and IZI1551-conditioned A375 cells. (c) Systematic *in silico* knock-out analysis of each individual protein at 4, 16 and 48 h. The Akaike Information Criterion (AIC) for the reference model is scaled to 0. Parameters playing an important role in the optimization are displayed in black (parameter $\geq$ 0). Parameters having no effect on the model optimization are displayed in red (parameter $]leq$ 0). (d) Parental and IZI1551-conditioned A375 cells were treated with IZI1551 (50 ng/ml). After 4 h, 16 h and 48 h the status of caspase-8, caspase-3, IκBα, p-p65(S536), p65, FLIPL and XIAP was determined by Western-blot analysis. GAPDH served as loading control. One out of five independently performed experiments is shown.

investigated, we identified the optimal network structure based on the Bayesian Information Criterion (BIC) (Figure 31a; BIC: red box).(Burnham & Anderson, 2004) The optimal network (Figure 31a; #Parameters) contained 29 parameters comprising 19 parameters with equal values for both cell types, 4 reactions equal to 0, and 6 cell type-specific reactions, while the initial network contained 58 non-zero parameters (2 cell lines x 29 parameters). The goodness-of-fit of the reduced model was assessed by the mean squared error (Figure 31a; MSE: red box). The total runtime of the experiment was 6.2 h for assessing the 676 model variants, i.e. approximately 13 s per individual model. The reduced mathematical model with only 31 non-zero parameters for both cell lines was shown to be able to describe the experimental data (Figure 31b) to a similar extent as the complete mathematical model which is considering different parameters for both cell types. These results show that our modeling pipeline is able identify major putative differences in parameter values between parental and conditioned cells (Table S2).

The relevance of these differences can be further analyzed by parameter comparison, as displayed by parameter values sorted by increasing difference between both cell types (Figure 31c). This allows estimating if a reaction is cell type-specific and/or essential. The reactions which strongly differ between both cell types (Figure 31c, right columns) are mainly linked to the NFκB activating and apoptosis inducing pathways (Table S2). Reactions from the anti-apoptotic proteins FLIP (FLIP$-|$Casp8cl) and also BCL2 (BCL2$-|$Casp3cl) were close to 0 in the parental and conditioned cells, meaning that the inhibition strength of both proteins would not be enough to inhibit apoptosis. The strongest difference between parameter values can be observed in the reaction XIAP$-|$Casp3cl, absent in parental cells (k = 0.0135) but highly active in conditioned cells (k = 0.8715). Considering these modelling results, one would expect that IZI1551-conditioned cells upregulated the apoptosis inhibitor XIAP to acquire resistance to the treatment. Accordingly, the regularized model predicts IκB super repressor (IκB-SR) to partially, and XIAP knockout to fully re-sensitize conditioned cells (Figure 31d).



5.4.5 DBN modelling correctly predicts melanoma cell re-sensitization to IZI551 by targeting NFκB or XIAP

Given these model predictions, we wanted to verify whether NFκB-driven up-regulation of XIAP plays a major role in conferring TRAIL resistance in IZI1551-conditioned A375 melanoma cells. As predicted, ectopic expression of a non degradable IκBα (S32/36A)-SR mutant, preventing NFκB activation, was able to partially re-sensitize conditioned cells to IZI1551 (Figure 32a), and coincided with increased XIAP-depletion (Figure 32b). An enhanced turnover of XIAP in parental versus conditioned A375 cells was also evident when we monitored loss of

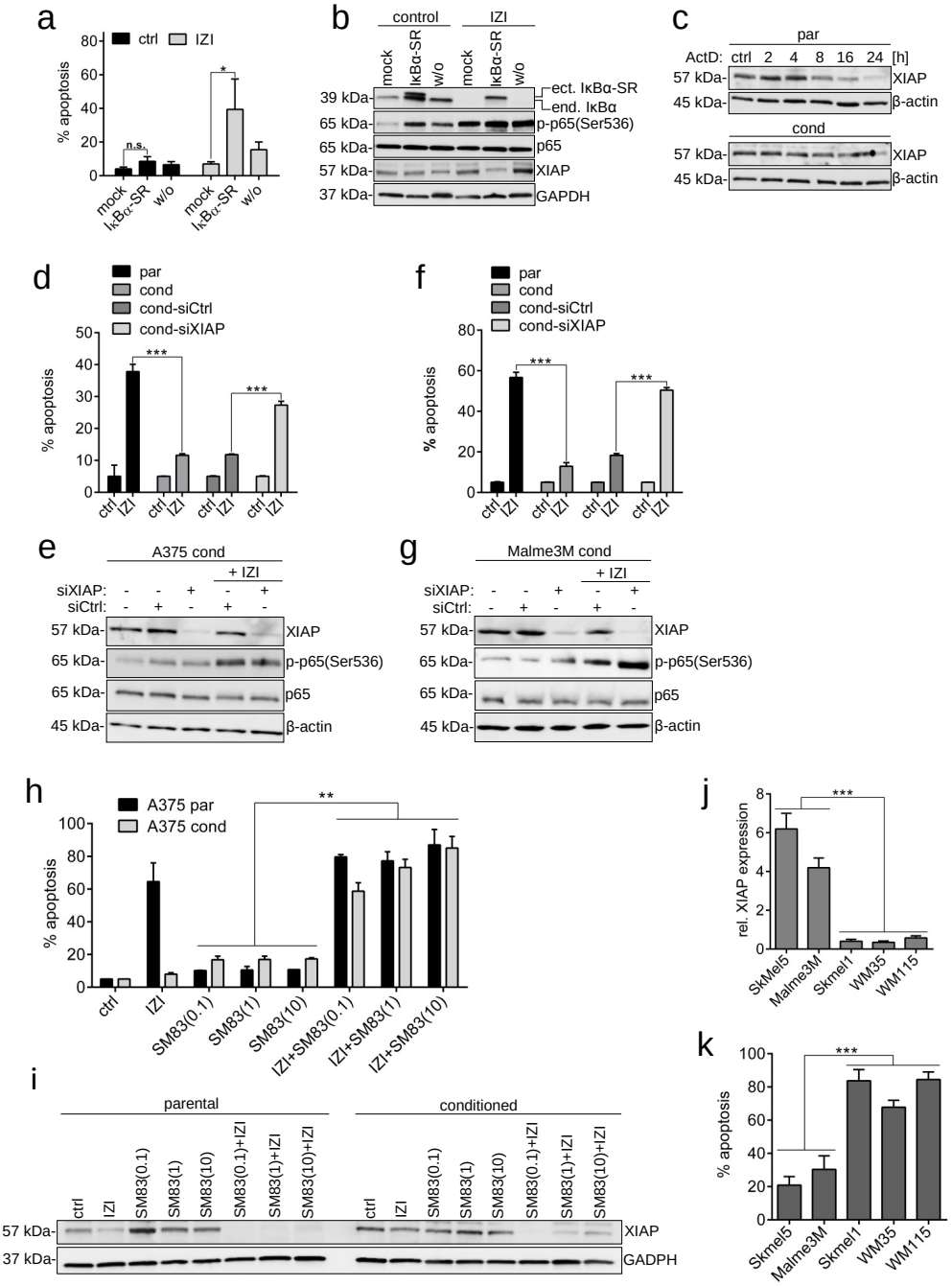
FIGURE 31: Model analysis predicts that dysregulated XIAP and IκBα drive IZI1551 resistance in melanoma. (a) Combined optimization of the grouped L1 and L1/2 regularization algorithms. BIC: The Bayesian Information Criterion (BIC) was used to obtain the best model structure. #Parameters: The number of non-null parameters for each model variant. MSE: Logarithm of the mean squared error indicating the quality of the fit compared to the experimental data. X-axis (left to right): increasing the L1/2 regularization. Y-axis (top to bottom): increasing the L1 grouped regularization. Tiles: blue the smallest (best) and yellow the largest (worst) values for the BIC, Parameters and the MSE L1 vs L1/2 regularization. The red box indicates the model with the best BIC. (b) Comparison of the simulated node activity obtained in the optimal model and the protein quantification for the parental and IZI1551-conditioned A375 cells. Blue dots represent the simulated node activity; green dots the average of 5 measurements with standard error of the mean. (c) Optimal parameter values for both cell lines. The parameters are sorted from the lowest (left) to the highest (right) difference in parameter values between cell lines. (d) Model predictions based on the optimized mathematical model simulating the effect of the IκBα super repressor (IκBα-SR) and XIAP knock-out (XIAP KO) on % apoptosis being induced upon IZI1551 treatment of parental and IZI1551-conditioned A375 cells.

endogenous XIAP upon transcriptional inhibition by Actinomycin D (ActD). While XIAP started to vanish after eight hours of ActD treatment in parental A375 cells, it stayed stable at least for 16 hours in conditioned cells (Figure 32c). Intriguingly, transient knock-down of XIAP using RNA interference was able to almost fully re-sensitize conditioned A375 cells to IZI1551, confirming predictions of the DBN model (Figure 32d and Figure 32e).

Most importantly, and without any previous molecular analysis, siRNA-mediated XIAP knock down antagonized its upregulation by NFκB and consequently fully re-sensitized conditioned Malme3M to IZI1551 (Figure 32f and Figure 32g), confirming that XIAP might be a key player in conferring TRAIL resistance in *mut*BRAF melanoma cells, and that the DBN developed here was able to predict this key player correctly. Accordingly, co-treatment with the SMAC mimetic (SM83) was able to re-sensitize IZI1551-conditioned melanoma cells through depletion of XIAP (compare Figure 5h and 5i).

To investigate whether XIAP expression level might serve as a biomarker predicting responsiveness to TRAIL receptor-activating agonistic molecules in general, we correlated semi-quantitative XIAP protein expression level with IZI1551 responsiveness in five different *mut*BRAF melanoma cell lines. Strikingly, only cell lines expressing very low XIAP levels, Skmel1, WM35 and WM115, induced pronounced apoptosis in response to IZI1551 treatment, whereas Malme3M and Skmel5, expressing elevated XIAP protein level, only moderately underwent apoptotic cell death (compare Figure 32j and Figure 32k).

In summary, we have identified hexavalent TRAIL receptor agonist IZI1551 to be superior in actively inducing cell death in *mut*BRAF melanoma compared to conventional trimeric TRAIL but also compared to specific targeted *mut*BRAF kinase inhibitors as used in the clinic. Above this, we have established a regularized DBN model that was able to predict key players of TRAIL susceptibility correctly and helped to identify XIAP to serve as a potential biomarker for TRAIL treatment responsiveness in *mut*BRAF melanoma.



5.5 Discussion

Defensive mechanisms against cell death render melanoma resistant to current therapeutic outlines with targeted kinase inhibitors. The molecular mechanism leading to intrinsic or acquired resistance against BRAF-inhibitors is still controversial since pro- and anti- apoptotic functions depend on cellular context, target proteins, and cross-talk of different pathways (Sullivan & Flaherty, 2013, 2014; Moriceau *et al.*, 2015). Accordingly, neither treatment of two *mut*BRAF melanoma cell lines with the *mut*BRAF inhibitor dabrafenib alone nor in combination with the MEK inhibitor trametinib yielded significant cell death. In contrast,

FIGURE 32: Depletion of XIAP re-sensitizes melanoma cells to IZI1551. (a) IZI1551-conditioned A375 melanoma cells were transiently transfected with an IκBα super repressor (IB-SR) or the empty vector (mock) and treated with IZI1551 (50 ng/ml). After 16 h apoptosis was determined using a CDDE (*p $\leq$ 0.05; n.s. = not significant) and (b) the status of IκBα, p-p65(Ser536), p65 and XIAP monitored by Western-blot analysis. GAPDH served as loading control. (c) Transcription was inhibited in parental and IZI1551-conditioned A375 cells by addition of Actinomycin D (ActD, 1 μM) for the indicated time points. Protein level of XIAP was monitored by Western-blot analysis. GAPDH served as loading control. (d) XIAP was depleted from parental and IZI1551-conditioned A375 and (f) Malme3M cells using RNAi for 72 h. 16 h after treatment with IZI1551 (50 ng/ml) apoptosis was determined using a CDDE (***p $\leq$ 0.001) and (e and g) the status of XIAP, p-p65(Ser536) and p65 monitored by Western-blot analysis. β-actin served as loading control. (h) Parental and IZI-conditioned A375 melanoma cells were treated with IZ1551 (50 ng/ml) or increasing doses of the smac mimetic SM83 (0.1, 1, 10 μM) alone or in combination. After 16 h apoptosis was determined using a CDDE (**p $\leq$ 0.01) and (i) the status of XIAP monitored by Western-blot analysis. GAPDH served as loading control. (j) For five unstimulated mutBRAF melanoma cell lines the relative expression of XIAP was determined by semi-quantitative Western-blot analysis and (k) the apoptotic response to IZI1551 (50 ng/ml) determined after 16 h using a CDDE (***p $\leq$ 0.001). For CDDE the mean $\pm$ SD of three independently performed experiments is shown. WBs represent one out of three independently performed experiments.

we demonstrated that the TRAIL receptor agonist IZI1551 potently induced cell death in parental *mut*BRAF as well as dabrafenib-conditioned *mut*BRAF melanoma cells lines, while sparing untransformed primary cells of the skin. It therefore appears that active and tumor-selective induction of apoptosis through death receptor activation might be a promising first or second line treatment alternative for *mut*BRAF melanoma, administered either alone or in combination with targeted *mut*BRAF inhibitors.

However, different mechanisms of intrinsic TRAIL resistance have also been observed in cancer cells, especially in melanoma (Dimberg *et al.*, 2013; De Miguel *et al.*, 2016). IZI1551-specific acquired resistance coincided with two major features, namely down-regulation of the initiator caspase-8 which is indispensable for downstream execution of apoptotic processes, and constitutive activation of the anti-apoptotic transcription factor NFκB. Low caspase-8 levels have been reported to form non-functional heterodimers with the FLICE inhibitory protein (FLIP) that are more stable than the functional caspase-8 homodimers and may lead to NFκB activation instead of cell death induction (Hughes *et al.*, 2009; Tummers & Green, 2017). In turn, FLIP is transcriptionally regulated by several transcription factors, including NFκB and its expression has been correlated to drug resistance in a wide range of human malignancies (Yongping *et al.*, 2016; Zang *et al.*, 2014; Huang *et al.*, 2016). Thus, low caspase-8 level together with FLIP may shift the balance at the TRAIL receptors from pro-apoptotic signaling to anti-apoptotic signal transduction via NFκB activation. It is known that constitutive NFκB activation is linked to tumor maintenance and drug resistance (Dutta *et al.*, 2006; Huang *et al.*, 2016; Braeuer *et al.*, 2006; Müller *et al.*, 2014). Studies investigating the role of NFκB in tumor pathogenesis and the mechanisms regulating its activity, revealed that multiple factors are involved in anti-apoptotic responses, and a better understanding of the molecular mechanisms could lead to new targets identification and prognostic biomarkers (Bassères & Baldwin,

2006). Accordingly, the canonical NFκB signaling pathway was included into a DBN modeling approach aiming to identify the key differences within the signal transduction networks of parental IZI1551-sensitive versus conditioned IZI1551-resistant *mut*BRAF melanoma cell lines and the molecular mechanism leading to acquired TRAIL resistance.

DBNs as well as the related probabilistic Boolean networks (PBNs) are specifically suited to quantitatively model large-scale regulatory and signaling networks based on steady state expression or activity data (De Landtsheer *et al.*, 2017; Trairatphisan *et al.*, 2014). Based on a minimal parametrization (1 parameter per interaction) and a relatively simple algebraic formalism, they obtain superior speed over more detailed kinetic models (e.g. ODE-based) while still preserving a good predictive power (De Landtsheer *et al.*, 2017; Lommel *et al.*, 2016). The continuous variables of the DBNs/PBNs thereby allow quantitative modeling and predictions in contrast to the classical Boolean approaches with only qualitative read-out. In this study we combined DBN modeling with a sophisticated regularization strategy aiming for sparse and context-sensitive networks and show the performance of this strategy in the detection of cell line specific deregulations of a signaling network. For ODE based models, the L1 regularization can be used to demonstrate the connections between the deregulation of signal transduction networks and the pathophysiology in cancers (Merkle *et al.*, 2016; Lucarelli *et al.*, 2018). In contrast to their single regularization, our method includes both cell type comparisons and network pruning as part of the overall optimization problem in the form of regularization functions, therefore providing a more stable solution than methods based on independent optimization and unsupervised clustering or multi-step model selection methods. Combining both selection methods resulted in a total of 676 model variants which could efficiently be scanned with the very fast DBN implementation in the FALCON toolbox. This resulted in the simultaneous network pruning, contextualization and parameter fitting which are tasks which usually are only performed sequentially in other modeling frameworks.

Analysis of the regularized model revealed that a subset of reactions, mainly linked to NFκB and anti-apoptotic signaling, were strongly upregulated in conditioned IZI1551-resistant cells, whereas the essential nodes in the parental cell lines were identified to be mainly pro-apoptotic proteins. The model accurately predicted IκBSR to partially and XIAP knockout to fully re-sensitize conditioned cells. The continuous regularization paths within the innovative strategy make sure that the top performing models located in the same region in the BIC landscape will have a similar parametrization and thus yield similar predictions.

Following these predictions, we confirmed that NFκB inhibition by ectopic expression of IκBα-SR mutant partially reduced cellular XIAP levels in conditioned melanoma cells coinciding with partial re-sensitization to IZI1551. More importantly, direct depletion of anti-apoptotic XIAP fully rescued the TRAIL-resistant phenotype, not only in the A375 cell line used for model parameterization, but also in another *mut*BRAF cell line, Malme3M. XIAP, an NFκB-dependent member of the inhibitor of apoptosis (IAP) family, inhibits apoptotic cell death through binding to the executioner caspase-3, -7 and -9, and has been shown to be upregulated

in many human tumors (Kashkar *et al.*, 2006; Vogler *et al.*, 2008; Flanagan *et al.*, 2015). Conversely, XIAP was shown to enhance NFκB activation constituting a positive feedback loop to prevent apoptosis (Hörnle *et al.*, 2011; Chawla-Sarkar *et al.*, 2004; Evans *et al.*, 2016). Accordingly, co-application of XIAP-inhibiting SMAC mimetics has successfully been used to sensitize different TRAIL-resistant tumor cells (Lecis *et al.*, 2010; Raulf *et al.*, 2014) to apoptotic cell death. To that effect, we showed co-stimulation with the SMAC mimetic SM83 to fully re-sensitize melanoma cells to IZI1551 that had acquired secondary resistance to the TRAIL-agonist via XIAP depletion.

Conclusively it turned out that the DBN model combined with the regularization strategy accurately predicted XIAP to be the key player in conferring TRAIL resistance. This became even more evident when we could correlate elevated XIAP expression level in melanoma cells to intrinsic TRAIL resistance, indicating that XIAP may serve as a biomarker for TRAIL responsiveness of *mut*BRAF melanoma. Here, we provide evidence that alterations in the abundance of the NFκB and XIAP proteins change the sensitivity of resistant melanoma cells to IZI1551 treatment. Our studies show that the resistance mechanism is conserved between A375 and Malme3M cells.

Taken together, these results indicate that essential network and cell type-specific reactions can be identified using protein measurements and regularized optimization of a DBN model. The underlying mechanism involved upregulation of NFκB, and predictions identified XIAP as the key player of TRAIL (IZI1551) resistance. Based on the fact that numerous SMAC mimetics are already used in clinical trials for numerous cancers, including leukemia, lymphoma, and solid tumors as single agents or combination therapies (Fulda, 2015), one could envisage SM83-IZI1551 combinations for the treatment of kinase-inhibitor resistant melanoma patients in the future.

Importantly, incorporation of multiple regularization functions in optimization problems, including logical networks modeling, may provide a promising avenue for future studies to assess the effects of drug combinations and eventually to identify responders to selected combination therapies in a personalized approach.

5.6 Methods

Unless stated otherwise, results of Cell Death Detection ELISA and flow cytometry analysis are presented as mean ± SD of 3 independently performed experiments. Western-blot analyses represent one out of 3 independently performed experiments. Statistical analysis of biochemical data was performed using Student's t-test.

5.6.1 Cells and Reagents

Human melanoma cell lines (A375, Malme3M, WM1366, WM1346, Skmel5, Skmel1, WM35, WM115), were maintained in RPMI 1640 medium (Gibco, #61870-010) with 10% FCS (Gibco, #10270-106) in a humified atmosphere of 5% CO2 at 37C. A375, Malme3M, WM1366, and WM1346 were conditioned to 5 ng/ml IZI1551, A375 and Malme3M to 1 μM dabrafenib over a period six months, adding fresh compound every other day. Primary cells were purchased from Cell Systems and used at passage 4. Keratinocytes (#FC-0007) were maintained in DermalifeK Complete Medium (Cell Systems, #LN-0027), fibroblasts (#FC-0001) in DMEM (Gibco, #41965-039) and melanocytes (#FC-0030) in Melanocyte Growth Medium (M2, Promocell, #C-24300). For cell death induction 50 ng/ml IZI1551 (University of Stuttgart), 30 μM Cisplatin (TEVA-Deutschland, #2615.03.01), 1 μM Actinomycin-D (Sigma, #A1410), 0.1-1-10 μM SM83 (Baliopharm), 10 μM dabrafenib, or 1 μM trametinib (both Selleckchem, #S2807 and #S2673) was added to cells.

5.6.2 Plasmids, Cloning and siRNA transfection

For transient expression of IκBα-SR-S32/36A, 6x10^6 A375 cells were electroporated with 20 μg of the plasmid pBK-CMV-IκBα-SR or the empty pBK-CMV vector and investigated 24 h later.

Gene silencing was facilitated by transfecting 5x10^4 cells with 40 pmol siRNA for XIAP- 5'-CGAGCAGGGUUUCUUUAUATT-3' (Ambion, #AM51331), or lacZ-5'-GCGGCUGCCGGAAUUUACCTT-3' (MWG Eurofins) using Lipofectamine 2000 (Thermo Scientific, #11668019), 72 h prior to stimulation.

5.6.3 3D melanoma spheroids

Melanoma spheroids were generated using the hanging drop method (Vörsmann *et al.*, 2013). Briefly, 5x10^4 GFP-expressing melanoma cells were resuspended in 5 ml of medium containing 20% methyl cellulose (Sigma, #M0512). 40 drops of 25 μl containing 250 cells were spotted on the lid of a 10 cm cell culture dish and incubated for 14 days at 37C, 5% CO2. For *in vitro* stimulation 160 melanoma spheroids were collected in a 2 cm culture dish previously coated with 1% agarose.

5.6.4 Flow cytometry

5x10^5 cells were blocked in PBS/2% BSA for 30 min, and incubated with the primary antibodies against TRAIL receptors R1, R2, R3, R4 (huTRAILR1-M271, huTRAILR2- M413. huTRAILR3-M430, huTRAILR4-M444, Amgen) at 2.5 μg/ml in PBS/2% BSA, for 1 h on ice.

After washing twice with PBS/1% BSA, 2 μg/ml of the secondary goat anti-mouse-488 antibody (Thermo Scientific #A-11001, RRID: AB-2534069) in PBS/2% BSA were added for 30 min at 4C. Subsequently, cells were washed twice with PBS/2% BSA and subjected to FACS analysis (LSR II, Becton Dickinson). Excitation wavelength used was 488 nm, the emitted green fluorescence (l_{max} 520 nm) was detected using (FL1) band-pass filter.

5.6.5 Determination of cell death and clonogenic outgrowth

Apoptosis was determined in a Cell Death Detection ELISA (CDDE, Roche, #11920685001) according to the manufacturers protocol. The enrichment of mono480 and oligonucleosomes released into the cytosol is calculated: absorbance of samples/absorbance of control cells at 450 nm (Tecan M200). An enrichment factor of 2 corresponds to 10% apoptosis as determined by AnnexinV-FITC/PI FACS analysis (FACSAria III, Becton Dickinson). For clonogenic assay, $2x10^4$ Malme3M or $8.5x10^2$ A375 cells were seeded into 6-well plates for 8 days or until control cells had reached confluency. Subsequently, cells were stained with crystal violet (0.1 w/v in 20 % Methanol) for 15 min at RT. Cells were washed and crystal violet dissolved from cells with 0.1 M KH_2PO_4/EtOH for 5 min at RT and color intensity of supernatants measured at 595 nm (Tecan M200).

5.6.6 Western-blot analysis

Cells were lysed in lysis buffer (50 mM HEPES, pH 7.5; 150 mM NaCl; 10 % glycerol; 1% Triton-X-100; 1.5 mM $MgCl_2$; 1 mM EGTA; 100 mM NaF; 10 mM pyrophosphate, 0.01 % NaN_3, phosSTOP and Complete). After centrifugation, supernatants were collected and the protein content determined by DC Protein assay kit (BioRad). 60-80 μg of protein extracts were subjected to 4-15% gradient SDS-PAGE (BioRad), blotted onto nitrocellulose membranes and incubated with antibodies directed against PARP, XIAP (BD-Biosciences; #551025, RRID:AB-394009; #610717, RRID:AB-398040), caspase-3, IκBα, NFκB-p65, p-p65(S536), AKT, p-AKT(S473), ERK1/2, p-ERK1/2(T202/Y204) (Cell Signaling; #9665, RRID:AB-2069872; #4814, RRID:AB-390781; #8242, RRID:AB-10859369; #3033, RRID:AB-331284; #2920, RRID:AB-1147620; #4060, RRID:AB-2315049; #9102, RRID:AB-330744; #4376, RRID:AB-331772), FLIP, (Sigma #PRS2437, RRID:AB-259702), and caspase-8 (Adipogen #AG-20B-0057, RRID:AB-2490271), respectively. Equal loading was monitored by re-probing membranes with antibodies against GADPH (Cell Signaling #2118, RRID:AB-561053), α-tubulin (Thermo Scientific #MS-581-P1, RRID:AB-144075), or -actin (Cell Signaling #4970, RRID:AB-2223172). HRP506-conjugated secondary antibodies were purchased from GE-Healthcare (Anti-mouse-HRP, RRID:AB-772210; Anti-rabbit-HRP, RRID:AB-772206). Bands were visualized by applying chemiluminescense SuperSignal detection systems (Thermo Scientific, #34087 and #34076). Protein expression was determined by calculating the ratio

between the protein intensities and either α-tubulin, -actin, or GADPH as housekeeping proteins using ImageQuant 5.2 software (GE Healthcare). Blots derive from the same experiment and have been processed in parallel.

5.6.7 Mathematical modeling

Quantitative protein expression values as determined by Western-blot analysis were then normalized across all cell lines, experimental conditions, and time-points, to the [0-1] interval, independently for each protein, and the average and standard error of five replicates was calculated. Bayesian modelling was performed using these averaged normalized relative expression values as input.

We generated a Dynamic Bayesian Network, a type of probabilistic logical network model of the main pathways hypothesized to play a role in apoptosis resistance from literature. In this type of network model, nodes representing the relative activity of signaling molecules are linked by simple logical functions. These functions can perform the basic AND, OR and NOT operations, and be parametrized with proportionality constants. Contextualized network models include parameter values (k) for each edge, representing the relative strength of this edge. In contrast with strictly Boolean models, which are qualitative, probabilistic logical models are thus able to provide quantitative estimates. Their formulation however is not as mathematically complex as Ordinary Differential Equations, which are classically used for pharmacological models, therefore their computational cost remains low. We used the Matlab toolbox FALCON (De Landtsheer *et al.*, 2017) to contextualize this network with the protein measurements. Briefly, FALCON uses gradient descent to optimize the set of parameter values minimizing the mean of squared error (MSE) between the simulated node intensities of a Dynamic Bayesian Network (Raulf *et al.*, 2014; Fulda, 2015) and the corresponding measured normalized protein expressions. After fitting this network model on steady-state protein measurements, quantitative information can be retrieved, informing on both the relative activity of the signaling molecules in the different experimental conditions, and the strength of interactions between molecules.

To select for context-specific interactions, we optimized the network for both parental and IZI-conditioned cell lines in a single optimization, and included in the objective function two regularization terms to materialize cross-context modeling assumptions. Firstly, we used a fractional norm on parameter values in an effort to prune the network from interactions not strictly necessary to fit the dataset in neither of the cell lines. The use of a fractional norm is dictated by the probabilistic nature of the modeling framework: for each node, the sum of incoming edge strengths must be equal to exactly 1, which renders the L1-norm ineffective to induce sparsity. Secondly, we used a group-level L_1-norm across cellular contexts (Yuan & Lin, 2006), using a strategy similar to Merkle *et al.* (2016) (Merkle *et al.*, 2016). The optimal parameter set K of P parameters $\{k_1, k, ..., k_P\}$ was recovered using the objective function:

$$\underset{K\in[0,1]^P}{\arg\min}\Big(\frac{1}{N}\sum_{i=1}^{N}(X_i-\hat{X}_i)^2+\lambda_1\sum_{k=1}^{P}\sqrt{k}+\lambda_2\sum_{g=1}^{P}\sum_{j=1}^{J^g}|k_j^g-\bar{k^g}|\Big) \tag{5.1}$$

with N the number of individual data points in dataset X and $\hat{X}$ being the set of corresponding values in the simulated network model. The first term is the mean squared error (MSE), the second term is the L_q semi-norm with $q = 0.5$, while the third term is the grouped L_1-norm, with J_g the number of members in group g, and $\bar{k^g}$ the mean value for parameter k in group g. The scalars λ_1 and λ_2 are tuning hyper-parameters controlling the regularization strengths for each of the regularization objectives.

During the systematic knockout experiments, we evaluated models using the Akaike Information Criteria (AIC), which balances goodness-of-fit with model size. AIC is calculated as: $Nlog(MSE) + 2P$. We recovered the MSE and optimal parameter sets for 676 combinations of λ_1 and λ_2 values (screening each one from 2^{-10} to 2^2 by half-log steps), and computed for each of these the Bayesian Information Criterion (BIC) as $Nlog(MSE) + log(N)P$, and the topology of the optimal network, using a minimal threshold of 0.01 for keeping edges and their corresponding parameters in the model.

5.6.8 Data availability

The dataset generated and analyzed during the current study and the model topology including the scripts are available in the GitHub repository, https://github.com/sysbiolux/FALCON/tree/master/FALCON/ExampleDatasets/DelMistro2018. All other relevant data are available from the authors.

Acknowledgments

We thank TUD, CRTD, FACS and imaging facilities for support, advice, and technical assistance.

Competing Interests

R.E.K. is a consultant to Sunrock Biopharma (Santiago de Compostela) and Oncomatryx (Bilbao). M.H., M.S., and R.E.K. are named inventors on patent applications covering the scTRAIL technology. Other than that, the authors declare no competing interests.

Author contributions

Conceptualization, G.D.M, P.L., and D.K.; Methodology, G.D.M., P.L., **S.D.L.**, M.H., M.S.,T.S., and D.K.; Software P.L., **S.D.L.**, and T.S.; Validation, G.D.M., P.L., and D.K.; Formal Analysis, G.D.M, P.L, and D.K.; Investigation, G.D.M., A.Z., I.M., M.H., and M.S.; Writing Original Draft, G.D.M, P.L. and D.K.; Writing Review & Editing, G.D.M, P.L., **S.D.L.**, T.S., R.E.K., S.B., and D.K.; Visualization: G.D.M, P.L., **S.D.L.** and D.K.; Supervision and Project Administration, D.K.; Funding Acquisition, D.K. and T.S. G.D.M. and P.L. contributed equally to this work and share the first authorship.

Funding

This project has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 642295 (MEL-PLEX), the Federal Ministry of Education and Research (BMBF: FKZ 031A423A, Melanoma Sensitivity), and the Luxembourg National Research Fund (FNR) within the project Melanoma Sensitivity (BMBF/BM/7643621).

Chapter 6

General Discussion

6.1 Modeling signaling pathways with DBNs

Cancer is the consequence of the successful amplification of genetically damaged cells, which acquire characteristics enabling them to proliferate, evade innate cell death mechanisms and the immune system, induce angiogenesis, and become invasive and metastatic. Understanding the interplay of different driver mutations, the epigenetic factors participating in the evolution of tumors, the specific strategies cancer cells employ to reprogram energy metabolism, and the crucial role played by the tumor micro-environment, brings forward new possibilities for creating therapeutic strategies for patients with cancer (Hanahan & Weinberg, 2011). However, the size and complexity of the regulatory networks of eukaryotic cells prevents, in most cases, to reach the level of mechanical understanding needed to fully identify the best therapeutic option for each tumor.

Signaling pathways are networks of tightly regulated signal-processing chemical cascades, which are responsible for the coordination of the cells' activities and in general the adequacy of a cell behavior and response to the environmental conditions. Many signaling pathways consist of surface receptors, phosphorylation cascades, and their effect is often the modification of the expression levels of different genes, via the action of transcription factors. However, most pathways cross-talk with each other and interact in non-linear ways, which makes the effect of a specific modification difficult to predict without the full context information, such as the state of activation of the other signaling pathways. Many cancers are characterized by a limited number of driver mutations, i.e. mutations in signal-processing proteins which occur independently in different patients. Two well-known examples are the constitutive activation of the EGFR receptor in non-small-cell lung cancer (Lynch *et al.*, 2004) and of the BRAF protein in melanoma (Shtivelman *et al.*, 2014). The high incidence of these mutations is an indication that these proteins play a key role in carcinogenesis and have therefore been investigated as targets for the design of anti-cancer therapies. This effort has led to the discovery and clinical use of targeted therapies (for example monoclonal antibodies directed against the

mutated form of the protein, or small-molecule kinase inhibitors specifically active against the mutated protein), which brought many improvements to the care of cancer patients over the past decades. However, while the response rate of patients to targeted agents is in general very high, relapse is still frequent, resulting in low survival rates for many cancer types as patients develop resistance to these specific therapies. Many molecular processes have been proposed to explain the acquisition of resistance by tumor cells, however it is likely that the spectrum of possible mechanisms is large and that the care of individual patients requires the identification of a specific drug, or drug combination, to be maximally efficient. Thus, there is an urgent need for tools and methods able to identify these secondary targets, to stratify patients into appropriate functional categories, and to predict the appropriate treatment for each patient.

Systems biology has emerged as one way to gain insight into the specific mechanisms of the diverse cancer subtypes. Network models, in particular, can be powerful in identifying new targets and prioritize mechanisms which play roles in the evolution of the disease. Computational methods have enabled the analysis of very large, multi-dimensional datasets, and many databases have been set up, aggregating data, with scopes ranging from pan-cancer to subtype-specific. Often, these methods propose either to analyze systems and their dynamics in a very specified manner, for example with ODEs, or to retrieve overall characteristics with a statistical model. In the former case, mathematical and computational complexity can make the analysis grow out of the range of feasibility. Also, the building of precise quantitative models requires an amount of data that is not always available when studying large-scale systems with many components and interactions. In the latter, the lack of interpretability of the computed parameters sometimes results in models with predictive power, but limited usefulness for the understanding of disease mechanisms and the design of therapeutic strategies.

While kinetic modeling continues to be a standard of the industry, qualitative, parameter-free modeling strategies have become more common, as they allow systems-level insights into previously inscrutable aspects of the studied system. In particular, logical models can be used to discover regulatory motifs and infer the general input-output behavior of a system. The simplest form of logical modeling is the Boolean formalism, which only considers the presence or absence of an interaction between two molecules, and the direction of the interaction (activation or inhibition). These models have shown that they can be adequate representations of the studied systems, and have helped elucidate many questions about the general behavior of regulatory systems and have even hinted at first principles concerning the organization of life itself. Many extensions of the Boolean mathematical formalism exist, most notably Probabilistic Boolean Networks (PBNs), fuzzy logic networks, or Petri nets. These extensions aim at relaxing the strict on/off constrain on the nodes' values into a quantitative framework, while keeping the simplicity of a formulation of the network as a series of rules, which is the main advantage of this type of modeling, as it allows both faster computations and easier understanding of the biological meaning of the interactions (Le Novère, 2015). These logical

modeling frameworks have also been applied with success to biological problems, most notably PBNs (Trairatphisan *et al.*, 2013).

Here, we propose to model signaling pathways in cancer with Dynamic Bayesian Networks (DBNs), a general type of probabilistic modeling that extends the strictly feed-forward Bayesian Networks into cyclic structures, allowing for positive and negative feedback loops. DBNs have rarely been used to model signaling pathways, and therefore no efficient tool exists for their simulation or their contextualization in light of biological measurements. However, the mathematical similarity between PBNs and DBNs has been demonstrated, and hints at a similar usefulness for DBN modeling of biological systems. In PBNs, the computation of the long-term probability distributions of the nodes requires the simulation of the networks over extended number of generations in order to create a Markov Chain, which can be computationally costly for large networks. In contrast, DBNs compute the expected limit probability of each node directly given its parents, so that the network quickly converges to its steady-state. We propose that these steady-states correspond to cellular states of quasi-stability, so that the probability distributions of the nodes in the network relate to the activity of signal-processing molecules in the cells.

6.1.1 FALCON toolbox

The main contribution of this thesis is the conceptualization, design, implementation and testing of the FALCON modeling toolbox. This toolbox consists of a series of scripts and related functions that reformulate information concerning the prior knowledge about the topology of a certain regulatory network and the empirically determined concentration of the different molecular forms of the components of these networks into an optimization problem aiming at inferring the relative probabilities of the molecules to interact in the specifically tested contexts. By using gradient descent, we are able to learn the best parametrization for the problem, i.e. the parameter set for which the predictions of the simulated system are the most in accordance with the biological measurements. The node values represent the probabilities for molecules to be in their active state, and they can also be understood as the normalized average activities of the nodes, in a system where a population of many molecules is represented by a single node. The computed parameters, or weights, represent the probabilities for the upstream nodes to influence the downstream nodes, and they can also be viewed as the relative influences of the parent nodes on their children nodes. Assessing the state of the different molecules in a regulatory system and the relative strength of the different interactions is useful to estimate the flow of the signal transduction. This flow can however be different depending the context in which the system evolves, i.e. the presence or absence of external stimuli exerting their actions on the different molecules. For this reason, we designed FALCON to be able to integrate multiple experimental conditions in one single optimization problem. This makes the toolbox adequate for the modeling of systems under multiple experimental

conditions, which can be particularly useful to analyze the results of drug screens in cell lines or tumor tissue.

FALCON was conceived to take maximal advantage of the Matlab computing language and its matrix-oriented design. In order to attain the fastest possible speed, we implemented a pipeline in which the interaction network is reformulated as a DBN which is both reduced (unnecessary nodes are removed from the structure) and expanded (multi-input logical interactions are decomposed into their constituent two-input functions). Efficient learning of the parameters is done via gradient descent of the objective function, which consists of the mean squared error between the model predictions (the inferred node values at steady-state) and the experimentally determined concentrations of the biological entities they represent. The chosen optimizer *fmincon* uses the mid-point method to compute empirical gradients, and explicitly uses the hard constrains on the values of the parameter set stemming from the application of the rules of probabilities on the parameters. We initialized the networks from different initial configurations to test for the existence of local minima in the optimization problem. By applying a mode of node initialization based on a normal distribution of initial parameter guesses around their center value, we optimized the speed of parameter learning, especially for large models. Furthermore, we implemented an efficient parallelization scheme to compute multiple optimizations, starting from different sets of random initial configurations, on different cores in parallel.

The choice of optimization method should ideally depend of the nature of the problem being optimized. In this regard, we implemented a gradient-descent algorithm using the interior-point method, a type of finite-difference approximation of the analytical gradient. While this approach is straightforward, it might not be the most efficient. It has been shown(Raue *et al.*, 2013) that in the case of ODEs, it is possible to assess at the same time the solution of as system of ODEs and the local sensitivities of these solutions (Leis & Kramer, 1988), with gains of speed of up to ten times. While the implementation of such second-order solving is not straightforward in our case, as the dynamism of our system is not an inherent part of the modeling but more an artifact of the Bayesian system used, it will be of prime interest to evaluate further improvements of the computational cost of our method.

The choice of normalization scheme ultimately determines the way the data is used for contextualization of the network information. In this regard, we chose to normalize the data to the $[0-1]$ range so that the highest value for each analyte is assigned the value 1 and the lowest, 0. However, a criticism can be formulated, as this scheme implicitly assumes that all of the analyzed molecules vary across experiments in the same range of relative concentrations. Indeed, in the event that a molecule varies by a small amount, this change would carry the same amount of information as the total removal of another protein from the system. In this case, our framework would wrongly weight these two changes in the same way while their biological meaning is probably different. While we have not fully tested the range of possible normalization schemes (and data transformations), we plan to do it in the future in the hope

of identifying the best possible data pre-processing pipeline to maximize the usefulness of our toolbox.

We implemented two ways to assess the accuracy of our parameter estimates. Firstly, FALCON can generate a new set of target node values with the same structure and statistical properties as the original experimental data, a process known as re-sampling. By repeatedly comparing the parameter estimates obtained by re-optimizing the model on this new dataset with the original estimates, a measure of the confidence in our estimates can be produced. While this process can be imperfect due to the inherent technical error on the value of the measurements and our assumptions about the distribution of this error, it is nevertheless the only way to mimic the acquisition of new data, a process which can be expensive, time consuming, or could not be feasible at all depending the nature of the biological system under study (for example, clinical biopsies). A possible extension of this method would be to combine the new synthetic data with the existing one, an iterative process known as data augmentation (Van Dyk & Meng, 2001). Such a process, common in the fields of image recognition and speech analysis, enables to train large models with limited quantities of data and has recently been applied to time-series prediction and phylogenetic analyses (Le Guennec *et al.*, 2016; Rodrigue & Aris-Brosou, 2011). Secondly, we implemented a flexible framework to perform bootstrapping analyses. Bootstrapping is the process in which a new dataset is produced by resampling from the original dataset, with replacement. Such modified dataset contains only measures that were present in the original dataset, but however gives a larger relative importance to some of the measurements and cancels others. Bootstrapped analyses are a way to control for the design of the biological experiment, i.e. to assess to what extent do the inferred parameter values depend on the presence or absence of a particular experimental condition or set of measured proteins. A third method, not implemented during this doctoral thesis, might be to use stochastic gradient descent (for example Metropolis-coupled Gibbs sampler) to infer the parameter estimates. By carefully designing the parameter learning procedure as a probabilistic process, the posterior distributions of the parameter estimates can be retrieved directly and used for statistical testing. However, such procedures generally necessitate large computational resources to generate the Markov Chains to sample from, and might be overly dependent on the specifications of the prior distributions, therefore needing larger quantities of data for training.

We tested FALCON on a series of different models, aiming at reproducing previous predictions generated using a PBN framework. In all cases, we determined that the results of FALCON were equivalent to the PBN estimates, with a gain of speed of several orders of magnitude. Specifically, we were able to generate the same predictions as optPBN (Trairatphisan *et al.*, 2016) on a small model of deregulated PDGF signaling is gastro-intestinal cancer, given either the presence or absence of a specific mutation in the receptor, and the concentration of two inhibitory drugs. Based on the proteomic profiles upon single perturbations (either mutated receptor or drug), FALCON was able to successfully predict the activity of different molecules

in the signaling pathway when both perturbations are present at the same time. During this experiment, we also compared the speed of our computations with another software named CellNetOptR and we determined that FALCON is more than 40 times faster.

FALCON was designed with practicality in mind, and for this reason several types of input formats are accepted. As Excel is one of the most-used software for handling data structures, both the network topology information and the experimental measurements can be presented in the .xls or .xlsx format. Nevertheless, FALCON also accepts flat files with the .txt and the .csv extensions, as well as the .sif format for networks, which is the standard for several network analysis software, notably Cytoscape. One drawback of the toolbox is the lack of support of the SBML format, which is used for many types of modeling formalisms and software in the biological modeling community. Also, the use of FALCON requires the access to the proprietary Matlab software, which might represent an obstacle for some researchers. For this reason, we are in the process of implementing the FALCON toolbox as a Python package. To have a Python version available will increase the usability of the results of this research on multiple platforms and the availability of this tool to more researchers.

The interest in the analysis of regulatory systems consists not only in the fitting of a specific model on specific data, but in the understanding of the global functional specificities of the system as a whole. Indeed, most often regulatory systems are investigated for the express reason that we want to control the system in a particular way, i.e. transfer the system from one state representing an unfavorable configuration to another, more favorable one, for example from disease to health. In order to identify new control points for the regulatory system being studied, the characteristics of the inferred model need to be investigated systematically. These control points might represent new targets for the design of anti-cancer therapies. We implemented global parameter sensitivity analysis in FALCON, in order to assess the identifiability of the parameters in the models. Indeed, in large networks, the interplay of the different parts of the network results in an apparent redundancy, i.e. a situation where the same set of functional behaviors can be simulated with different parametrizations. In such a case, the actual value of the model parameters is not knowable unless we make additional assumptions for the model or collect additional data. By systematically testing the ability of the network to reproduce the entire series of functional information contained in the dataset, but with slight perturbations of one parameter at at time around its optimal value, it is possible to detect which parts of the model are the least constrained by the data. Such information, in turn, could be useful when designing experimental setups or when merging different datasets together.

Furthermore, the nature of the chemotherapeutic agents often dictates which type of biological target can be efficiently altered in clinical practice. For example, transcription factors have long been considered 'undruggable', with the exception of ligand-inducible nuclear receptors, because of the lack of existing mechanisms by which we could control them. In contrast, most current targeted therapies target receptors expressed at the surface of the cells, for example

with monoclonal antibodies, or are directed towards kinases in the cytoplasm, with small-molecule inhibitors. These agents can act either by blocking the target completely by limiting its ability to become activated by a certain molecule, by binding in an inhibitory pocket, or by blocking the target's interaction with its own target molecules. In order to search for such nodes in the network, we implemented a pipeline to perform *in silico* knock-outs, which can be either at the level of individual nodes or individual interactions, and assess the ability of the resulting network to maintain its functions. The resulting profiles highlight the nodes and interactions for which a chemical intervention is predicted to have the most profound effect, as well as the ones where efficient blocking or inhibition does not in itself lead to a change in the functionality of the regulatory network.

One clear drawback of the mathematical framework used here is its inherent limitation to the study of systems at steady-state. It could be argued that living systems, being continuously metabolically active, only reach a steady-state when they die (or not even then). Although the Bayesian Networks we use are Dynamic, the strict updating scheme used forces all interactions to occur at the same frequency. In our framework, the rates of the reactions and the relative rarity of the events are represented by the probabilistic weights of the interactions, in such a way that the set of states traversed by the model before reaching its steady-state might not represent the real system, as the different processes taking place in living systems occur across widely varying time-scales. For example, significant changes occur over milliseconds for concentrations of certain ions, over seconds to minutes for signaling pathways, and over hours for gene expression levels. Therefore, the progression of these different processes in living cells can be better viewed as series of rapid changes followed by periods of relative stability. The assumption of modeling frameworks like ours, which perform the comparison between biological data and networks at the steady-state, is that at the time of measurements, the fast-acting regulatory subsystems, which are represented in the model, have reached an equilibrium in their activation levels while the influence of the slower-acting subsystems, which are not represented in the model, is not yet relevant. In the case of the investigation of cancer signaling pathways, the assumption is that upon receptor activation by a ligand or kinase blocking by a small-molecule inhibitor, the various signaling pathways reach an equilibrium before gene expression changes leads to down-regulation of the constituents of the signaling pathway, by either modifying the expression level of the targets themselves, or their inhibitors or activators. Ideally, a framework could be designed in which the different interactions of the DBN occur at different frequencies, thereby mimicking the dynamic behavior of real biological systems and allowing the study of their dynamical behavior. An example of such framework is logical ODE modeling, available in the software package CNetOptR. However, such modeling framework requires at least one additional parameter per interaction, and might therefore result in increased computational time. Also, the availability of high-quality time-resolved phosphoproteomics datasets is very limited, as such data is difficult and costly to produce. Thus, these additional temporal parameters might be unidentifiable and not result in increased

accuracy of the predictions or usability of the models to investigate new clinical therapies. Such improvements therefore fall outside of the scope of the present work.

Finally, the performance of the DBN framework for the modeling of regulatory systems is extremely dependent on the quality of the topological information. This information is usually retrieved from databases, and is aggregated and curated in light of the current knowledge to produce the networks used for modeling. However, this knowledge is necessarily inaccurate and partial, in the sense that it is unlikely that the final network comprises all the interactions that are relevant in a certain context, and none of the ones that are irrelevant. This dependency is an inherent disadvantage of network-based approaches, and could be mitigated with techniques to optimize the structure of a regulatory network and contextualize its characteristics at the same time. While such optimization can be achieved in some cases with regularization, this technique is only able to remove unnecessary parametrizations from the optimization problem in light of the data, but not to achieve the full scope of topological modifications that could be achieved with, for example, a topological search with a genetic algorithm. Nevertheless, while the high dependency on prior knowledge networks exists, its effects on the quality of the modeling results can be controlled by carefully choosing of the studied system and the research question.

6.1.2 Regularization for model selection

All models are an imperfect representation of the underlying system. However, the goal of modeling is usually not to have the best possible virtual representation of reality but to be able to answer a finite set of research questions. Both the knowledge used to construct the model structure and the measurements used to contextualize it contain an unknown quantity of error. Therefore, to be useful, a model must be of a certain size, which is dependent on the quantity and quality of the data available for its parametrization. For this reason, regularization is often used in statistical modeling to limit the number of features of interest and their weights in the final model, with the explicit goal to reduce overfitting, i.e. increase the robustness and the accuracy of the predictions. However, the most used forms of regularization (the L1 norm, the L2 norm, and their combination) are not immediately applicable in the probabilistic context of DBNs. We implemented a partial-norm regularization, which penalizes the multiplicity of parent nodes for all nodes of the network. We designed an analysis pipeline in which the same prior-knowledge DBN is iteratively contextualized on the same dataset with various regularization strengths. We observed that such screening identifies the critical values of the regularization strength where the parametrization changes, and by using measures of model size adequacy like the Bayesian Information Criterion, we are able to retrieve a configuration of the DBN which minimizes simultaneously its error and its size. By applying such analytic method on DBN models of signaling pathways, it becomes possible to prune out the interactions that are probably not relevant in the context in which the measurements were taken.

Furthermore, regularization of an optimization problem can be used to incorporate additional assumptions about the underlying system. For example, one frequent research question is the identification of significative differences between different biological contexts, for example different cell lines or patients. We implemented a form of group-wise parameter regularization in FALCON, in the form of a penalty proportional to the L1 norm of the differences between parameter values within the same group of parameters. We designed an analysis pipeline in which a background DBN model is contextualized for multiple cell lines in parallel, as part of the same optimization problem. To identify interactions which might be differentially parametrized between the cell lines, we group the network parameters across the contexts and apply group-wise regularization. We observed that, as in the case with partial norm regularization, such screening identifies the critical values of the regularization strength where the configuration of the parameters within one group, representing the same interaction for different cell lines, changes. As regularization strength increases, the parts of the network for which the differences between the parameter values is small become more homogeneous unless this difference is well supported by the data. By applying this strategy on drug screens performed on multiple cell lines, it becomes possible to assess which interactions are probably the most differentially active between the different cell lines.

However, the implicit assumption of the latter regularization scheme is that within the set of cell lines, each interaction in the regulatory network is either similar for all cell lines, or that the differences between the parameter values are randomly distributed. However, the genetic insults giving rise to cancer induce perturbations in molecular interactions which are likely to occur in discrete steps. Indeed, driver mutations often result in the complete loss of the function of a certain protein, for example p53, or a certain level of constitutive enzymatic activity, for example the common mutations of genes of the RAS family. Therefore, we implemented a regularization scheme in FALCON which favors the formation of an unknown number of clusters within each group of parameters in the parameter space. We accomplished this by computing a measure of the difference between the observed local densities of the parameters and their expected densities under the assumption of a uniform distribution. In practice, when applied to a DBN model in the form of a penalty in the objective function, this uniformity-based procedure encourages the smoothing of group-wise differences in parametrization of the model. We applied FALCON to a DBN model of the signaling of colorectal cancer and phosphoproteomic measurements across the signaling pathways for 14 cell lines, both published in Eduati *et al.* (2017). We observed that our final, regularized model groups the parameters to a single value for most interactions, but points to the heterogeneity of the interaction's strength in other cases. Some of these interactions were related to the regulation of the Akt-PI3K pathway, and were correlated to the cell lines sensitivities to a series of inhibitors of the MAPK pathway. These results indicate that the assumption that the group-wise differences in parametrization are mostly due to noise and can be smoothed out unless they are supported by the data, helps in reducing the size of the model several-fold and in detecting the parts of the model that are the most critical in reproducing the differences in the functional behavior

of different cell lines. These differences, in turn, might reflect important features associated with the specific pathway configuration, for example drug sensitivity.

We evaluated this new uniformity-based regularization scheme against the k-means algorithm, a standard method for cluster inference, as well as with two other methods for assessing the uniform spread of the parameters in the parameter space. We concluded that the computational cost associated with our method was negligible in comparison with other methods. Importantly, we also concluded that this cost would scale linearly with the number of dimension tested, while it would grow faster with the other methods. For example, while the time-complexity of the k-means algorithm is known to be $O(n^2)$, our algorithm most probably has a complexity of $O(n)$, as long as we keep the dimensions independent (i.e. the parameter space has a number of dimensions equal to the number of parameters and all dimensions are orthogonal to all others).

Critically, the structure of the groupings induced by the use of regularization is highly dependent on the quality of the data used for contextualizing the models, as well as on the accuracy of the prior topological information. Therefore, a possible extension of the current framework is the assessment of the robustness of these grouping in the light of slightly different data, for example by using bootstrapping. Alternatively, perturbed data (data to which a subjective amount of error is added artificially) could be used for the same goal. Such an assessment would allow to provide statistical estimates of the reliability of the groupings.

There are different ways in which regularization could be incorporated in the objective function of a contextualization algorithm within a network model of cancer signaling pathways. For example, our computations of uniformity-based and group-wise penalties operate across all parameter groups independently. However, regulatory pathways can display a level of structure in which molecular interactions are organized by motifs. Incorporating multiple parameters into the same penalty function, in such a way that multiple pathway-level differences in parametrizations can either compensate or reinforce each other, might help in individualizing the most significant differences between functionally diverging cell lines or patient profiles.

Similarly, in our framework the steady-state DBN models represent the biological system in a state of quasi-equilibrium. Therefore, time-resolved models can only be constructed by assuming a succession of quasi-equilibria at which time the measurements were made, fitting an independent model for each timepoint. This procedure of stacking models through time is not ideal, as it wrongly assumes time-independence of the model parameters. We can probably safely assume that as regulatory systems are stable and maintain homeostasis in general, in most stable systems the majority of parameters will remain constant through time or change only slowly. We implemented a regularization scheme in which we penalized the same DBN model for different timepoints while incorporating a penalty proportional to the sum of the differences in parametrization between successive time-points. When applied to the modeling of cancer signaling pathways, the goal of this procedure is to smooth out the small

differences in parametrization between DBN models of successive quasi-equilibria, in order to help individualizing the most crucial changes, i.e. the timepoints at which significant pathway re-wiring occurs as a consequence of secondary regulatory effects, like long-term changes in the expression level of a compensatory mutation or the acquisition of a secondary mutation in a cell population.

6.1.3 Using multi-dimensional regularization to infer cell line-specific interactions

While the clinical response rates for targeted small-molecule kinase inhibitors is typically high, the high relapse rates point to unknown defense mechanisms against cell death that render many cancer types resistant to current therapeutic regimes. One example of this situation is *BRAF*-mutated melanoma, representing more than 50% of diagnosed metastatic melanoma cases and for which multiple mechanisms have been proposed leading to intrinsic or acquired resistance against BRAF-inhibitors. Similarly, while TRAIL, and in general TRAIL receptor agonists have been investigated as potential cancer agents for decades, they generally failed during clinical trials due to the high levels of acquired resistance observed in cancer cells. Overall, the clinical needs of many cancer patients are not met yet, but increasing knowledge of the potential resistance mechanisms points to possible future therapies and stratification strategies based on the adaptation and combination of several targeted therapeutics.

We adapted a subpopulation of A375 melanoma to low concentrations of IZI1551, a new multimeric TRAIL receptor agonist. We observed that while in the naive A375 cells, apoptotic cell death could be efficiently triggered by larger doses of IZI1551, the conditioned cells were largely resistant. We compared the proteomic profiles of the parental and adapted cell lines and modeled the differential configurations of the regulatory network, including the main components of the apoptotic and MAPK signaling pathways. We applied a sophisticated regularization scheme, comprised of dual penalization of the model size across the dimensions of the individual interactions and the group-wise differences between both cell lines. By screening an array of regularization strengths, we individualize one configuration of the model in which the goodness-of-fit of the overall model is balanced with its size, according to the Bayesian Information Criterion. Analysis of this regularized model revealed that a subset of reactions, mainly linked to NFκB and anti-apoptotic signaling, were strongly up-regulated in conditioned IZI1551-resistant cells, whereas the pro-apoptotic signaling seemed up-regulated in the parental cell line. The model predicted inhibition of NFκB via IκBSR expression to partially and XIAP knockout to fully re-sensitize the conditioned cells to IZI1551. These predictions were validated *in vitro*, not only in the A375 cell line used for model parameterization, but also in Malme3M, a different BRAF-mutated cell line.

Our results show that DBN modeling, combined with the regularization strategy used, accurately predicted XIAP to be the key player in conferring TRAIL resistance in BRAF-mutated

melanoma cell lines. Our study also shows that this resistance mechanism is conserved between A375 and Malme3M cells. If the correlation between elevated XIAP expression and intrinsic TRAIL resistance in melanoma cells is revealed to be valid across a wider panel of cell lines and patient-derived material, this would indicate that XIAP may serve as a biomarker for TRAIL responsiveness in patients diagnosed with melanoma. The success of our strategy to identify the main control points in a small DBN model of signaling pathways indicates that such an approach might play a larger role in the future, notably in the use of large drug screens to investigate resistance mechanisms in cancer. Potentially, integration of the above-mentioned regularization schemes (across time or across large sets of cell lines) within the existing regularization analysis pipeline could unveil more details about the particular resistance and adaptation strategies of cancer cells to therapeutic agents, therefore revealing for each patients subset the best treatment options.

6.2 Perspectives

In this work, a tool for the modeling of signaling pathways is designed and applied to the simulation of perturbed regulatory networks in cancer. In the future, such modeling approaches are likely to have a growing impact on the design of new therapies. As such, logical modeling, and in particular DBN modeling will likely find their best application in the explorative, large-scale modeling of regulatory systems, at the cellular or tissular levels, while more detailed modeling techniques, like ODE systems, might be preferred either for the detailed modeling of the mechanism of action of drugs in smaller, highly-parametrized systems, or for inherently dynamic processes, like pharmacodynamic whole-body models for example.

In particular, FALCON could be applied to multi-pathway cancer models parametrized with patient-specific data. High-throughput measurements could be obtained from sets of patient-derived material or cultured spheroids, and inform caretakers about the specific pathways active in the patients cancer cells. Such information could help prioritize the personal combination therapies which have the highest likelihood of generating a long-term complete response in the patients. In this context, the use of regularized objective functions, taking into account the inaccuracy of the prior-knowledge DBN, the technical error in the measurements, and the distribution of the model parameters amongst the diagnosed patients, should play a determinant role in ensuring the success of such personalized cancer treatments.

Bibliography

Aagaard, Lars, & Rossi, John J. 2007. RNAi therapeutics: principles, prospects and challenges. *Advanced Drug Delivery Reviews*, **59**(2-3), 75–86.

Albert, R. 2005. Scale-free networks in cell biology. *Journal of Cell Science*, **118**(21), 4947–4957.

Albert, Réka, & Thakar, Juilee. 2014. Boolean modeling: A logic-based dynamic approach for understanding signaling and regulatory networks and for making useful predictions. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, **6**(5), 353–369.

Alberts, Bruce, Johnson, Alexander, Lewis, Julian, Raff, Martin, Roberts, Keith, & Walter, Peter. 2002. *Molecular Biology of the Cell, 4th edition*. Garland sc edn. New York: Garland Science (NY).

Aldridge, Bree B., Burke, John M., Lauffenburger, Douglas A., & Sorger, Peter K. 2006. Physicochemical modelling of cell signalling pathways. *Nature Cell Biology*, **8**(11), 1195–1203.

Alon, Uri. 2007. Network motifs: Theory and experimental approaches. *Nature Reviews Genetics*, **8**(6), 450–461.

Anand, Preetha, Kunnumakara, Ajaikumar B., Sundaram, Chitra, Harikumar, Kuzhuvelil B., Tharakan, Sheeja T., Lai, Oiki S., Sung, Bokyung, & Aggarwal, Bharat B. 2008. Cancer is a preventable disease that requires major lifestyle changes. *Pharmaceutical Research*, **25**(9), 2097–2116.

Anderson, T W, & Darling, A D. 1954. A test of goodness of fit. *Journal of the American Statistical Association*, **49**(268), 765–769.

Anton, Delphine, Burckel, Hélène, Josset, Elodie, & Noel, Georges. 2015. Three-dimensional cell culture: A breakthrough in vivo. *International Journal of Molecular Sciences*, **16**(3), 5517–5527.

Ashdown, Martin L., Robinson, Andrew P., Yatomi-Clarke, Steven L., Ashdown, Martin Luisa, Allison, Andrew, Abbott, Derek, Markovic, Svetomir N., & Coventry, Brendon J. 2015. Chemotherapy for late-stage cancer patients: meta-analysis of complete response rates. *F1000 Research*, **4**(0), 232.

Ayyadurai, V. A S, & Dewey, C. Forbes. 2011. CytoSolve: A scalable computational method for dynamic integration of multiple molecular pathway models. *Cellular and Molecular Bioengineering*, **4**(1), 28–45.

Barabási, Albert-László. 2007. Network medicine From obesity to the diseasome. *New England Journal of Medicine*, **357**(4), 404–407.

Barabási, Albert-László, Gulbahce, Natali, & Loscalzo, Joseph. 2011. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics*, **12**(1), 56–68.

Bassères, D. S., & Baldwin, A. S. 2006. Nuclear factor-κB and inhibitor of κB kinase pathways in oncogenic initiation and progression. *Oncogene*, **25**(51), 6817–6830.

Boespflug, Amélie, Caramel, Julie, Dalle, Stephane, & Thomas, Luc. 2017. Treatment of NRAS-mutated advanced or metastatic melanoma: rationale, current trials and evidence to date. *Therapeutic Advances in Medical Oncology Review*, **9**(7), 481–492.

Bollag, Gideon, Hirth, Peter, Tsai, James, Zhang, Jiazhong, Ibrahim, Prabha N., Cho, Hanna, & Al., Et. 2010. Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature*, **467**(3), 596–599.

Boole, George. 1854. *An investigation of the laws of thought.* Project Gutenberg.

Bornholdt, Stefan, & Bornholdt, Stefan. 2005. Less is more in modeling large genetic networks. *Science*, **310**(October), 449–451.

Bourdon, Jérémie, & Roux, Olivier. 2016. Computational methods in systems biology. *Pages 1–2 of: Biosystems*, vol. 149.

Box, George E P. 1976. Science and statistics. *Journal of the American Statistical Association*, **71**(356), 791–799.

Braeuer, Susanne J, Büneker, Chirlei, Mohr, Andrea, & Zwacka, Ralf Michael. 2006. Constitutively activated nuclear factor-kappaB, but not induced NF-kappaB, leads to TRAIL resistance by up-regulation of X-linked inhibitor of apoptosis protein in human cancer cells. *Molecular Cancer Research*, **4**(10), 715–28.

Bühlmann, Peter, Rütimann, Philipp, van de Geer, Sara, & Zhang, Cun Hui. 2013. Correlated variables in regression: Clustering and sparse estimation. *Journal of Statistical Planning and Inference*, **143**(11), 1835–1858.

Burnham, Kenneth P., & Anderson, David R. 2002. *Model selection and multimodel inference, a practical information-theoretic approach.* New York: Springer-Verlag.

Burnham, Kenneth P., & Anderson, David R. 2004. Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, **33**(2), 261–304.

Calegari, Silvia, & Ciucci, Davide. 2006. Integrating fuzzy logic in ontologies. *Pages 66–73 of: Proceedings of the Eighth International Conference on Enterprise Information Systems.* SciTePress - Science and and Technology Publications.

Cargnello, M., & Roux, P. P. 2011. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiology and Molecular Biology Reviews*, **75**(1), 50–83.

Cerami, Ethan G., Gross, Benjamin E., Demir, Emek, Rodchenkov, Igor, Babur, Özgün, Anwar, Nadia, Schultz, Nikolaus, Bader, Gary D., & Sander, Chris. 2010. Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Research*, **39**(SUPPL. 1), 685–690.

Chakrabarty, Broto, & Parekh, Nita. 2016. NAPS: Network analysis of protein structures. *Nucleic Acids Research*, **44**(W1), W375–W382.

Chao, Ann, Connell, Cari J, Mccullough, Marjorie L, Jacobs, Eric J, Flanders, W Dana, Rodriguez, Carmen, & Calle, Eugenia E. 2005. Meat consumption and risk of colorectal cancer. *JAMA*, **293**(2), 172–182.

Chawla-Sarkar, M., Bae, S. I., Reu, F. J., Jacobs, B. S., Lindner, D. J., & Borden, E. C. 2004. Downregulation of Bcl-2, FLIP or IAPs (XIAP and survivin) by siRNAs sensitizes resistant melanoma cells to Apo2L/TRAIL-induced apoptosis. *Cell Death and Differentiation*, **11**(8), 915–923.

Chen, Edwin, Staudt, Louis M., & Green, Anthony R. 2012. Janus kinase deregulation in leukemia and lymphoma. *Immunity*, **36**(4), 529–541.

Chickering, David Maxwell, Heckerman, David, & Meek, Christopher. 2004. Large-sample learning of Bayesian networks. *Journal of Machine Learning Research*, **5**(1999), 1287–1330.

Choudhary, Ashish, Datta, Aniruddha, Bittner, Michael L, & Dougherty, Edward R. 2006. Intervention in a family of Boolean networks. *Bioinformatics*, **22**(2), 226–32.

Coley, William B. 1891. Contribution to the knowledge of sarcoma. *Annals of Surgery*, **14**(199), 199–200.

Croce, Carlo M. 2008. Oncogenes and cancer. *The New England Journal of Medicine*, **358**(5), 502–11.

Czitrom, Veronica. 1999. One-factor-at-a-time versus designed experiments. *The American Statistician*, **53**(2), 126.

De Landtsheer, Sébastien, Trairatphisan, Panuwat, Lucarelli, Philippe, & Sauter, Thomas. 2017. FALCON: a toolbox for the fast contextualization of logical networks. *Bioinformatics*, **33**(21), 3431–3436.

De Landtsheer, Sébastien, Lucarelli, Philippe, & Sauter, Thomas. 2018. Using regularization to infer cell line specificity in logical network models of signaling pathways. *Frontiers in Physiology*, **9**(MAY), 1–13.

De Miguel, D, Lemke, J, Anel, Alberto, Walczak, H, & Martinez-Lostao, L. 2016. Onto better TRAILs for cancer treatment. *Cell Death and Differentiation*, **4**(5), 1–15.

Deininger, Michael W. N., & Druker, Brian J. 2003. Specific targeted therapy of chronic myelogenous leukemia with imatinib. *Pharmacological reviews*, **55**(3), 401–423.

Dimberg, Lina Y, Anderson, Charles K, Camidge, Ross, Behbakht, Kian, Thorburn, Andrew, Ford, Heide L, Anderson, Kenneth C., Camidge, Ross, Behbakht, Kian, Thorburn, Andrew, & Ford, Heide L. 2013. On the TRAIL to successful cancer therapy? Predicting and counteracting resistance against TRAIL-based therapeutics. *Oncogene*, **32**(11), 1341–1350.

Dondelinger, Frank, Lèbre, Sophie, & Husmeier, Dirk. 2013. Non-homogeneous dynamic Bayesian networks with Bayesian regularization for inferring gene regulatory networks with gradually time-varying structure. *Machine Learning*, **90**(2), 191–230.

Dutta, J., Fan, Y., Gupta, N., Fan, G., & Gélinas, C. 2006. Current insights into the regulation of programmed cell death by NF-κB. *Oncogene*, **25**(51), 6800–6816.

Edmondson, Rasheena, Broglie, Jessica Jenkins, Adcock, Audrey F., & Yang, Liju. 2014. Three-dimensional cell culture systems and their applications in drug discovery and cell-based biosensors. *ASSAY and Drug Development Technologies*, **12**(4), 207–218.

Eduati, Federica, Doldàn-Martelli, Victoria, Klinger, Bertram, Cokelaer, Thomas, Sieber, Anja, Kogera, Fiona, Dorel, Mathurin, Garnett, Mathew J., Blüthgen, Nils, & Saez-Rodriguez, Julio. 2017. Drug resistance mechanisms in colorectal cancer dissected with cell type-specific dynamic logic models. *Cancer Research*, canres.0078.2017.

El-Chaar, Nader N., Piccolo, Stephen R., Boucher, Kenneth M., Cohen, Adam L., Chang, Jeffrey T., Moos, Philip J., & Bild, Andrea H. 2014. Genomic classification of the RAS network identifies a personalized treatment strategy for lung cancer. *Molecular Oncology*, **8**(7), 1339–1354.

Euler, Leonhard. 1736. *Solutio problematis ad geometrian situs pertinentis.*

Evans, M. K., Sauer, S. J., Nath, S., Robinson, T. J., Morse, M. A., & Devi, G. R. 2016. X-linked inhibitor of apoptosis protein mediates tumor cell resistance to antibody-dependent cellular cytotoxicity. *Cell Death & Disease*, **7**(1), e2073.

Fan, Yue, Wang, Xiao, & Peng, Qinke. 2017. Inference of gene regulatory networks using Bayesian nonparametric regression and topology information. *Computational and Mathematical Methods in Medicine*, **2017**.

Fidler, Isaiah J, Hart, Ian R, Fidler, Isaiah J, & Hart, Ian R. 1982. Biological diversity in metastatic Neoplasms : origins and implications. *Science*, **217**(4564), 998–1003.

Fink, Susan L., & Cookson, Brad T. 2006. Caspase-1-dependent pore formation during pyroptosis leads to osmotic lysis of infected host macrophages. *Cellular Microbiology*, **8**(11), 1812–1825.

Fitzmaurice, Christina, & The Global Burden of Disease Cancer Collaboration. 2017. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015. *JAMA Oncology*, **3**(4), 524.

Flanagan, L., Kehoe, J., Fay, J., Bacon, O., Lindner, A. U., Kay, E. W., Deasy, J., McNamara, D. A., & Prehn, J. H.M. 2015. High levels of X-linked Inhibitor-of-Apoptosis Protein (XIAP) are indicative of radio chemotherapy resistance in rectal cancer. *Radiation Oncology*, **10**(1), 1–9.

Frede, Dorothea. 1985. *The sea-battle reconsidered. Oxford studies in ancient philosophy.* Oxford.

Friedman, Nir, Linial, Michal, Nachman, Iftach, & Pe'er, Dana. 2000. Using Bayesian networks to analyze expression data. *Proceedings of the fourth annual international conference on Computational molecular biology - RECOMB '00*, 127–135.

Friend, Stephen H., Bernards, Rene, Rogelj, Snezna, Weinberg, Robert A., Rapaport, Joyce M., Albert, Daniel M., & Dryja, Thaddeus P. 1986. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature*, **323**(6089), 643–646.

Fröhlich, Holger, Bahamondez, Gloria, Götschel, Frank, & Korf, Ulrike. 2015. Dynamic Bayesian network modeling of the interplay between EGFR and hedgehog signaling. *PLoS ONE*, **10**(11), 1–14.

Fukuhara, Hiroshi, Ino, Yasushi, & Todo, Tomoki. 2016. Oncolytic virus therapy: A new era of cancer treatment at dawn. *Cancer Science*, **107**(10), 1373–1379.

Fulda, Simone. 2015. Promises and challenges of Smac mimetics as cancer therapeutics. *Clinical Cancer Research*, **21**(22), 5030–5036.

Giacomantonio, Clare E., & Goodhill, Geoffrey J. 2010. A Boolean model of the gene regulatory network underlying Mammalian cortical area development. *PLoS computational biology*, **6**(9), e1000936.

Gilpin, Leilani H., Bau, David, Yuan, Ben Z., Bajwa, Ayesha, Specter, Michael, & Kagal, Lalana. 2018. Explaining explanations: an approach to evaluating interpretability of machine learning. *ArXiv*, 1806.00069v2.

Glorot, Xavier, & Bengio, Yoshua. 2010. Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS)*, **9**, 249–256.

Gottwald, Siegfried. 2001. *A treatise on many-values logic.* London: King's College London.

Griffin, Merope, Scotto, Daniele, Josephs, Debra H, Mele, Silvia, Crescioli, Silvia, Bax, Heather J, Pellizzari, Giulia, Wynne, Matthew D, Nakamura, Mano, Hoffmann, M, Ilieva, Kristina M, Cheung, Anthony, Spicer, James F, Lacy, Katie E, & Karagiannis, Sophia N. 2017. BRAF inhibitors : resistance and the promise of combination treatments for melanoma. *Oncotarget*, **8**(44), 78174–78192.

Grzegorczyk, Marco, & Husmeier, Dirk. 2011. Improvements in the reconstruction of time-varying gene regulatory networks: Dynamic programming and regularization by information sharing among genes. *Bioinformatics*, **27**(5), 693–699.

Hagberg, Aric A., Schult, Daniel A., & Swart, Pieter J. 2008. Exploring network structure, dynamics, and function using NetworkX. *Pages 11–16 of:* Varoquaux, G., Vaught, T., & Millman, J. (eds), *Proceedings of the 7th Python in Science Conference.*

Hanahan, Douglas, & Weinberg, Robert a. 2011. Hallmarks of cancer: the next generation. *Cell*, **144**(5), 646–674.

Harrell, Frank. 1995. *Regression modeling strategies.* Nashville, TN: Springer.

Hartwig, Torsten, Montinaro, Antonella, von Karstedt, Silvia, Sevko, Alexandra, Surinova, Silvia, Chakravarthy, Ankur, Taraborrelli, Lucia, Draber, Peter, Lafont, Elodie, Arce Vargas, Frederick, El-Bahrawy, Mona A., Quezada, Sergio A., & Walczak, Henning. 2017. The TRAIL-induced cancer secretome promotes a tumor-supportive immune microenvironment via CCR2. *Molecular Cell*, **65**(4), 730–742.e5.

He, Kaiming, Zhang, Xiangyu, Ren, Shaoqing, & Sun, Jian. 2015. Deep residual learning for image recognition. *ArXiv*, 1512.03385v1.

Henry, Conor M., & Martin, Seamus J. 2017. Caspase-8 acts in a non-enzymatic role as a scaffold for assembly of a pro-inflammatory FADDosome complex upon TRAIL stimulation. *Molecular Cell*, **65**(4), 715–729.e5.

Hill, Steven M., Lu, Yiling, Molina, Jennifer, Heiser, Laura M., Spellman, Paul T., Speed, Terence P., Gray, Joe W., Mills, Gordon B., & Mukherjee, Sach. 2012. Bayesian inference of signaling network topology in a cancer cell line. *Bioinformatics*, **28**(21), 2804–2810.

Hoadley, Katherine A, Yau, Christina, Wolf, Denise M, Cherniack, Andrew D, Tamborero, David, Ng, Sam, Leiserson, Max D.M., Niu, Beifang, McLellan, Michael D., Uzunangelov, Vladislav, Zhang, Jiashan, Kandoth, Cyriac, Akbani, Rehan, Shen, Hui, Omberg, Larsson, Chu, Andy, Margolin, Adam A, van't Veer, Laura J., Lopez-Bigas, Nuria, Laird, Peter W, Raphael, Benjamin J, Ding, Li, Robertson, A Gordon, Byers, Lauren A, Mills, Gordon B, Weinstein, John N, Van Waes, Carter, Chen, Zhong, Collisson, Eric A, Benz, Christopher C., Perou, Charles M, & Stuart, Joshua M. 2014. Multiplatform analysis of 12 cancer Types reveals molecular classification within and across tissues of origin. *Cell*, **158**(4), 929–944.

Hocking, R R. 1976. The analysis and selection of variables in linear regression. *Biometrics*, **32**(1), 1–49.

Hodgkin, A. L., & Huxley, A. F. 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Bulletin of Mathematical Biology*, **52**(1-2), 25–71.

Hörnle, M, Peters, N, Thayaparasingham, B, Vörsmann, H, Kashkar, H, & Kulms, D. 2011. Caspase-3 cleaves XIAP in a positive feedback loop to sensitize melanoma cells to TRAIL-induced apoptosis. *Oncogene*, **30**(5), 575–587.

Huang, Sui, & Ingber, Donald E. 2000. Shape-dependent control of cell growth, differentiation, and apoptosis: switching between attractors in cell regulatory networks. *Experimental Cell Research*, **261**(1), 91–103.

Huang, Ying, Yang, Xiang, Xu, Tianrui, Kong, Qinghong, Zhang, Yaping, Shen, Yuehai, Wei, Yunlin, Wang, Guanlin, & Chang, Kwen-Jen. 2016. Overcoming resistance to TRAIL-induced apoptosis in solid tumor cells by simultaneously targeting death receptors, c-FLIP and IAPs. *International Journal of Oncology*, 153–163.

Huerta, Edmundo Bonilla, Duval, Béatrice, & Hao, Jin Kao. 2008. Fuzzy logic for elimination of redundant information of microarray data. *Genomics, Proteomics and Bioinformatics*, **6**(2), 61–73.

Hughes, Michelle A., Harper, Nicholas, Butterworth, Michael, Cain, Kelvin, Cohen, Gerald M., & MacFarlane, Marion. 2009. Reconstitution of the death-inducing signaling complex reveals a substrate switch that determines CD95-mediated death or survival. *Molecular Cell*, **35**(3), 265–279.

Hutt, Meike, Marquardt, Lisa, Seifert, Oliver, Siegemund, Martin, Müller, Ines, Kulms, Dagmar, Pfizenmaier, Klaus, & Kontermann, Roland E. 2017. Superior properties of Fc-comprising scTRAIL fusion proteins. *Molecular Cancer Therapeutics*, sep.

Jacob, Laurent, Obozinski, Guillaume, & Vert, Jean-Philippe. 2009. Group lasso with overlap and graph lasso. *Pages 1–8 of: Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09*. Montreal: ACM.

Jenatton, Rodolphe, Audibert, Jean-Yves, & Bach, Francis. 2009. Structured variable selection with sparsity-inducing norms. *Journal of Machine Learning Research*, **12**(apr), 2777–2824.

Jessy, Thomas. 2011. Immunity over inability: the spontaneous regression of cancer. *Journal of Natural Science, Biology and Medicine*, **2**(1), 43.

Joshi-Tope, G., Gillespie, M., Vastrik, I., D'Eustachio, P., Schmidt, E., de Bono, B., Jassal, B., Gopinath, G. R., Wu, G. R., Matthews, L., Lewis, S., Birney, E., & Stein, L. 2005. Reactome: A knowledgebase of biological pathways. *Nucleic Acids Research*, **33**(DATABASE ISS.), 428–432.

Junttila, Melissa R., & De Sauvage, Frederic J. 2013. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature*, **501**(7467), 346–354.

Kandoth, Cyriac, McLellan, Michael D., Vandin, Fabio, Ye, Kai, Niu, Beifang, Lu, Charles, Xie, Mingchao, Zhang, Qunyuan, McMichael, Joshua F., Wyczalkowski, Matthew A., Leiserson, Mark D.M., Miller, Christopher A., Welch, John S., Walter, Matthew J., Wendl, Michael C., Ley, Timothy J., Wilson, Richard K., Raphael, Benjamin J., & Ding, Li. 2013. Mutational landscape and significance across 12 major cancer types. *Nature*, **502**(7471), 333–339.

Karmarkar, N. 1984. A new polynomial-time algorithm for linear programming. *Combinatorica*, **4**(4), 373–395.

Kashkar, Hamid, Seeger, Jens-michael, Hombach, Andreas, Deggerich, Anke, Yazdanpanah, Benjamin, Utermöhlen, Olaf, Heimlich, Gerd, Abken, Hinrich, & Krönke, Martin. 2006. XIAP targeting sensitizes Hodgkin lymphoma cells for cytolytic T-cell attack. *Blood*, **108**(10), 3434–3440.

Kauffman, Stuart. 1969. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, **22**(3), 437–467.

Ketchen, David, & Shook, Christopher. 1996. The application of cluster analysis in strategic management research: An analysis and critique. *Strategic Management Journal*, **17**(6), 441–458.

Khattak, Muhammad, Fisher, Rosalie, Turajlic, Samra, & Larkin, James. 2013. Targeted therapy and immunotherapy in advanced melanoma: an evolving paradigm. *Therapeutic Advances in Medical Oncology*, **5**(2), 105–118.

Kholodenko, Boris N. 2006. Cell-signalling dynamics in time and space. *Nature Reviews Molecular Cell Biology*, **7**(3), 165–76.

Klamt, Steffen, Saez-Rodriguez, Julio, & Gilles, Ernst D. 2007. Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC Systems Biology*, **1**(1), 1–13.

Knudson, A. G. 1971. Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences*, **68**(4), 820–823.

Kolmogorov, A. N. 1956. *Foundations of the theory of probability.* Chelsea Publishing Company (NY).

Kosko, Bart. 1990. Fuzziness vs. probability. *International Journal of General Systems*, **17**(2-3), 211–240.

Kourou, Konstantina, Papaloukas, Costas, & Fotiadis, Dimitrios I. 2017. Integration of pathway knowledge and dynamic bayesian networks for the prediction of oral cancer recurrence. *IEEE Journal of Biomedical and Health Informatics*, **21**(2), 320–327.

Kutmon, Martina, Riutta, Anders, Nunes, Nuno, Hanspers, Kristina, Willighagen, Egon L., Bohler, Anwesha, Mélius, Jonathan, Waagmeester, Andra, Sinha, Sravanthi R., Miller, Ryan, Coort, Susan L., Cirillo, Elisa, Smeets, Bart, Evelo, Chris T., & Pico, Alexander R. 2016. WikiPathways: Capturing the full diversity of pathway knowledge. *Nucleic Acids Research*, **44**(D1), D488–D494.

Lähdesmäki, Harri, Hautaniemi, Sampsa, Shmulevich, Ilya, & Yli-Harja, Olli. 2006. Relationships between probabilistic Boolean networks and dynamic Bayesian networks as models of gene regulatory networks. *Signal Processing*, **86**(4), 814–834.

Lake, David, Corrêa, Sonia A. L., & Müller, Jürgen. 2016. Negative feedback regulation of the ERK1/2 MAPK pathway. *Cellular and Molecular Life Sciences*, **73**(23), 4397–4413.

Lazebnik, Yuri. 2002. Can a biologist fix a radio? - Or, what I learned while studying apoptosis. *Cancer Cell*, **2**(3), 179–182.

Le Guennec, Arthur, Malinowski, Simon, & Tavenard, Romain. 2016. Data augmentation for Time series classification using convolutional neural networks. *2nd ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data.*

Le Novère, Nicolas. 2015. Quantitative and logic modelling of molecular and gene networks. *Nature Reviews Genetics*, **16**(3), 146–158.

Lecis, D, Drago, C, Manzoni, L, Seneci, P, Scolastico, C, Mastrangelo, E, Bolognesi, M, & Anichini, A. 2010. Novel SMAC-mimetics synergistically stimulate melanoma cell death in combination with TRAIL and Bortezomib. *British Journal of Cancer*, **102**(12), 1707–1716.

Leis, Jorge R., & Kramer, Mark A. 1988. The simultaneous solution and sensitivity analysis of systems described by ordinary differential equations. *ACM Transactions on Mathematical Software*, **14**(1), 45–60.

Li, Banghe, Shen, Yuefeng, & Li, Bo. 2008. Quasi-steady-state laws in enzyme kinetics. *Journal of Physical Chemistry A*, **112**(11), 2302–2321.

Li, Fangting, Long, Tao, Lu, Ying, Ouyang, Qi, & Tang, Chao. 2004. The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 4781–4786.

Li, Peng, Zhang, Chaoyang, Perkins, Edward J., Gong, Ping, & Deng, Youping. 2007. Comparison of probabilistic Boolean network and dynamic Bayesian network approaches for inferring gene regulatory networks. *BMC Bioinformatics*, **8**(SUPPL. 7), S13.

Linkermann, Andreas. 2014. Necroptosis. *New England Journal of Medicine*, **370**(5), 455–465.

Lommel, Maiti J., Trairatphisan, Panuwat, Gabler, Karoline, Laurini, Christina, Muller, Arnaud, Kaoma, Tony, Vallar, Laurent, Sauter, Thomas, & Schaffner-Reckinger, Elisabeth. 2016. L-plastin Ser5 phosphorylation in breast cancer cells and in vitro is mediated by RSK downstream of the ERK/MAPK pathway. *FASEB Journal*, **30**(3), 1218–1233.

Lucarelli, Philippe, Schilling, Marcel, Kreutz, Clemens, Vlasov, Artyom, Boehm, Martin E, Iwamoto, Nao, Steiert, Bernhard, Lattermann, Susen, Wäsch, Marvin, Stepath, Markus, Matter, Matthias S, Heikenwälder, Mathias, Hoffmann, Katrin, Deharde, Daniela, Damm, Georg, Seehofer, Daniel, Muciek, Maria, Gretz, Norbert, Lehmann, Wolf D, Timmer, Jens, & Klingmüller, Ursula. 2018. Resolving the combinatorial complexity of Smad protein complex formation and its link to gene expression. *Cell Systems*, **6**(1), 75–89.

Luke, Jason J., Flaherty, Keith T., Ribas, Antoni, & Long, Georgina V. 2017. Targeted agents and immunotherapies: optimizing outcomes in melanoma. *Nature Reviews Clinical Oncology*, **14**(8), 463–482.

Lum, Julian J., DeBerardinis, Ralph J., & Thompson, Craig B. 2005. Autophagy in metazoans: Cell survival in the land of plenty. *Nature Reviews Molecular Cell Biology*, **6**(6), 439–448.

Lynch, T J, Bell, D W, Sordella, R, Gurubhagavatula, S, Okimoto, R A, Brannigan, B W, Harris, P L, Haserlat, S M, Supko, J G, Haluska, F G, Louis, D N, Christiani, D C, Settleman, J, & Haber, D A. 2004. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *The New England Journal of Medicine*, **350**(21), 2129–2139.

MacNamara, Aidan, Terfve, Camille, Henriques, David, Bernabé, Beatriz Peñalver, & Saez-Rodriguez, Julio. 2012. State-time spectrum of signal transduction logic models. *Physical biology*, **9**(4), 045003.

Margolin, Kim. 2016. The promise of molecularly targeted and immunotherapy for advanced melanoma. *Current Treatment Options in Oncology*, **17**(9), 1–14.

Markov, A. 1954. *The theory of algorithms*. Vol. 42. Math-Net.Ru.

Marquart, John, Chen, Emerson Y., & Prasad, Vinay. 2018. Estimation of the percentage of US patients with cancer who benefit from genome-driven oncology. *JAMA Oncology*, **4**(8), 1093–1098.

Marusyk, Andriy, & Polyak, Kornelia. 2010. Tumor heterogeneity: causes and consequences. *Biochimica et Biophysica Acta Reviews on Cancer*, **1805**(1), 105–117.

Massey, Frank J. 1951. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, **46**(253), 68–78.

Mendoza, L., Thieffry, D., & Alvarez-Buylla, E. R. 1999. Genetic control of flower morphogenesis in Arabidopsis thaliana: a logical analysis. *Bioinformatics*, **15**(7-8), 593–606.

Mendoza, Michelle C, Er, E Emrah, & Blenis, John. 2011. The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. *Trends in Biochemical Sciences*, **36**(6), 320–328.

Merkle, Ruth, Steiert, Bernhard, Salopiata, Florian, Depner, Sofia, Raue, Andreas, Iwamoto, Nao, Schelker, Max, Hass, Helge, Wäsch, Marvin, Böhm, Martin E., Mücke, Oliver, Lipka, Daniel B., Plass, Christoph, Lehmann, Wolf D., Kreutz, Clemens, Timmer, Jens, Schilling, Marcel, & Klingmüller, Ursula. 2016. Identification of cell type-specific differences in erythropoietin receptor signaling in primary erythroid and lung cancer cells. *PLoS Computational Biology*, **12**(8), 1–34.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. 2002. Network motifs: simple building blocks of complex networks. *Science*, **298**(5594), 824–827.

Mizushima, Noboru, & Komatsu, Masaaki. 2011. Autophagy: renovation of cells and tissues. *Cell*, **147**(4), 728–741.

Morgan, Josh Lyskowski, Berger, Daniel Raimund, Wetzel, Arthur Willis, & Lichtman, Jeff William. 2016. The Fuzzy Logic of network connectivity in mouse visual thalamus. *Cell*, **165**(1), 192–206.

Moriceau, Gatien, Hugo, Willy, Hong, Aayoung, Shi, Hubing, Kong, Xiangju, Yu, Clarissa C., Koya, Richard C., Samatar, Ahmed A., Khanlou, Negar, Braun, Jonathan, Ruchalski, Kathleen, Seifert, Heike, Larkin, James, Dahlman, Kimberly B., Johnson, Douglas B., Algazi, Alain, Sosman, Jeffrey A., Ribas, Antoni, & Lo, Roger S. 2015. Tunable-combinatorial mechanisms of acquired resistance limit the efficacy of BRAF/MEK cotargeting but result in melanoma drug addiction. *Cancer Cell*, **27**(2), 240–256.

Morris, Luc G T, & Chan, Timothy a. 2015. Therapeutic targeting of tumor suppressor genes. *Cancer*, **121**(9), 1357–1368.

Morris, Melody K., Saez-Rodriguez, Julio, Sorger, Peter K., & Lauffenburger, Douglas A. 2010. Logic-based models for the analysis of cell signaling networks. *Biochemistry*, **49**(15), 3216–3224.

Müller, Ines, Beissert, Stefan, & Kulms, Dagmar. 2014. Anti-apoptotic NFκB and "gain of function" mutP53 in concert act pro-apoptotic in response to UVB+IL-1 via enhanced TNF production. *The Journal of Investigative Dermatology*, 851–860.

Murphy, K P. 2002. *Dynamic Bayesian Networks: representation, inference and learning*. Ph.D. thesis, University of California, Berkeley.

Müssel, Christoph, Hopfensitz, Martin, & Kestler, Hans a. 2010. BoolNet–an R package for generation, reconstruction and analysis of Boolean networks. *Bioinformatics*, **26**(10), 1378–80.

Noble, Denis. 2006. *The music of life: Biology beyond the genome*. Oxford: Oxford University Press.

Paluncic, Jasmina, Kovacevic, Zaklina, Jansson, Patric J., Kalinowski, Danuta, Merlot, Angelika M., Huang, Michael L.H., Lok, Hiu Chuen, Sahni, Sumit, Lane, Darius J.R., & Richardson, Des R. 2016. Roads to melanoma: key pathways and emerging players in melanoma progression and oncogenic signaling. *Biochimica et Biophysica Acta - Molecular Cell Research*, **1863**(4), 770–784.

Pardoll, Drew M. 2016. The blockade of immune checkpoints in cancer immunotherapy. *Nature Reviews Cancer*, **12**(4), 252–264.

Pearce, Alison, Haas, Marion, Viney, Rosalie, Pearson, Sallie-Anne, Haywood, Philip, Brown, Chris, & Ward, Robyn. 2017. Incidence and severity of self-reported chemotherapy side effects in routine care: A prospective cohort study. *PLOS ONE*, **12**(10), e0184360.

Pearl, J. 2014. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Elsevier.

Penfold, Christopher A., & Wild, David L. 2011. How to infer gene networks from expression profiles, revisited. *Interface Focus*, **1**(6), 857–870.

Petry, Sebastian. 2006. Simultaneous regression shrinkage , variable selection and clustering of predictors with OSCAR. *Biometrics*, **64**(November), 1–17.

Prill, Robert J., Marbach, Daniel, Saez-Rodriguez, Julio, Sorger, Peter K., Alexopoulos, Leonidas G., Xue, Xiaowei, Clarke, Neil D., Altan-Bonnet, Gregoire, & Stolovitzky, Gustavo. 2010. Towards a rigorous assessment of systems biology models: The DREAM3 challenges. *PLoS ONE*, **5**(2).

Raue, Andreas, Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., Timmer, J., Schilling, M., Timmer, J., Raue, Andreas, Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingm??ller, U., & Timmer, J. 2009. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, **25**(15), 1923–1929.

Raue, Andreas, Schilling, Marcel, Bachmann, Julie, Matteson, Andrew, Schelke, Max, Kaschek, Daniel, Hug, Sabine, Kreutz, Clemens, Harms, Brian D., Theis, Fabian J., Klingmüller, Ursula, & Timmer, Jens. 2013. Lessons learned from quantitative dynamical modeling in systems biology. *PLoS ONE*, **8**(9).

Raulf, N., El-Attar, R., Kulms, D., Lecis, D., Delia, D., Walczak, H., Papenfuss, K., Odell, E., & Tavassoli, M. 2014. Differential response of head and neck cancer cell lines to TRAIL or Smac mimetics is associated with the cellular levels and activity of caspase-8 and caspase-10. *British Journal of Cancer*, **111**(10), 1955–1964.

Rawlings, J. S. 2004. The JAK/STAT signaling pathway. *Journal of Cell Science*, **117**(8), 1281–1283.

Rebane, George, & Pearl, Judea. 2013. The recovery of causal poly-trees from statistical data. *ArXiv*, mar, 1304.2736.

Richmond, Ann, & Ueda, Yukiko. 2013. NF-κB activation in melanoma. *Nature Reviews Cancer*, **13**(2), 83–96.

Rigden, Daniel J., Fernández-Suárez, Xosé M., & Galperin, Michael Y. 2016. The 2016 database issue of nucleic acids research and an updated molecular biology database collection. *Nucleic Acids Research*, **44**(D1), D1–D6.

Rodrigue, Nicolas, & Aris-Brosou, Stéphane. 2011. Fast bayesian choice of phylogenetic models: Prospecting data augmentation-based thermodynamic integration. *Systematic Biology*, **60**(6), 881–887.

Roser, Max, & Ritchie, Hannah. 2019. *Cancer*.

Rumelhart, David E., Hinton, Geoffry E., & William, Ronald J. 1986. Learning representations by back-propagating errors. *Nature*, **323**(6058), 533–536.

Rybak, Adrian P., Bristow, Robert G., & Kapoor, Anil. 2015. Prostate cancer stem cells: deciphering the origins and pathways involved in prostate tumorigenesis and aggression. *Oncotarget*, **6**(4), 1900–1919.

Saadatpour, Assieh, Wang, Rui Sheng, Liao, Aijun, Liu, Xin, Loughran, Thomas P., Albert, István, & Albert, Réka. 2011. Dynamical and structural analysis of a t cell survival network identifies novel candidate therapeutic targets for large granular lymphocyte leukemia. *PLoS Computational Biology*, **7**(11).

Sabers, C. J., Martin, M. M., Brunn, G. J., Williams, J. M., Dumont, F. J., Wiederrecht, G., & Abraham, R. T. 1995. *Isolation of a protein target of the FKBP12-rapamycin complex in mammalian cells*.

Saez-Rodriguez, Julio, Simeoni, Luca, Lindquist, Jonathan A., Hemenway, Rebecca, Bommhardt, Ursula, Arndt, Boerge, Haus, Utz Uwe, Weismantel, Robert, Gilles, Ernst D., Klamt, Steffen, & Schraven, Burkhart. 2007. A logical model provides insights into T cell receptor signaling. *PLoS Computational Biology*, **3**(8), 1580–1590.

Saez-Rodriguez, Julio, Alexopoulos, Leonidas G., Zhang, Ming Sheng, Morris, Melody K., Lauffenburger, Douglas A., & Sorger, Peter K. 2011. Comparing signaling networks between normal and transformed hepatocytes using discrete logical models. *Cancer Research*, **71**(16), 5400–5411.

Saleh, Maya, Vaillancourt, John P., Graham, Rona K., Huyck, Matthew, Srinivasula, Srinivasa M., Alnemri, Emad S., Steinberg, Martin H., Holan, Vikki, Baldwin, Clinton T., Hotchkiss, Richard S., Buchman, Timothy G., Zehnbauer, Barbara A., Hayden, Michael R., Farrer, Lindsay A., Roy, Sophie, & Nicholson, Donald W. 2004. Differential modulation of endotoxin responsiveness by human caspase-12 polymorphisms. *Nature*, **429**(6987), 75–79.

Schivo, Stefano, Scholma, Jetse, van der Vet, Paul E., Karperien, Marcel, Post, Janine N., van de Pol, Jaco, & Langerak, Rom. 2016. Modelling with ANIMO: between fuzzy logic and differential equations. *BMC Systems Biology*, **10**(1), 56.

Schlatter, Rebekka, Schmich, Kathrin, Vizcarra, Ima Avalos, Scheurich, Peter, Sauter, Thomas, Borner, Christoph, Ederer, Michael, Merfort, Irmgard, & Sawodny, Oliver. 2009. ON/OFF and beyond - A Boolean model of apoptosis. *PLoS Computational Biology*, **5**(12).

Schwarz, Gideon. 1978. Estimating the dimension of a model. *Annals of Statistics*, **6**(2), 461–464.

Scott, Ronald Bodley. 1970. Cancer chemotherapy - the first twenty-five years. *British Medical Journal*, **4**(5730), 259–265.

Sever, R., & Brugge, J. S. 2015. Signal transduction in cancer. *Cold Spring Harbor Perspectives in Medicine*, **5**(4), a006098–a006098.

Shannon, P. 2003. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, **13**(11), 2498–2504.

She, Yiyuan. 2010. Sparse regression with exact clustering. *Electronic Journal of Statistics*, **4**, 1055–1096.

Sherbenou, Daniel W, & Druker, Brian J. 2007. Applying the discovery of the Philadelphia chromosome. *The Journal of Clinical Investigation*, **117**(8), 2067–2074.

Shmulevich, I., & Dougherty, E.R. 2002. From Boolean to probabilistic Boolean networks as models of genetic regulatory networks. *Proceedings of the IEEE*, **90**(11), 1778–1792.

Shmulevich, Ilya, Dougherty, Edward R., Kim, Seungchan, & Zhang, Wei. 2002. Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, **18**(2), 261–274.

Shtivelman, Emma, Davies, Michael A., Hwu, Patrick, Yang, James, Lotem, Michal, Oren, Moshe, Flaherty, Keith T., Fisher, David E., Shtivelman, Emma, Davies, Michael A., Hwu, Patrick, Yang, James, Lotem, Michal, Oren, Moshe, Flaherty, Keith T., & Fisher, David E. 2014. Pathways and therapeutic targets in melanoma. *Oncotarget*, **5**(7), 1701–1752.

Siegel, Rebecca L, Miller, Kimberly D, & Jemal, Ahmedin. 2016. Cancer statistics, 2016. *CA: a Cancer Journal for Clinicians*, **66**, 7–34.

Siegemund, Martin, Schneider, Felix, Hutt, Meike, Seifert, Oliver, Müller, Ines, Kulms, Dagmar, Pfizenmaier, Klaus, & Kontermann, Roland E. 2018. IgG-single-chain TRAIL fusion proteins for tumour therapy. *Scientific Reports*, **8**(1), 1–11.

Simon, Noah, Friedman, Jerome, Hastie, Trevor, & Tibshirani, Robert. 2012. A sparse-group lasso. *Journal of Computational and Graphical statistics*, **22**(2), 231–245.

Smith, Michael P., Sanchez-Laorden, Berta, O'brien, Kate, Brunton, Holly, Ferguson, Jennifer, Young, Helen, Dhomen, Nathalie, Flaherty, Keith T., Frederick, Dennie T., Cooper, Zachary A., Wargo, Jennifer A., Marais, Richard, & Wellbrock, Claudia. 2014. The immune microenvironment confers resistance to MAPK pathway inhibitors through macrophage-derived TNF-alpha. *Cancer Discovery*, **4**(10), 1214–1229.

Sobol, I. M. 1976. Uniformly distributed sequences with an additional uniform property. *USSR Computational Mathematics and Mathematical Physics*, **16**(5), 236–242.

Sobol, I. M. 1990. Sensitivity analysis for non-linear mathematical models. *Mathematical Modeling and Computational experiment*, **1**(may), 407–414.

Sperandio, S., de Belle, I., & Bredesen, D. E. 2000. An alternative, nonapoptotic form of programmed cell death. *Proceedings of the National Academy of Sciences*, **97**(26), 14376–14381.

Spitzer, Matthew H., & Nolan, Garry P. 2016. Mass Cytometry: Single Cells, Many Features. *Cell*, **165**(4), 780–791.

Steiert, Bernhard, Timmer, Jens, & Kreutz, Clemens. 2016. L1 regularization facilitates detection of cell type-specific parameters in dynamical systems. *Bioinformatics*, **32**(17), i718–i726.

Sullivan, Ryan J., & Flaherty, Keith T. 2013. Resistance to BRAF-targeted therapy in melanoma. *European Journal of Cancer*, **49**(6), 1297–1304.

Sullivan, Ryan J., & Flaherty, Keith T. 2014. Major therapeutic developments and current challenges in advanced melanoma. *British Journal of Dermatology*, **170**(1), 36–44.

Szklarczyk, Damian, Morris, John H., Cook, Helen, Kuhn, Michael, Wyder, Stefan, Simonovic, Milan, Santos, Alberto, Doncheva, Nadezhda T., Roth, Alexander, Bork, Peer, Jensen, Lars J., & Von Mering, Christian. 2017. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Research*, **45**(D1), D362–D368.

Takeshige, Kazuhiko, Baba, Misuzu, Tsuboi, Shigeru, Noda, Takeshi, & Ohsumi, Yoshinori. 1992. Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction. *The Journal of cell biology*, **119**(2), 301–311.

Terfve, Camille, Cokelaer, Thomas, Henriques, David, MacNamara, Aidan, Goncalves, Emanuel, Morris, Melody K., van Iersel, Martijn, Lauffenburger, Douglas A., & Saez-Rodriguez, Julio. 2012. CellNOptR: A flexible toolkit to train protein signaling networks to data using multiple logic formalisms. *BMC Systems Biology*, **6**(jan), 133.

Thayaparasingham, B, Kunz, a, Peters, N, & Kulms, D. 2009. Sensitization of melanoma cells to TRAIL by UVB-induced and NF-kappaB-mediated downregulation of xIAP. *Oncogene*, **28**(3), 345–362.

Thiele, Ines, Swainston, Neil, Fleming, Ronan M.T., Hoppe, Andreas, Sahoo, Swagatika, Aurich, Maike K., Haraldsdottir, Hulda, Mo, Monica L., Rolfsson, Ottar, Stobbe, Miranda D., Thorleifsson, Stefan G., Agren, Rasmus, Bölling, Christian, Bordel, Sergio, Chavali, Arvind K., Dobson, Paul, Dunn, Warwick B., Endler, Lukas, Hala, David, Hucka, Michael, Hull, Duncan, Jameson, Daniel, Jamshidi, Neema, Jonsson, Jon J., Juty, Nick, Keating, Sarah, Nookaew, Intawat, Le Novère, Nicolas, Malys, Naglis, Mazein, Alexander, Papin, Jason A., Price, Nathan D., Selkov, Evgeni, Sigurdsson, Martin I., Simeonidis, Evangelos, Sonnenschein, Nikolaus, Smallbone, Kieran, Sorokin, Anatoly, Van Beek, Johannes H.G.M., Weichart, Dieter, Goryanin, Igor, Nielsen, Jens, Westerhoff, Hans V., Kell, Douglas B., Mendes, Pedro, & Palsson, Bernhard O. 2013. A community-driven global reconstruction of human metabolism. *Nature Biotechnology*, **31**(5), 419–425.

Thomas, S. J., Snowden, J. A., Zeidler, M. P., & Danson, S. J. 2015. The role of JAK/STAT signalling in the pathogenesis, prognosis and treatment of solid tumours. *British Journal of Cancer*, **113**(3), 365–371.

Tibshirani, Robert. 1996. Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society Series B*, **58**(1), 267–288.

Tibshirani, Robert, Saunders, Michael, Rosset, Saharon, Zhu, Ji, Knight, Keith, & Watson, I B M T J. 2005. Sparsity and smoothness via the fused LASSO. *Journal of the Royal Statistical Society Series B*, **67**(1), 91–108.

Tikhonov, A N. 1963. On the solution of incorrectly put problems and the regularisation method. *Pages 261–265 of: Outlines Joint Symposium Partial Differential Equations*, vol. 4.

Todd, Robert G, & Helikar, Tomáš. 2012. Ergodic sets as cell phenotype of budding yeast cell cycle. *PloS one*, **7**(10), e45780.

Tomasetti, Cristian, & Vogelstein, Bert. 2015. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science*, **347**(6217), 78–81.

Trairatphisan, Panuwat, Mizera, Andrzej, Pang, Jun, Tantar, Alexandru Adrian, Schneider, Jochen, & Sauter, Thomas. 2013. Recent development and biomedical applications of probabilistic Boolean networks. *Cell Communication and Signaling*, **11**(1), 46.

Trairatphisan, Panuwat, Mizera, Andrzej, Pang, Jun, Tantar, Alexandru Adrian, & Sauter, Thomas. 2014. optPBN: An optimisation toolbox for probabilistic Boolean networks. *PLoS ONE*, **9**(7), e98001.

Trairatphisan, Panuwat, Wiesinger, Monique, Bahlawane, Christelle, Haan, Serge, & Sauter, Thomas. 2016. A Probabilistic boolean network approach for the analysis of cancer-specific signalling: A case study of deregulated pdgf signalling in GIST. *PLoS ONE*, **11**(5), e0156223.

Tummers, Bart, & Green, Douglas R. 2017. Caspase-8: regulating life and death. *Immunological Reviews*, **277**(1), 76–89.

Van der leeuw, S. E. 2004. Why model? *Cybernetics and Systems*, **35**(2-3), 117–128.

Van Dyk, David A, & Meng, Xiao-li. 2001. The art of data augmentation. *Journal of Computational and Graphical Statistics*, **10**(1), 1–50.

Vanderbei, Robert J., & Carpenter, Tamra J. 1993. Symmetric indefinite systems for interior point methods. *Mathematical Programming*, **58**(1-3), 1–32.

Vignes, Matthieu, Vandel, Jimmy, Allouche, David, Ramadan-Alban, Nidal, Cierco-Ayrolles, Christine, Schiex, Thomas, Mangin, Brigitte, & de Givry, Simon. 2011. Gene regulatory network reconstruction using bayesian networks, the dantzig selector, the lasso and their meta-analysis. *PLoS ONE*, **6**(12).

Villaverde, Alejandro F., & Banga, Julio R. 2014. Reverse engineering and identification in systems biology: strategies, perspectives and challenges. *Journal of the Royal Society Interface*, **11**(91), 20130505.

Vinh, Nguyen Xuan, Chetty, Madhu, Coppel, Ross, & Wangikar, Pramod P. 2011. GlobalMIT: Learning globally optimal dynamic bayesian network with the mutual information test criterion. *Bioinformatics*, **27**(19), 2765–2766.

Vogel, Charles L, Cobleigh, Melody a, Tripathy, Debu, Gutheil, John C, Harris, Lyndsay N, Fehrenbacher, Louis, Slamon, Dennis J, Murphy, Maureen, Novotny, William F, Burchmore, Michael, Shak, Steven, Stewart, Stanford J, Press, Michael, Vogel, CL; Cobleigh, MA; Tripathy, D; Gutheil, JC; Harris, LN; Fehrenbacher, L; Slamon, DJ; Murphy, M; Novotny, WF; Burchmore, M; Shak, S; Stewart, SJ; Press, M;, Vogel, Charles L, Cobleigh, Melody a, Tripathy, Debu, Gutheil, John C, Harris, Lyndsay N, Fehrenbacher, Louis, Slamon, Dennis J, Murphy, Maureen, Novotny, William F, Burchmore, Michael, Shak, Steven, Stewart, Stanford J, & Press, Michael. 2002. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *Journal of clinical oncology*, **20**(3), 719–26.

Vogelstein, Bert, & Kinzler, Kenneth W. 2004. Cancer genes and the pathways they control. *Nature Medicine*, **10**(8), 789–799.

Vogler, Meike, Walczak, Henning, Stadel, Dominic, Haas, Tobias L., Genze, Felicitas, Jovanovic, Marjana, Gschwend, Jürgen E., Simmet, Thomas, Debatin, Klaus Michael, & Fulda, Simone. 2008. Targeting XIAP bypasses Bcl-2-mediated resistance to TRAIL and

cooperates with TRAIL to suppress pancreatic cancer growth in vitro and in vivo. *Cancer Research*, **68**(19), 7956–7965.

Vörsmann, H., Groeber, F., Walles, H., Busch, S., Beissert, S., Walczak, H., & Kulms, D. 2013. Development of a human three-dimensional organotypic skin-melanoma spheroid model for in vitro drug testing. *Cell Death and Disease*, **4**(7), e719.

Waltz, R A, Morales, J L, Nocedal, J, & Orban, D. 2006. An interior algorithm for nonlinear optimization that combines line search and trust region steps. *Mathematical Programming, Series A*, **107**(3), 391–408.

Wang, Rui Sheng, Saadatpour, Assieh, & Albert, Réka. 2012. Boolean modeling in systems biology: An overview of methodology and applications. *Physical Biology*, **9**(5).

Way, Gregory P., Sanchez-Vega, Francisco, La, Konnor, Armenia, Joshua, Chatila, Walid K., Luna, Augustin, Sander, Chris, Cherniack, Andrew D., Mina, Marco, Ciriello, Giovanni, Schultz, Nikolaus, Caesar-Johnson, Samantha J., Demchok, John A., Felau, Ina, Kasapi, Melpomeni, Ferguson, Martin L., Hutter, Carolyn M., Sofia, Heidi J., Tarnuzzer, Roy, Wang, Zhining, Yang, Liming, Zenklusen, Jean C., Zhang, Jiashan (Julia), Chudamani, Sudha, Liu, Jia, Lolla, Laxmi, Naresh, Rashi, Pihl, Todd, Sun, Qiang, Wan, Yunhu, Wu, Ye, Cho, Juok, DeFreitas, Timothy, Frazer, Scott, Gehlenborg, Nils, Getz, Gad, Heiman, David I., Kim, Jaegil, Lawrence, Michael S., Lin, Pei, Meier, Sam, Noble, Michael S., Saksena, Gordon, Voet, Doug, Zhang, Hailei, Bernard, Brady, Chambwe, Nyasha, Dhankani, Varsha, Knijnenburg, Theo, Kramer, Roger, Leinonen, Kalle, Liu, Yuexin, Miller, Michael, Reynolds, Sheila, Shmulevich, Ilya, Thorsson, Vesteinn, Zhang, Wei, Akbani, Rehan, Broom, Bradley M., Hegde, Apurva M., Ju, Zhenlin, Kanchi, Rupa S., Korkut, Anil, Li, Jun, Liang, Han, Ling, Shiyun, Liu, Wenbin, Lu, Yiling, Mills, Gordon B., Ng, Kwok Shing, Rao, Arvind, Ryan, Michael, Wang, Jing, Weinstein, John N., Zhang, Jiexin, Abeshouse, Adam, Armenia, Joshua, Chakravarty, Debyani, Chatila, Walid K., de Bruijn, Ino, Gao, Jianjiong, Gross, Benjamin E., Heins, Zachary J., Kundra, Ritika, La, Konnor, Ladanyi, Marc, Luna, Augustin, Nissan, Moriah G., Ochoa, Angelica, Phillips, Sarah M., Reznik, Ed, Sanchez-Vega, Francisco, Sander, Chris, Schultz, Nikolaus, Sheridan, Robert, Sumer, S. Onur, Sun, Yichao, Taylor, Barry S., Wang, Jioajiao, Zhang, Hongxin, Anur, Pavana, Peto, Myron, Spellman, Paul, Benz, Christopher, Stuart, Joshua M., Wong, Christopher K., Yau, Christina, Hayes, D. Neil, Parker, Joel S., Wilkerson, Matthew D., Ally, Adrian, Balasundaram, Miruna, Bowlby, Reanne, Brooks, Denise, Carlsen, Rebecca, Chuah, Eric, Dhalla, Noreen, Holt, Robert, Jones, Steven J.M., Kasaian, Katayoon, Lee, Darlene, Ma, Yussanne, Marra, Marco A., Mayo, Michael, Moore, Richard A., Mungall, Andrew J., Mungall, Karen, Robertson, A. Gordon, Sadeghi, Sara, Schein, Jacqueline E., Sipahimalani, Payal, Tam, Angela, Thiessen, Nina, Tse, Kane, Wong, Tina, Berger, Ashton C., Beroukhim, Rameen, Cherniack, Andrew D., Cibulskis, Carrie, Gabriel, Stacey B., Gao, Galen F., Ha, Gavin, Meyerson, Matthew, Schumacher, Steven E., Shih, Juliann, Kucherlapati, Melanie H., Kucherlapati, Raju S., Baylin, Stephen, Cope, Leslie, Danilova, Ludmila, Bootwalla, Moiz S., Lai, Phillip H., Maglinte, Dennis T., Van Den Berg, David J., Weisenberger, Daniel J., Auman, J. Todd, Balu, Saianand, Bodenheimer, Tom, Fan, Cheng, Hoadley, Katherine A., Hoyle, Alan P., Jefferys, Stuart R., Jones, Corbin D., Meng, Shaowu, Mieczkowski, Piotr A., Mose,

Lisle E., Perou, Amy H., Perou, Charles M., Roach, Jeffrey, Shi, Yan, Simons, Janae V., Skelly, Tara, Soloway, Matthew G., Tan, Donghui, Veluvolu, Umadevi, Fan, Huihui, Hinoue, Toshinori, Laird, Peter W., Shen, Hui, Zhou, Wanding, Bellair, Michelle, Chang, Kyle, Covington, Kyle, Creighton, Chad J., Dinh, Huyen, Doddapaneni, Harsha Vardhan, Donehower, Lawrence A., Drummond, Jennifer, Gibbs, Richard A., Glenn, Robert, Hale, Walker, Han, Yi, Hu, Jianhong, Korchina, Viktoriya, Lee, Sandra, Lewis, Lora, Li, Wei, Liu, Xiuping, Morgan, Margaret, Morton, Donna, Muzny, Donna, Santibanez, Jireh, Sheth, Margi, Shinbrot, Eve, Wang, Linghua, Wang, Min, Wheeler, David A., Xi, Liu, Zhao, Fengmei, Hess, Julian, Appelbaum, Elizabeth L., Bailey, Matthew, Cordes, Matthew G., Ding, Li, Fronick, Catrina C., Fulton, Lucinda A., Fulton, Robert S., Kandoth, Cyriac, Mardis, Elaine R., McLellan, Michael D., Miller, Christopher A., Schmidt, Heather K., Wilson, Richard K., Crain, Daniel, Curley, Erin, Gardner, Johanna, Lau, Kevin, Mallery, David, Morris, Scott, Paulauskis, Joseph, Penny, Robert, Shelton, Candace, Shelton, Troy, Sherman, Mark, Thompson, Eric, Yena, Peggy, Bowen, Jay, Gastier-Foster, Julie M., Gerken, Mark, Leraas, Kristen M., Lichtenberg, Tara M., Ramirez, Nilsa C., Wise, Lisa, Zmuda, Erik, Corcoran, Niall, Costello, Tony, Hovens, Christopher, Carvalho, Andre L., de Carvalho, Ana C., Fregnani, José H., Longatto-Filho, Adhemar, Reis, Rui M., Scapulatempo-Neto, Cristovam, Silveira, Henrique C.S., Vidal, Daniel O., Burnette, Andrew, Eschbacher, Jennifer, Hermes, Beth, Noss, Ardene, Singh, Rosy, Anderson, Matthew L., Castro, Patricia D., Ittmann, Michael, Huntsman, David, Kohl, Bernard, Le, Xuan, Thorp, Richard, Andry, Chris, Duffy, Elizabeth R., Lyadov, Vladimir, Paklina, Oxana, Setdikova, Galiya, Shabunin, Alexey, Tavobilov, Mikhail, McPherson, Christopher, Warnick, Ronald, Berkowitz, Ross, Cramer, Daniel, Feltmate, Colleen, Horowitz, Neil, Kibel, Adam, Muto, Michael, Raut, Chandrajit P., Malykh, Andrei, Barnholtz-Sloan, Jill S., Barrett, Wendi, Devine, Karen, Fulop, Jordonna, Ostrom, Quinn T., Shimmel, Kristen, Wolinsky, Yingli, Sloan, Andrew E., De Rose, Agostino, Giuliante. 2018. Machine learning detects pan-cancer Ras pathway activation in the cancer genome atlas. *Cell Reports*, **23**(1), 172–180.e3.

Weigel, Ronald J, & Mcdougall, I Ross. 2006. The role of radioactive iodine in the treatment of well-differentiated thyroid cancer. *Surgical Oncology Clinics of North America*, **15**, 625–638.

Wilczynski, B., & Dojer, N. 2009. BNFinder: exact and efficient method for learning Bayesian networks. *Bioinformatics*, **25**(2), 286–287.

Wittmann, Dominik M, Krumsiek, Jan, Saez-Rodriguez, Julio, Lauffenburger, Douglas a, Klamt, Steffen, & Theis, Fabian J. 2009. Transforming Boolean models to continuous models: methodology and application to T-cell receptor signaling. *BMC Systems Biology*, **3**(98).

Wullschleger, Stephan, Loewith, Robbie, & Hall, Michael N. 2006. TOR signaling in growth and metabolism. *Cell*, **124**(3), 471–484.

Xing, Linlin, Guo, Maozu, Liu, Xiaoyan, Wang, Chunyu, Wang, Lei, & Zhang, Yin. 2017. An improved Bayesian network method for reconstructing gene regulatory network based on candidate auto selection. *BMC Genomics*, **18**(Suppl 9), 17–30.

Yongping, Shao, Le, Kaitlyn, Cheng, Hanyin, & Aplin, Andrew E. 2016. NFκB-regulation of c-FLIP promotes TNFα-mediated RAF inhibitor resistance in melanoma. *Journal of Investigative Dermatology*, **135**(7), 1839–1848.

Yu, Le, Marshall, S, Forster, T, & Ghazal, P. 2006. Modelling of macrophage gene expression in the interferon pathway. *Pages 45–46 of: Genomic Signal Processing and Statistics, 2006. GENSIPS2006.*

Yuan, Ming, & Lin, Yi. 2006. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society Series B*, **68**(1), 49–67.

Zadeh, L. A. 1965. Fuzzy sets. *Information and Control*, **8**, 338–353.

Zang, Fenglin, Wei, Xiyin, Leng, Xue, Yu, Man, & Sun, Baocun. 2014. C-FLIP(L) contributes to TRAIL resistance in HER2-positive breast cancer. *Biochemical and Biophysical Research Communications*, **450**(1), 267–273.

Zhang, Lin, Baladandayuthapani, Veerabhadran, Mallick, Bani K., Manyam, Ganiraju C., Thompson, Patricia A., Bondy, Melissa L., & Do, Kim-Anh. 2014. Bayesian hierarchical structured variable selection methods with application to molecular inversion probe studies in breast cancer. *Journal of the Royal Statistical Society Series C*, **63**(4), 595–620.

Zhao, Peng, Rocha, Guilherme, & Yu, Bin. 2009. The composite absolute penalties family for grouped and hierarchical variable selection. *Annals of Statistics*, **37**(6 A), 3468–3497.

Zhou, S F, Di, Y M, Chan, E, Du, Y M, Chow, V D W, Xue, C C L, Lai, X S, Wang, J C, Li, C G, Tian, M, & Duan, W. 2008. Clinical pharmacogenetics and potential application in personalized medicine. *Current Drug Metabolism*, **9**(8), 738–784.

Zou, Hui, & Hastie, Trevor. 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B*, **67**(2), 301–320.

Zou, Min, & Conzen, Suzanne D. 2005. A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. *Bioinformatics*, **21**(1), 71–79.