LSF Research Working Paper Series

N°. 10-09

Date: September 2010

Title: An Optimal Control Approach to Portfolio Optimisation with Conditioning Information

Author(s)*: Boissaux M., Schiltz J.

Abstract: In the classical discrete-time mean-variance context, a method for portfolio optimisation using conditioning information was introduced in 2001 by Ferson and Siegel ([1]). The fact that there are many possible signals that could be used as conditioning information, and a number of empirical studies that suggest measurable relationships between signals and returns, causes this type of portfolio optimisation to be of practical as well as theoretical interest.

Ferson and Siegel obtain analytical formulae for the basic unconstrained portfolio optimisation problem. We show how the same problem, in the presence of a risk-free asset and given a single conditioning information time series, may be expressed as a general constrained infinite-horizon optimal control problem which encompasses the results in [1] as a special case. Variants of the problem not amenable to closed-form solutions can then be solved using standard numerical optimal control techniques.

We extend the standard finite-horizon optimal control sufficiency and necessity results of the Pontryagin Maximum Principle and the Mangasarian sufficiency theorem to the doubly-infinite horizon case required to cover our formulation in its greatest generality. As an application, we rephrase the previously unsolved constrained-weight variant of the problem in [1] using the optimal control framework and derive the specific necessary conditions applicable. Finally, we carry out simulations involving numerical solution of the resulting optimal control problem to assess the extent to which the use of conditioning information brings about practical improvements in the field of portfolio optimisation.

Keywords: Optimal Control, Portfolio Optimization

JEL Classification: C02, C61, C65, G11

***Corresponding Author's Address:** Tel. : +352 46 66 44 5716; Fax : 352 46 66 44 6835.
E-mail address: marc.boissaux@uni.lu, jang.schiltz@uni.lu

An Optimal Control Approach to Portfolio Optimisation with Conditioning Information

Marc Boissaux
marc.boissaux@uni.lu

Jang Schiltz
jang.schiltz@uni.lu

Luxembourg School of Finance
University of Luxembourg, Luxembourg

This version: March 28, 2011
Original version: September 27, 2010

Abstract

In the classical discrete-time mean-variance context, a method for portfolio optimisation using conditioning information was introduced in 2001 by Ferson and Siegel ([1]). The fact that there are many possible signals that could be used as conditioning information, and a number of empirical studies that suggest measurable relationships between signals and returns, causes this type of portfolio optimisation to be of practical as well as theoretical interest.

Ferson and Siegel obtain analytical formulae for the basic unconstrained portfolio optimisation problem. We show how the same problem, in the presence of a risk-free asset and given a single conditioning information time series, may be expressed as a general constrained infinite-horizon optimal control problem which encompasses the results in [1] as a special case. Variants of the problem not amenable to closed-form solutions can then be solved using standard numerical optimal control techniques.

We extend the standard finite-horizon optimal control sufficiency and necessity results of the Pontryagin Maximum Principle and the Mangasarian sufficiency theorem to the doubly-infinite horizon case required to cover our formulation in its greatest generality. As an application, we rephrase the previously unsolved constrained-weight variant of the problem in [1] using the optimal control framework. Finally, we provide an illustration involving numerical solution of the resulting optimal control problem to assess the extent to which the use of conditioning information brings about practical improvements in the field of portfolio optimisation.

Keywords: Optimal Control, Portfolio Optimization
JEL Codes: C02, C61, C65, G11

1 Introduction

The present section introduces the problem of conditioned mean-variance portfolio optimisation and provides a brief survey of the existing literature. A number of indicators that could serve as conditioning information are briefly referenced and the GRAI indicator used in the later numerical illustration is discussed in slightly greater depth. Finally, the plan for the remaining sections of the paper is set.

1.1 Problem history

Early discussions of the mean-variance portfolio optimisation problem with signalling were largely theoretical and based on asset pricing models such as proposed by Lucas ([2]), in which accumulating information known to the agents is part of the makeup of market-clearing equilibrium prices. This type of setup was discussed using functional analysis tools in a 1987 paper by Hansen and Richard ([3]), who distinguish in particular between the classical unconditional efficient frontier and the ("conditional") efficient frontier involving mean and variance estimates conditioned on the available market information. The authors then show that any return on the unconditional frontier will be located on the conditional frontier as well (that is, that any unconditionally efficient portfolio is conditionally efficient), but that the converse is not in general true. This result then suggests that, even in the presence of information, it is preferable to optimise the mean-variance tradeoff between *unconditional* moments as the resulting portfolios will be efficient in both domains.

The discussion in [3], although far-reaching, is given from an asset pricing point of view and does not yield concrete optimal portfolio weights for the type of problem examined. The paper assumes that a conditional efficient frontier can be built and focuses on its relationship to the classical frontier; it does not interest itself in how both frontiers could be built so as to take conditioning information into account.

That additional step was taken by Ferson and Siegel in [1]. The authors add a financial interpretation to the theoretical preference for continued use of unconditional moments shown in [3]: indeed this is placed within the frequently encountered context of an uninformed investor paying a portfolio manager, who may have access to additional information, to manage their investments. The uninformed investor will then gauge their manager's investment performance using only the basic (unconditioned) information available to them. In this constellation, it is advisable for the manager to ensure that their portfolio choices will be efficient in the classical unconditional manner as investors will otherwise regard them as inferior to portfolio managers who choose not to use conditioning information. This abets the theoretical conclusion of [3] and leads Ferson and Siegel to aim for a conditioned optimisation of unconditional portfolio moments.

The way in which [1] introduce conditioning information is very generic and the present paper will thus emulate it. The article exists within a discrete-time myopic investment world, such that only two time instants are considered - an initial time t, at which the investment choice is made, and a final time $t+1$, at which the investment returns are examined. The authors assume the existence of a vector of signals s; what makes these signals interesting to the portfolio manager is that there exists some measurable relationship $\mu(s)$ between the signal as observed at the initial time and the return as revealed at the final time.[1] The fundamental signal-return relationship is then

$$r_{t+1} = \mu(s_t) + \epsilon, \tag{1}$$

where the time indices on r and on s will be suppressed in what follows. Here ϵ is a noise term whose conditional mean given s is assumed to be zero. There is no specific requirement on the functional form of $\mu(s)$. Note that, with this model,

$$E[r|s] = \mu(s) + E[\epsilon|s] = \mu(s)$$

and that $\mu(s)$ is regarded as deterministic once s is known i.e. contributes nothing to the conditional variance of returns.

Ferson and Siegel then obtain expressions for unconditional portfolio mean and variance given the signal-return relation (1) holds, and prove that the optimal portfolio weights in the presence of n risky assets and a risk-free asset with return r_f equal

$$u'(s) = \frac{\mu_P - r_f}{E[(\mu(s) - r_f e)'\Lambda(s)(\mu(s) - r_f e)]}(\mu(s) - r_f e)'\Lambda(s), \tag{2}$$

where μ_P is the required expected unconditional portfolio return, e is an n-vector of ones and $\Lambda(s) = [(\mu(s) - r_f e)(\mu(s) - r_f e)' + \Sigma^2]^{-1}$: here Σ^2 (which may in its greatest generality be a function of s) is the conditional covariance matrix $E[\epsilon\epsilon'|s]$.

Although the proofs given in [1] are nonconstructive, the results can be obtained through a variational argument from first principles. Examples of such arguments, which use stochastic Lagrange multipliers as analysed in [5], [6], are given by Chiang in [7], which obtains closed-form expressions for optimal weights within benchmark tracking variants of the basic mean-variance optimisation problem considered in [1]. Chiang also reports the results of an empirical study in the currency markets, using pure macroeconomic conditioning information such as forward premia or depreciation rates, and obtains marked improvements with respect to the classical solutions; however, a short exercise covering US equity markets yields less convincing figures.

Zhou ([8]) applies the analysis from [1] and [7] to Grinold's fundamental law of active portfolio management ([9]) which, as is pointed out, implies a context in which managers attempt to maximise conditional value-added through use of the information to which they have access. [8] obtains an alternative value-maximising solution for unconditional value added and provides a numerical

[1] Note it is this lagged relationship, with the optimisation opportunities it entails, which constitutes the principal fundamental difference between the present setup and market factor models such as the APT ([4]).

illustration showing improvement with respect to the classical strategy derived in [9].

In an empirical paper, Basu et al. ([10]) condition on both business cycle predictors (such as treasury bill rate or credit spread) and pure investor sentiment indicators (obtained from a regular survey carried out by the University of Michigan) to optimise portfolios of trading strategies. Both market timing and momentum-based strategies are used and the results are tested, in particular, with respect to the October 1987 and post-Internet bubble stock market crashes. The authors find that the strategies do manage to signal position reductions in times of crisis, but that this is only the case if the full set of signals are used. A companion paper ([11]) carries out a similar study involving, in the role of portfolio assets, the S&P500 and US treasury indices, gold and the federal rate as a risk-free proxy. Conditioning information is provided by the VIX volatility index (see e.g. [12]) and futures data for each of the portfolio assets. Results are entirely compatible with those mentioned for [10].

Another study on optimisation of a portfolio of trading strategies is carried out by Luo et al. ([13]), who do not use the Ferson and Siegel result, but optimise expected quadratic utility, from which they subtract an additional term that introduces a penalty for investment gearing and is specified as a function of the signal. The resulting strategy, at the very least related to that of [1], is then applied to a portfolio of trading strategies in the FX markets with a signal obtained through PCA reduction of indicators taken from various asset markets. As a result, improved performance is again seen with respect to the benchmark, which is made up of an equally weighted basket of the strategies in the portfolio.

1.2 Possible indicators

There is intuitive appeal in the definition or construction of indices that may be seen to represent the attitude of investors in the widest sense. The above literature review gives a good indication of the scope of different indicators that could meaningfully be used in the role of the conditioning signal. Some research literature focusing exclusively on the elaboration of such indices also exists. The angle is then generally not one of portfolio optimisation, but a more economic one in which it is attempted (for its own sake) to separate the different but related factors of overall market risk levels, investor risk aversion and (more short-term) investor sentiment: see e.g. [14]. The resulting task of separating different types of psychological motivation of market participants is very difficult, especially as the terminology is not entirely unambiguous: for instance, risk aversion is considered constant through time in classical finance theory (see e.g. [15]). The paper [16] offers a useful survey of indices fitting within this research strand and tests various indices for their power in predicting financial crises.

A number of indices are formulated in ways that could perhaps best be labelled as ad hoc. These either directly consist of individual macroeconomic quantities chosen for an intuitive link they may have to a certain type of risk, or some aggregation of several such quantities through either an averaging scheme or an application of principal components analysis. Good examples of such indices are given by the PCA index used by Luo et al. in the study [13] cited above, as well as a PCA indicator proposed by Deutsche Bundesbank.

Theoretically motivated indices tend to belong to one of two strands. On the one hand, there are indices that use the implied option volatility quoted in the markets in some way. The predominant approach here is to extract the risk-neutral distribution of future returns from the quoted volatility surface and then to compare its expectation of extreme returns with that obtained econometrically from historical returns data, see [14] for a discussion of this type of index.

On the other hand and based on an earlier industry paper by Persaud, Kumar and Persaud ([17]) introduced a type of risk aversion index which was subsequently labelled as *global risk aversion index*, or GRAI. The intuition behind this is that it is possible to distinguish between price changes due to changes in "global" (overall market) risk and price changes due to changes in investor risk aversion, the first being directly proportional to the change in overall risk and the second, proportional to the risk of each individual asset. Accordingly, the rank correlation between (current) price changes and (previous) asset risk is expected to be significant when risk aversion has changed, but negligible when the root cause of the changes is a change in overall market risk.

The discussion by Misina (see [18], [19]) formalises the theoretical motivation given in the original paper by attempting to fit the developments within the framework of a CAPM market model. It finds that this is only possible when independence of returns is assumed, such that any cross-correlations between assets are eliminated. This is a significant assumption and [19] attempts to justify it by describing a data normalisation procedure.

Thus the theoretical solidity of the GRAI model may be imperfect. However, it is certainly the case that the elaboration of an unassailable theoretical foundation to justify the separation of market time series into factors subject to different psychological drivers is a very difficult, if not impossible project. Accordingly, the intuitive appeal behind GRAI has led to a number of researchers and financial organisations building their own versions of the index. In particular, a variant of GRAI which uses the coefficient of the regression between price changes and risk (as expressed using volatility) has become known as *risk aversion index* or RAI; [16] report that this variant was originally proposed by Wilmot et al. when establishing a risk appetite index for Crédit Suisse in 2004. The State Street investor confidence index (ICI) can also be seen as a type of RAI since it is theoretically constructed using a CAPM foundation similar to that referenced above; the main difference is that the ICI focuses on changes in investor holdings as well as in prices, and thus presupposes access to relevant data.

1.3 Plan

The remainder of this paper is structured as follows. Section 2 establishes the generic extensions to standard results in the contingent fields of the calculus of variations and optimal control that are needed to support the general, doubly infinite signal support version of the conditioned optimisation problem. Section 3 uses these results to replicate results previously obtained by Ferson and Siegel ([1]), and to obtain a new open-loop expression for the specific problem of constrained conditioned portfolio optimisation in the presence of a risk-free asset. Section **??** reports the setup of a numerical illustration that was carried out, and gives the results obtained. Section 4 concludes.

2 Theoretical background

The present section introduces notation and shows how the classical optimal control results of the Pontryagin Minimum Principle and the Mangasarian sufficiency theorem can be extended to a problem domain that is doubly infinite. For a textbook treatment of the calculus of variations and optimal control, see e.g. [20].

A typical optimal control problem on a fixed finite interval $I_S = [s^-, s^+]$ takes on the following format:

$$\begin{aligned} &\text{minimise} && J_{I_S}(x(s), u(s)) = \int_{s^-}^{s^+} L(x(s), u(s), s) ds && (3)\\ &\text{subject to} && \dot{x}(s) = f(x(s), u(s), s), && (4)\\ & && x(s^-) = x^-, x(s^+) = x^+, && (5)\\ &\text{and} && u(s) \in U \ \forall s \in I_S, && (6) \end{aligned}$$

where $x \in (\mathcal{C}[s^-, s^+])^m$, both $(L(x(s), u(s), s)$ and $f(x(s), u(s), s))$ are jointly continuous with respect to both $x(s)$ and $u(s)$ and the control constraint set $U \subseteq \mathbb{R}^n$ is convex (and constant in s). The functional $J_{I_S}(x, u)$ is often called the *cost function*. Note the cost function notation is parameterised by the interval on which the function applies: this will be important to develop the limit case.

One notion relevant to the type of control problem discussed in this paper is that of autonomousness:

Definition 2.1 *Consider the type of optimal control problem defined by* (3)*,* (4)*,* (5) *and* (6) *above. The problem is said to be* autonomous *if none of the state differential equation component functions* $f_i(x(s), u(s), s), i \in \{1, \ldots m\}$*, depend explicitly on* s*.*

Remark Clearly, any non-autonomous problem can be transformed into an autonomous one by defining an additional state $x_{m+1}(s)$ such that $x_{m+1}(s^-) = s^-$ and $\dot{x}_{m+1}(s) = 1$: thus $x_{m+1}(s) = s \ \forall s \in I_S$. This shows that problem autonomousness is largely a question of problem formulation, and that any results for autonomous problems apply to non-autonomous problems as well, as long as the additional transformation state $x_{m+1}(s)$ is kept in mind.

At this point it is assumed that the interval $I_S = [s^-, s^+]$ is closed with both s^- and s^+ finite, and that the problem as given above is not to be understood as the restriction of a similar problem set over an infinite interval of the independent variable. Then any pair of state variable and control variable $(x(s), u(s))$ which verifies the problem constraints (4), (5) and (6) is called an *admissible pair*. Clearly, the optimal pair $(x^*(s), u^*(s))$ which minimises (3) and solves the optimal control problem will be chosen from the set of admissible pairs for the given independent variable boundary points s^- and s^+. We denote this set by $\mathcal{A}_{I_S}$, where the definition of I_S determines whether the relevant problem domain is closed or not.

Typically, the independent variable s is taken to denote time, in which case one usually sets $s^- = 0$. It is necessary to introduce two new quantities before giving the classical formulation of necessary optimality conditions for the finite horizon case. First, associate one *costate* variable (sometimes called an *adjoint* variable) to each state variable in the problem. Each costate variable is a function of the independent variable s and noted $\lambda_i(s)$ for $i \in \{1, 2, \dots m\}$. The column vector whose components are the costates is noted $\lambda(s) = (\lambda_1(s), \dots \lambda_m(s))$ and called the *costate vector*. Costates are then used to define the *Hamiltonian function* as

$$\mathcal{H}(x(s), u(s), \lambda_0, \lambda(s), s) = \lambda_0 L(x(s), u(s), s) + \lambda(s) \cdot f(x(s), u(s), s),$$

where $\cdot$ denotes the $\mathbb{R}^m$-scalar product and λ_0 is a scalar constant.
It can be proved that the following holds:

Theorem 2.2 *[First-order (Euler-Lagrange) necessary conditions, finite horizon]*

Consider the previous problem (3), (4), (5) *and* (6), *with unconstrained controls such that $U = \mathbb{R}^n$ and problem domain $I_S = [s^-, s^+]$. Let $(x^*(s), u^*(s))$ be an admissible pair for this problem. If this pair minimises the cost functional in* (3), *there exists a constant λ_0^* and a continuously differentiable costate vector function $\lambda^*(s) = (\lambda_1^*(s), \dots \lambda_m^*(s)) \in (\mathcal{C}^1[s^-, s^+])^m$ such that the quadruple $(x^*(s), u^*(s), \lambda_0^*, \lambda^*(s))$ satisfies the system*

$$\dot{x}^*(s) = f(x^*(s), u^*(s), s) \text{ with } x^*(s^-) = x^- \text{ and } x^*(s^+) = x^+, \tag{7}$$

$$\dot{\lambda}^*(s) = -\mathcal{H}_x(x^*(s), u^*(s), \lambda_0^*, \lambda^*(s), s) \text{ and} \tag{8}$$

$$\mathcal{H}_u(x^*(s), u^*(s), \lambda_0^*, \lambda^*(s), s) = 0, \tag{9}$$

with $(\lambda_0^*, \lambda^*(s)) \neq (0, \cdots 0)$, for all $s \in I_S$.

Proof See e.g. [21]. The proof generalises that of the equivalent result in the calculus of variations in that *admissible variations* $\epsilon h(s)$, for which the problem conditions (4), (5) and (6) continue to be verified, are added to the control function, following which the cost functional *Gâteaux derivative*

$$\delta J_{I_S}(x(s), u(s)) = \lim_{\epsilon \to 0} \frac{d}{d\epsilon}\Big(J_{I_S}(x(s), u(s) + \epsilon h(s))\Big)$$

is calculated and set to zero. Clearly, in variational calculus, variations are applied to the state rather than the control; otherwise this *variational procedure* has been left unchanged. Given this context, the theorem conditions are often called the Euler-Lagrange equations for optimal control.

Finally, note that the condition on $(\lambda_0^*, \lambda^*(s)) \neq (0, \cdots 0)$ is not given in the classical variational theory but implied by the later statement of the Euler-Lagrange conditions for a doubly infinite horizon (theorem 2.15) as obtained from the corresponding Pontryagin result. ∎

Note the above are necessary but not sufficient conditions. They are also of first order and thus determine *stationary* candidate functionals only, without indicating wheter the functional concerned implies a minimum or a maximum of the objective function. In this context, the Legendre-Clebsch condition gives a second order result analogous to the finite-dimensional case.

Theorem 2.3 *[Second-order (Legendre-Clebsch) necessary conditions, finite horizon]*

Consider the previous problem (3), (4), (5) *and* (6). *Let* $(x^*(s), u^*(s))$ *be an admissible pair for this problem. If this pair minimises the cost functional in* (3)*, the matrix of second derivatives of the problem Hamiltonian with respect to the controls and at the optimal pair is positive semi-definite:*

$$z'\Big(\mathcal{H}_{uu}(x^*(s), u^*(s), \lambda_0^*, \lambda^*(s), s)\Big)z \geq 0,$$

for all $s \in [s^-, s^+]$ *and for all* $z \in \mathbb{R}^n$.

Proof Again, see e.g. [21]. ∎

Remark For $\lambda_0^* = 0$, the problem is called *singular*; for $\lambda_0^* = 1$, it is called *regular*. The singular case arises when the first-order stationarity condition is trivially satisfied by any admissible control on some nonempty subinterval $(\theta, \psi) \subseteq [s^-, s^+]$. This corresponds to the case where the matrix $\mathcal{H}_{uu}$ is singular over (θ, ψ), as can directly be proven by contradiction: note that this case is not excluded by the classical Legendre-Clebsch requirement of positive semidefiniteness given above. In this situation, the cost function $L(x(s), u(s), s)$ is irrelevant as far as necessary optimality conditions are concerned and hence it is possible to assume $\lambda_0^* = 0$. This situation implies the existence of *singular arcs* and can lead to complicated discussions: see [22] for a survey. The possibility of dealing with a singular problem cannot in general be evaluated a priori except in specific cases (e.g. if both cost function integrand and state equation are linear in the controls, see [20] or [22]) and this question thus needs to be resolved on a case-by-case basis. This is often possible by assuming that $\lambda_0^* = 0$ and then checking whether this necessarily entails $\lambda^*(s) = 0\ \forall s \in I_S$, in which case the problem is shown to be regular by contradiction with the necessity conditions.

In the applications found in the present paper, the independent variable s corresponds to an observed signal (rather than time) and the domain boundaries s^- and s^+ may in the general case correspond to $-\infty$ and $+\infty$ so as to generate expectation integrals over probability distributions whose support is typically the entire real line. This situation does not typically occur in time-based optimal control problems given that such problems are always initiated at a certain point in time, which can in each case be made to correspond to the time $t = 0$, without loss of generality, by appropriate definition of the time axis.

However, optimal control problems with a terminal point at infinite time (*infinite horizon* problems) have naturally arisen in various disciplines such as continuum mechanics [23], advertising expenditure models (for an overview, see e.g. [24], [25]) or economics, where an early consumption optimisation problem over an infinite horizon was introduced by Ramsey in 1928 [26]. The following will justify the extension of various necessity and sufficiency results linked with this infinite-horizon case to the *doubly infinite horizon* case which covers the type of problem under discussion to the greatest required level of generality.

To this end, initially extend the ideas of admissible pairs and optimality to the doubly infinite horizon case. First describe the corresponding version of the previous optimal control problem set by (3), (4), (5) and (6). Based on previous parameterised notation for the cost function J_{I_S}, where the interval I_S may be finite, infinite at one end or doubly infinite, the infinite horizon problem (P) is to

$$\text{optimise} \qquad J_{I_S}(x(s), u(s)) = \int_{s^-}^{s^+} L(x(s), u(s), s) ds \text{ as } s^- \to -\infty, s^+ \to +\infty \tag{10}$$

$$\text{subject to} \qquad \dot{x}(s) = f(x(s), u(s), s)\ \forall s \in I_S, \tag{11}$$

$$\lim_{s \to -\infty} x(s) = x^-, \lim_{s \to +\infty} x(s) = x^+, \tag{12}$$

$$\text{and} \qquad u(s) \in U\ \forall s \in I_S. \tag{13}$$

Remark 1) The above formulation only allows for the most straightforward type of boundary conditions: thus it is assumed that both state variable limits exist and that they verify equality relationships. Note that, for the doubly infinite case, existence of an initial boundary limit value and an equality condition for that same boundary are both imperative if the problem is to be meaningful. However, more general terminal conditions could be applied to some of the states; typically these would include weaker restrictions that have to be verified for 'large' values of s only, or they might leave some variables free. See e.g. [27], [28]. Still, it is suggested that the portfolio type problems discussed in this paper do not require conditions of this type. Indeed, any portfolio optimisation will involve either minimising an undesirable (risk metric) quantity (such as variance (e.g. [29]), kurtosis (e.g. [30])) subject to another quantity (typically desirable, such as expected return (e.g. [29]) or skewness (e.g. [31]) being held fixed or vice versa. In each case, a problem setting of this type necessitates the existence of a limit terminal condition, verified either as an equality or possibly as an inequality (for instance, one could require expected return to reach at least a minimum level): hence all more general terminal condition types are discarded for the purpose of the present work. Inequality conditions are not considered as they will not change the problem with respect to equality conditions for any plausibly shaped efficient frontier.

2) Note that the definition of admissible pairs, given above for the finite problem domain case, directly generalises to the infinite domain.

In the discussion of the infinite-horizon problem which follows, the above version of the control problem will be used as a referential point and referred to as the *generic* problem, (P). This means that, in the doubly infinite-horizon context, both finite-horizon problems and problems of the singly infinite-horizon type (which initiate at a finite point θ) are regarded as restrictions of the general problem given above. That perception slightly modifies these problems with respect to their standalone versions since boundary conditions at finite points no longer hold *ex ante*, but only once an optimal state is given. Accordingly the classical definition of admissible pairs has to be adapted so as to eliminate the requirement for boundary conditions at finite points:

Definition 2.4 *Consider the restriction of the generic optimal control problem (P) given by* (10)*,* (11)*,* (12) *and* (13) *to any interval* $I \subseteq \mathbb{R}$*. Let* $(x(s), u(s))$ *be any pair of state and control variables defined on* I*, with* $x(s)$ *absolutely continuous on* I *and* $u(s)$ *measurable on* I*. If the pair* $(x(s), u(s))$ *verifies both the differential equation* (11) *and the condition that* $u(s) \in U \ \forall s \in I$*, it is called an* admissible pair for the I-restricted problem *if* $\lim_{s\to-\infty} x(s) = x^-$ *in case* I *has no finite lower bound, and* $\lim_{s\to+\infty} x(s) = x^+$ *in case* I *has no finite upper bound. The set of admissible pairs for the* I*-restricted problem is noted* $\mathcal{A}^r_I$*.*

Remark 1) Note that admissible pairs $(x(s), u(s))$ for the full problem (10), (11), (12) and (13) continue to correspond to the original definition and need to verify the problem's initial and terminal limit conditions given in (12). Thus the set of admissible pairs in this case is indifferently noted as either $\mathcal{A}_{(-\infty,+\infty)}$ or $\mathcal{A}^r_{(-\infty,+\infty)}$.

2) As defined, $(x(s), u(s)) \in \mathcal{A}_{(-\infty,+\infty)}$ implies that both $(x(s), u(s)) \in \mathcal{A}^r_{(-\infty,\theta]}$ and $(x(s), u(s)) \in \mathcal{A}^r_{[\theta,+\infty)}$ $\forall\theta \in \mathbb{R}$ but not the converse. This is because admissibility for the $(-\infty, \theta]$- and the $[\theta, +\infty)$-restricted problems does not require the relevant state variables to take any specific value at θ and so we may have $(x_1(s), u_1(s)) \in \mathcal{A}^r_{(-\infty,\theta]}$ and $(x_2(s), u_2(s)) \in \mathcal{A}^r_{[\theta,+\infty)}$ for which $x_1(\theta) \neq x_2(\theta)$.

The problem goal has been kept vague in (10) as it is necessary to be careful regarding its specification. Indeed it is not legitimate to simply adapt the finite-horizon goal (3) by letting the interval limits s^- and s^+ tend to $-\infty$ and $+\infty$ respectively as the improper integral given by the cost function limit

$$J_{(-\infty,+\infty)}(x(s),u(s)) = PV\int_{-\infty}^{+\infty} L(x(s),u(s),s)ds$$

may not be defined for all admissible pairs: here, as for all integrals with double infinite limits that appear in the present paper, the integral is valued by its Cauchy principal value - see e.g.[32] for details. Clearly, the transition to an infinite horizon problem is simplest if this is the case as the previous control problem can then directly be rephrased; however making this assumption does not lead to a sufficiently general theory. Analogously to [33] and [28], it is necessary to introduce different criteria for infinite horizon optimality over the full real line. To this end introduce the *cost function surplus*

$$\Delta(\theta,\psi) = \int_{\theta}^{\psi} L(x^*(s),u^*(s),s)ds - \int_{\theta}^{\psi} L(x(s),u(s),s)ds, \tag{14}$$

with $(x(s),u(s)) \in \mathcal{A}_{(-\infty,+\infty)}$ any admissible pair and $(x^*(s),u^*(s)) \in \mathcal{A}_{(-\infty,+\infty)}$ the candidate pair for optimality. Note that the aim is to look for candidate pairs which make the cost function surplus negative in various meaningful ways. In increasing order of strength, the following definition proposes two of the four optimality criteria given in [28], followed by the criterion applicable when the cost function limit $J_{(-\infty,+\infty)}(x(s),u(s))$ exists: this is called *strong optimality* in [33]. These definitions are now restated for reference.

Definition 2.5 *[Optimality criteria, singly infinite horizon]*
(i) Let $I_S = [\theta,+\infty)$ the signal support, and $\psi > \theta$ any real number. Let $I = [\theta,\psi] \subset \mathbb{R}$. An admissible pair $(x^*(s),u^*(s)) \in \mathcal{A}_{[\theta,+\infty)}$ is said to be finitely optimal *if every admissible pair $(x(s),u(s))$ for the I-restricted problem which satisfies $x(\psi) = x^*(\psi)$ causes the inequality*

$$J_I(x^*(s),u^*(s)) \leq J_I(x(s),u(s))$$

to be verified.
(ii) An admissible pair $(x^*(s),u^*(s)) \in \mathcal{A}_{[\theta,+\infty)}$ is said to be optimal according to the overtaking criterion (OT-optimal) *if there exists $\psi \geq \theta$ such that,*

$$\forall s^+ > \psi, \Delta(\theta,s^+) \leq 0.$$

(iii) An admissible pair $(x^*(s),u^*(s)) \in \mathcal{A}_{[\theta,+\infty)}$ is said to be strongly optimal *if the corresponding improper integral converges, i.e.*

$$\Big|\int_{\theta}^{+\infty} L(x(s),u(s),s)ds\Big| < +\infty$$

and, given any other admissible pair $(x(s),u(s)) \in \mathcal{A}_{[\theta,+\infty)}$,

$$\lim_{s^+\to+\infty} \Delta(\theta,s^+) \leq 0.$$

The above definitions adapted to the case of a doubly infinite horizon are as follows.

Definition 2.6 *[Optimality criteria, doubly infinite horizon] (i) Let $I = [\theta, \psi] \subset \mathbb{R}$, where $-\infty < \theta < \psi < +\infty$, be any finite real interval. An admissible pair $(x^*(s), u^*(s)) \in \mathcal{A}_{(-\infty,+\infty)}$ is said to be* finitely optimal *if every admissible pair $(x(s), u(s))$ for the I-restricted problem which satisfies both $x(\theta) = x^*(\theta)$ and $x(\psi) = x^*(\psi)$ causes the inequality*

$$J_I(x^*(s), u^*(s)) \leq J_I(x(s), u(s))$$

to be verified.

(ii) An admissible pair $(x^*(s), u^*(s)) \in \mathcal{A}_{(-\infty,+\infty)}$ is said to be optimal according to the overtaking criterion (OT-optimal) *if there exist $\theta, \psi \in \mathbb{R}$ such that $\theta \leq \psi$ and*

$$\forall s^- < \theta, s^+ > \psi, \Delta(s^-, s^+) \leq 0.$$

(iii) An admissible pair $(x^*(s), u^*(s)) \in \mathcal{A}_{(-\infty,+\infty)}$ is said to be strongly optimal *if the corresponding improper integral converges, i.e.*

$$\left| PV \int_{-\infty}^{+\infty} L(x(s), u(s), s) ds \right| < +\infty$$

and, given any other admissible pair $(x(s), u(s)) \in \mathcal{A}_{(-\infty,+\infty)}$,

$$\lim_{\theta \to +\infty} \Delta(-\theta, \theta) \leq 0.$$

Remark 1) As with all improper integrals with double limits, the limit on Δ above is interpreted using the Cauchy principal value convention.

2) Note that, in the finite optimality case, initial and terminal conditions for $x(s)$ are obtained from the values of $x^*(s)$ and not known independently. Finite optimality means that an optimal pair is not dominated on any finite segment of the real axis by any other pair admissible for the corresponding restriction of the problem: this suggests the Bellman optimality principle, which will be discussed in more detail below.

3) For strong optimality, the limit on $\Delta(-\theta, \theta)$ clearly exists if $J_{(-\infty,+\infty)}$ absolutely converges.

4) The two remaining types of optimality discussed in e.g. [34] involve the $\limsup$ and $\liminf$ of cost functions for which the limit may not exist. These offer a level of generality which is thought unnecessary for the type of portfolio problem examined in what follows.

The statement about the increasing order of strength implied by the above definitions of optimality can be formalised directly:

Proposition 2.7 *In the context of the above generic optimal control problem, consider an admissible pair $(x^*(s), u^*(s))$.*

If $(x^*(s), u^*(s))$ is strongly optimal for the given problem, it is also OT-optimal. Additionally, OT-optimality implies finite optimality.

Proof Consider any optimal control problem admissible pair set by (10), (11), (12) and (13), and consider a pair $(x^*(s), u^*(s))$, which is both admissible and strongly optimal for this problem.

Then $(x^*(s), u^*(s))$ is OT-optimal for the same problem, as can be seen directly from the definition of OT-optimality and the definition of the double limit.

Next, assume that $(x^*(s), u^*(s))$ is OT-optimal for the given problem. Assume for a contradiction that $(x^*(s), u^*(s))$ is not finitely optimal for the same problem. Then, for any $\theta, \psi \in \mathbb{R}$, there exists $I = [\overline{\theta}, \overline{\psi}] \subseteq \mathbb{R}$ such that $\overline{\theta} < \theta$, $\psi < \overline{\psi}$ and a pair $(\hat{x}(s), \hat{u}(s))$, admissible for the I-restricted problem corresponding to the original problem, such that

$$\int_{\overline{\theta}}^{\overline{\psi}} L(x^*(s), u^*(s), s)ds - \int_{\overline{\theta}}^{\overline{\psi}} L(\hat{x}(s), \hat{u}(s), s) = \epsilon > 0 \tag{15}$$

with the boundary conditions $\hat{x}(\overline{\theta}) = x^*(\overline{\theta})$ and $\hat{x}(\overline{\psi}) = x^*(\overline{\psi})$. Now define

$$(\tilde{x}(s), \tilde{u}(s)) = \begin{cases} (\hat{x}(s), \hat{u}(s)) \text{ for } s \in [\overline{\theta}, \overline{\psi}] \\ (x^*(s), u^*(s)) \text{ for } s \in \mathbb{R} \backslash [\overline{\theta}, \overline{\psi}]. \end{cases}$$

Then $(\tilde{x}(s), \tilde{u}(s))$ is an admissible pair: the boundary conditions on $\hat{x}(s)$ ensure that $\tilde{x}(s)$ is continuous on $\mathbb{R}$ and that the state equations (11) are verified on the whole of $\mathbb{R}$. Also, the boundary conditions on $x^*(s)$ remain valid and $\tilde{u}(s) \in U$ on the entire real axis. Thus, since it was assumed that $(x^*(s), u^*(s))$ is OT-optimal, $\exists \theta, \psi$ such that, $\forall s^- < \theta$ and $\forall s^+ > \psi$,

$$\int_{s^-}^{s^+} L(x^*(s), u^*(s), s)ds - \int_{s^-}^{s^+} L(\tilde{x}(s), \tilde{u}(s), s)ds \leq 0. \tag{16}$$

But $\overline{\theta}$ and $\overline{\psi}$ are finite so choose $s^- = \min(\overline{\theta}, \theta)$ and $s^+ = \max(\overline{\psi}, \psi)$ in the above and decompose the integrals to obtain

$$\begin{aligned} (16) \Leftrightarrow \int_{s^-}^{\overline{\theta}} L(x^*(s), u^*(s), s)ds + \int_{\overline{\theta}}^{\overline{\psi}} L(x^*(s), u^*(s), s)ds +& \\ \int_{\overline{\psi}}^{s^+} L(x^*(s), u^*(s), s)ds - \int_{s^-}^{\overline{\theta}} L(x^*(s), u^*(s), s)ds -& \\ \int_{\overline{\theta}}^{\overline{\psi}} L(\hat{x}(s), \hat{u}^*(s), s)ds - \int_{\overline{\psi}}^{s^+} L(x^*(s), u^*(s), s)ds \leq 0& \\ \Leftrightarrow \int_{\overline{\theta}}^{\overline{\psi}} L(x^*(s), u^*(s), s)ds - \int_{\overline{\theta}}^{\overline{\psi}} L(\hat{x}(s), \hat{u}^*(s), s)ds \leq 0.& \end{aligned}$$

This constitutes a contradiction with respect to (15). Hence it can be concluded that the initial assumption (of $(x^*(s), u^*(s))$ not being finitely optimal) must be rejected, and thus the required implication has been shown. ∎

Extending the classical singly-infinite problem class to a doubly-infinite horizon will involve partitioning the real axis so the doubly-infinite problem is decomposed into a pair of singly-infinite ones. To complete the set of necessary definitions, it is necessary to spell out the meaning of optimality for the doubly-infinite problem restricted to $(-\infty, \theta]$ for any $\theta \in \mathbb{R}$.

Definition 2.8 *Let* $(x^*(s), u^*(s)) \in \mathcal{A}^r_{(-\infty,\theta]}$ *be any admissible pair for the* $(-\infty, \theta]$*-restricted problem. Now introduce a change of variable* $\tau = -s$ *and define mirrored state and control variables*

$$\bar{x}^*(\tau) = x^*(s)$$
$$\bar{u}^*(\tau) = u^*(s),$$

such that

$$\begin{aligned}\dot{\bar{x}}^*(\tau) &= \frac{d\bar{x}^*(\tau)}{d\tau} \\ &= \frac{dx^*(s)}{d\tau} \\ &= \frac{dx^*(s)}{ds}\frac{ds}{d\tau} \\ &= -\dot{x}^*(s) \\ &= -f(\bar{x}^*(\tau), \bar{u}^*(\tau), -\tau)\end{aligned}$$

and

$$\bar{x}^*(-\theta) = x^*(\theta).$$

Clearly, $(\bar{x}^*(\tau), \bar{u}^*(\tau))$ *is admissible for the (P)-related problem* $(\bar{P})$ *defined over the interval* $\bar{I}_S = [-\theta, +\infty)$*:*

$$\text{optimise} \quad J_{\bar{I}_S}(\overline{x}(\tau), \overline{u}(\tau)) = \int_{-\theta}^{\tau^+} L(\overline{x}(\tau), \overline{u}(\tau), -\tau)d\tau \ \text{ as } \tau^+ \to +\infty \tag{17}$$

$$\text{subject to} \quad \dot{\overline{x}}(\tau) = -f(x(\tau), u(\tau), -\tau) \ \forall \tau \in \bar{I}_S, \tag{18}$$

$$\lim_{\tau \to +\infty} \overline{x}(\tau) = \lim_{s \to -\infty} x(s) = x^-, \ \overline{x}(-\theta) = x^*(\theta), \tag{19}$$

$$\text{and} \quad u(\tau) \in U \ \forall \tau \in \bar{I}_S. \tag{20}$$

Then the admissible pair $(x^*(s), u^*(s)) \in \mathcal{A}^r_{(-\infty,\theta]}$ *is said to be optimal if and only if the* $\bar{P}$*-admissible pair* $(\bar{x}^*(\tau), \bar{u}^*(\tau))$ *is.*

Note the independent argument used in non-autonomous problem expressions is $-\tau = s$ rather than τ: this is necessary for invariance of later results with respect to the orientation of the axis, as will become clear. The change of sign seen in the state equations reflects the fact that $x(s)$ has been mirrored with respect to the vertical axis and has thus been made into an even function for the purpose of the $\bar{I}_S$-restriction.

In the above the variational (Euler-Lagrange) set of necessity conditions adapted to optimal control (theorem 2.2) was presented and the problem context extended to the infinite domain case. The next subsection will discuss the necessity theorem which is perhaps the most natural for optimal control problems with control constraints, i.e. the Pontryagin Minimum Principle.

2.0.1 The Pontryagin Minimum Principle.

It has been seen that the Euler-Lagrange equations taken from the calculus of variations can be transposed to the more general case of optimal control problems. However, their suitability for optimal control problems that go beyond the sort of problem expressible using the calculus of variations is naturally limited. One type of optimal control problem which cannot generally be treated in the classical framework arises when the control constraints embedded in (6) are activated, i.e. U is defined as a proper subset of $\mathbb{R}^n$. Looking at the Euler-Lagrange equations, it is immediately clear that the condition on Hamiltonian derivative with respect to the controls being zero need no longer hold as the extremum may be reached on the boundary of the admissible control set. In connection with this, Pontryagin ([35]) points out that closed admissible control sets, in particular, are of great interest in practice, and that the pertinent problems often lead to optimal controls on the set boundary. The Pontryagin Minimum Principle (PMP) provides a generalisation of the classical conditions which can still be used with constrained control sets.

As for the Euler-Lagrange equations, the initial formulation of the PMP covered a problem formulated over a closed finite interval $I_S = [s^-, s^+]$.

Theorem 2.9 *[Pontryagin Minimum Principle, finite horizon]*

Consider the previously given finite horizon optimal control problem (3), (4), (5) *and* (6).

Let $(x^*(s), u^*(s))$ be an admissible pair for the above problem: thus $u^*(s)$ is a measurable control defined on $I_S = [s^-, s^+]$ with associated optimal path $x^*(s)$. If $(x^*(s), u^*(s))$ solves the above problem i.e. is optimal, there exist a constant λ_0^* and and a continuous and piecewise continuously differentiable costate vector function $\lambda^*(s) = (\lambda_1^*(s), \dots, \lambda_m^*(s))$ such that, $\forall s \in I_S$,

$$(\lambda_0^*, \lambda_1^*(s), \dots \lambda_m^*(s)) \neq (0, 0, \dots 0) \text{ and}$$

$u^*(s)$ minimises the Hamiltonian $\mathcal{H}(x^*(s), u(s), \lambda_0^*, \lambda^*(s), s)$ for $u(s) \in U$, i.e.

$$\mathcal{H}(x^*(s), u^*(s), \lambda_0^*, \lambda^*(s), s) \leq \mathcal{H}(x^*(s), u(s), \lambda_0^*, \lambda^*(s), s) \ \forall u(s) \in U, s \in I_S.$$

The costates $\lambda_i^*(s)$, for $i \in \{1, 2, \dots m\}$, verify

$$\dot{\lambda}_i^*(s) = -\frac{\partial \mathcal{H}}{\partial x_i}$$

except at any points of discontinuity for $u^*(s)$.
Finally, scaling always offers the possibility to normalise λ_0^* so either $\lambda_0^* = 1$ or $\lambda_0^* = 0$.

Proof See [35]. ∎

Remark 1) The given version of the PMP diverges from its original formulation in ([35]), where a cost function is to be minimised but the attached condition requires maximising the Hamiltonian. In each case, the 'maximum' or 'minimum' qualifier refers to the maximisation or minimisation of the Hamiltonian, rather than the intended type of optimisation. All four possible formulations of the PMP are equivalent; for instance, Seierstad [34] presents the maximisation formulation of the principle attached to a maximisation optimal control problem. In each case, it is simple to switch between formulations by noticing that maximisation of J_I corresponds to minimisation of $-J_I$, and that any attached discussion of sufficiency conditions (see below) may require switching convexity for concavity and vice versa.

2) For the problems under consideration, it has been suggested that fixed end point values will always be required. If this is not that case, the end point state values need to be optimised as well: this leads to an additional type of equation in the PMP, called a *transversality condition* and again directly generalising an equivalent condition found for purely variational problems with free end point. See [36] for a classic discussion of transversality conditions in the infinite-horizon context.

The original discussion of the Minimum Principle [35] only briefly mentions the infinite-horizon case. The authors conclude that the result not only holds over the half-line $[0, +\infty)$, but over the entire line $(-\infty, +\infty)$: this corresponds to the case required for the problems in the present discussion. Given the extent to which practical control problems have required the full half-line (as before, see e.g. [26], [24], [25]), more detailed proofs of the Principle for this specific case have been proposed. See in particular Halkin [37] for an early generic infinite-horizon proof and [38] or [39] for examples of proofs of stronger results for specific problem classes. The original infinite-horizon formulation of the PMP then appears as follows:

Theorem 2.10 *[Pontryagin Minimum Principle, real half-line to $+\infty$]*

Consider the previously given optimal control problem (3), (4), (5) *and* (6) *with a terminal horizon, s^+, at infinity, that is a problem domain of $I_S = [s^-, +\infty)$:*

$$
\begin{array}{ll}
\textit{optimise} & J(u) = \displaystyle\int_{s^-}^{+\infty} L(x(s), u(s), s)ds \\
\textit{subject to} & \dot{x}(s) = f(x(s), u(s), s), \\
 & x(s^-) = x^-, \displaystyle\lim_{s \to +\infty}(x(s)) = x^+ \\
\textit{and} & u(s) \in U \ \forall s \in I_S,
\end{array}
$$

where s^- is fixed and finite, L and f are jointly continuous in $x(s)$ and $u(s)$, and $u(s)$ is a piecewise continuous function defined on I_S.

Let $(u^*(s), x^*(s))$ be an admissible pair for the above problem: thus $u^*(s)$ is a measurable control defined on $[s^-, +\infty)$ with associated optimal path $x^*(s)$. If $(u^*(s), x^*(s))$ solves the above problem i.e. is optimal, there exist a constant λ_0^* and a continuous and piecewise continuously differentiable vector function $\lambda^*(s) = (\lambda_1^*(s), \dots, \lambda_m^*(s))$ such that, $\forall s \geq s^-$,

$$(\lambda_0^*, \lambda_1^*(s), \dots \lambda_m^*(s)) \neq (0, 0, \dots 0) \text{ and}$$

$u^*(s)$ minimises the Hamiltonian $\mathcal{H}(x^*(s), u(s), \lambda_0^*, \lambda^*(s), s)$ for $u(s) \in U$, i.e.

$$\mathcal{H}(x^*(s), u^*(s), \lambda_0^*, \lambda^*(s), s) \leq \mathcal{H}(x^*(s), u(s), \lambda_0^*, \lambda^*(s), s) \ \forall u(s) \in U, s \in I_S.$$

The costates $\lambda_i^*(s)$, for $i \in \{1, 2, \dots m\}$, verify

$$\dot{\lambda}_i^*(s) = -\frac{\partial \mathcal{H}}{\partial x_i}$$

except at any points of discontinuity for $u^*(s)$.
Finally, scaling always offers the possibility to normalise λ_0^* so either $\lambda_0^* = 1$ or $\lambda_0^* = 0$.

Proof See [37]: this treats the maximisation version of the problem but, from the preceding discussion, the two formulations can be reconciled by a change of sign in λ_0^*, which has no impact on the result as it involves a normalisation of λ_0^* in any case.

The proof is in two steps, of which the first proves that optimality involving a lim sup-based criterion implies what the present work calls finite optimality. The second step then involves a limit argument based on an infinite sequence of domain endpoints s_i which has since become somewhat standard: for instance, Aseev et al. refer to it as the method of approximations ([40]). As such, it may be observed that Halkin's proof really covers the case of finite optimality even though it is formulated in terms of the previously mentioned lim sup-based criterion. This is of significance as the chain of optimality implications shown in proposition 2.7 entails that the PMP then holds for a pair which is optimal according to any of the three criteria introduced in definition 2.6. ∎

In line with the previous argument, it is necessary to generalise the previous version (theorem 2.10) of the Pontryagin Minimum Principle to the full real axis $(-\infty, +\infty)$. This initially requires an extension of the well-known Bellman optimality principle ([41]) to the full real axis.

Theorem 2.11 *[Bellman Optimality Principle, real half-line to $+\infty$]*
Let $(x^*(s), u^*(s)) \in \mathcal{A}^r_{[\theta, +\infty)}$ be an admissible pair for the $[\theta, +\infty)$-restricted problem based on the generic problem, $\forall \theta \in \mathbb{R}$.

If $(x^*(s), u^*(s))$ is optimal according to any applicable optimality criterion, then let $x^*_{[\theta,\psi]}(s)$ be the restriction of $x^*(s)$ to the closed interval $[\theta, \psi]$ for any finite $\psi \geq \theta$, and $u^*_{[\theta,\psi]}(s)$ be the restriction of $u^*(s)$ to the same interval $[\theta, \psi]$.

The pair $(x^*_{[\theta,\psi]}(s), u^*_{[\theta,\psi]}(s))$ now minimises the restriction of the problem's cost function $J_{[\theta,\psi]}(x, u)$ on the set $\mathcal{A}^r_{[\theta,\psi)}$ of admissible pairs $(x(s), u(s))$ for which the initial and final conditions $x(\theta) = x^*(\theta)$ and $x(\psi) = x^*(\psi)$ hold.

Proof This proof can essentially be found in Carlson and Haurie [33]. Note that the authors examine the maximisation problem, such that inequalities need to be reversed. Also, a slightly more general definition of the infinite horizon problem as starting from any point $\theta \in \mathbb{R}$ (and not necessarily at 0) is used in the present work. Accordingly all occurrences of the problem origin in the original proof have to be replaced by θ, and the initial condition, in particular, is now a function of the optimal path (i.e. $x(\theta) = x^*(\theta)$) rather than exogenously specified for every possible θ.

See the later proof of the Principle in the doubly infinite case for an adapted version of the same argument. ∎

Remark Here the possible 'applicable optimality criteria' are those proposed for singly-infinite problem domains (see definition 2.5).

Based on these two versions, an extension of the Bellman Optimality Principle to the full real axis by the adaptation of existing proofs is straightforward enough. The proof strategy involves a decomposition of the optimality concept on $(-\infty, +\infty)$ by means of optimality on real half-lines as defined above.

Theorem 2.12 *[Bellman Optimality Principle, full real axis]*

Let $(x^*(s), u^*(s)) \in \mathcal{A}_{(-\infty,+\infty)}$ be an admissible pair for the generic optimal control problem set in (10), (11), (12) *and* (13).

If $(x^*(s), u^*(s))$ is optimal according to any of the optimality criteria given in definition 2.6, then let $x^*_{(-\infty,\theta]}(s)$ be the restriction of $x^*(s)$ to the interval $(-\infty, \theta]$ for any finite θ, and $u^*_{(-\infty,\theta]}(s)$ be the restriction of $u^*(s)$ to the same interval $(-\infty, \theta]$. Furthermore let $x^*_{[\theta,+\infty)}(s)$ be the restriction of $x^*(s)$ to the interval $[\theta, +\infty)$ and $u^*_{[\theta,+\infty)}(s)$ be the restriction of $u^*(s)$ to the same interval $[\theta, +\infty)$.

Then both of the pairs $(x^*_{(-\infty,\theta]}(s), u^*_{(-\infty,\theta]}(s)) \in \mathcal{A}^r_{(-\infty,\theta]}$ and $(x^*_{[\theta,+\infty)}(s), u^*_{[\theta,+\infty)}(s)) \in \mathcal{A}^r_{[\theta,+\infty)}$ are finitely optimal.

Proof Notice initially that $(x^*_{(-\infty,\theta]}(s), u^*_{(-\infty,\theta]}(s)) \in \mathcal{A}^r_{(-\infty,\theta]}$ and $(x^*_{[\theta,+\infty)}(s), u^*_{[\theta,+\infty)}(s)) \in \mathcal{A}^r_{[\theta,+\infty)}$ directly from the definition, and that the finite optimality mentioned involves the definition applicable to the singly-infinite horizon: see definition 2.5.

Assume for a contradiction that the above result is not true. Then, for one of the two half-lines involved at least, the pair $(x^*(s), u^*(s))$ is not finitely optimal. Say, without loss of generality, that $(x^*(s), u^*(s))$ is not finitely optimal on $[\theta, +\infty)$ for some $\theta \in \mathbb{R}$. Then, from 2.11, one has for some $\psi \geq \theta$ and for some $(\hat{x}(s), \hat{u}(s)) \in \mathcal{A}^r_{[\theta,\psi]}$ that

$$\int_\theta^\psi L(\hat{x}(s), \hat{u}(s), s)ds < \int_\theta^\psi L(x^*(s), u^*(s), s)ds$$

with $\hat{x}(\theta) = x^*(\theta)$ and $\hat{x}(\psi) = x^*(\psi)$. There thus exists some $\epsilon > 0$ such that

$$\int_\theta^\psi L(\hat{x}(s), \hat{u}(s), s)ds < \int_\theta^\psi L(x^*(s), u^*(s), s)ds - \epsilon. \tag{21}$$

Following a standard pattern, now assemble the pair

$$(\tilde{x}(s), \tilde{u}(s)) = \begin{cases} (\hat{x}(s), \hat{u}(s)) & \text{for } s \in [\theta, \psi] \\ (x^*(s), u^*(s)) & \text{for } s \in (-\infty, \theta) \cup (\psi, +\infty). \end{cases}$$

Now $\tilde{x}(s)$ is continuous on $(-\infty, +\infty)$ since both $\tilde{x}(\theta) = \hat{x}(\theta) = x^*(\theta)$ and $\tilde{x}(\psi) = \hat{x}(\psi) = x^*(\psi)$. Also $\tilde{u}(s) \in U$ a.e. on $(-\infty, +\infty)$ since both $\hat{u}(s) \in U$ and $u^*(s) \in U$ a.e. Finally the original admissible pair $(x^*(s), u^*(s))$ is retained in both limits: it can be concluded that $(\tilde{x}(s), \tilde{u}(s)) \in \mathcal{A}_{(-\infty,+\infty)}$.

But finite optimality of $(x^*(s), u^*(s))$ on $(-\infty, +\infty)$ has been assumed. Hence there exist $\hat{\theta} < \theta$ and $\hat{\psi} > \psi$ such that, using the previous value of ϵ,

$$\begin{aligned} \int_{\hat{\theta}}^{\hat{\psi}} L(x^*(s), u^*(s), s) ds &< \int_{\hat{\theta}}^{\hat{\psi}} L(\tilde{x}(s), \tilde{u}(s), s) ds - \frac{\epsilon}{2} \\ &= \int_{\hat{\theta}}^{\theta} L(x^*(s), u^*(s), s) ds + \int_{\theta}^{\psi} L(\hat{x}(s), \hat{u}(s), s) ds + \\ &\qquad \int_{\psi}^{\hat{\psi}} L(x^*(s), u^*(s), s) ds + \frac{\epsilon}{2} \\ &< \int_{\hat{\theta}}^{\hat{\psi}} L(x^*(s), u^*(s), s) ds - \frac{\epsilon}{2} \text{ by (21).} \end{aligned}$$

However, this is contradictory. Accordingly, it can be concluded that $(x^*(s), u^*(s))$ must be optimal on $[\theta, \infty)$ for any $\theta \in \mathbb{R}$. Using a similar argument, it can be shown that $(x^*(s), u^*(s))$ additionally has to be optimal on $(-\infty, \theta]$ at the same time. Hence the required result has been proven. ∎

Remark The given optimality requirements involve finite optimality both (through the chain of implications given in proposition 2.7) in the theorem precondition and (explicitly) in the result. This is sufficient for the present purpose, which is to prove the necessity result of the PMP.

The availability of theorems 2.10 and 2.12 then allows for an easy generalisation of the Pontryagin Minimum Principle to the doubly-infinite horizon case. The only open question to solve before proceeding to the proof is whether the relations obtained in the infinite-horizon case (theorem 2.10) are invariant for the restricted problem over the half-line to $-\infty$. This can be answered in the affirmative.

Lemma 2.13 *Let $(x^*(s), u^*(s)) \in \mathcal{A}^r_{(-\infty,\theta]}$ be an admissible pair for the restricted problem on $(-\infty, \theta]$. If $(x^*(s), u^*(s))$ is optimal, the relations given in Pontryagin's Minimum Principle for the positive half-line (theorem 2.10) remain true.*

Proof Assume that $(x^*(s), u^*(s)) \in \mathcal{A}^r_{(-\infty,\theta]}$ is optimal. Then, using the definition of optimality (Definition 2.8) and its attached notation, $(\bar{x}^*(\tau), \bar{u}^*(\tau))$ is both admissible on the linked problem $\bar{P}$ (given by (17), (18), (19) and (20)) and optimal. Application of theorem 2.10 thus justifies the existence of a nonnegative $\overline{\lambda}_0$ and a continuous and piecewise continuously differentiable vector function $\overline{\lambda}(\tau) = (\overline{\lambda}_1(\tau), \dots \overline{\lambda}_m(\tau))$ for which the following set of relations is verified:

$(i)(\overline{\lambda}_0, \overline{\lambda}_1(\tau), \dots \overline{\lambda}_m(\tau)) \neq (0, \dots 0)\ \forall \tau \in [-\theta, \infty)$;

$(ii)\Big(\frac{\partial \bar{\lambda}}{\partial \tau}\Big)(\tau) = \dot{\bar{\lambda}}(\tau) = -\frac{\partial}{\partial \bar{x}}\mathcal{H}(\bar{x}^*(\tau), \bar{u}^*(\tau), \bar{\lambda}_0, \bar{\lambda}(\tau), -\tau)$ a.e. on $[-\theta, \infty)$;

and

$(iii)\mathcal{H}(\bar{x}^*(\tau), \bar{u}^*(\tau), \bar{\lambda}_0, \bar{\lambda}(\tau), -\tau) \leq \mathcal{H}(\bar{x}^*(\tau), \bar{u}(\tau), \bar{\lambda}_0, \bar{\lambda}(\tau), -\tau)\ \forall \tau \in [-\theta, \infty), \forall \bar{u}(\tau) \in U.$

Now *define* $\lambda_i(s) = -\bar{\lambda}_i(\tau), i \in \{1, 2, \dots m\}$ and $\lambda_0 = \bar{\lambda}_0$: as will be seen, this asymmetry is required to absorb the negative sign of f in $(\bar{P})$. Additionally recall the relations $\bar{x}(\tau) = x(s)$ and $\bar{u}(\tau) = u(s)$. Finally note that the use of $-\tau$ in the Hamiltonian is justified as the problem being looked at is obtained from the restriction of a problem with domain $\mathbb{R}$ and so both $f(x(s), u(s), s)$ and $L(x(s), u(s), s)$ are defined on $\mathbb{R}$.
To first verify (ii), write out

$$\begin{aligned}\dot{\bar{\lambda}}(\tau) &= -\frac{\partial}{\partial \bar{x}}\Big[\bar{\lambda}_0 L(\bar{x}(\tau), \bar{u}(\tau), -\tau) + \bar{\lambda}(\tau) \cdot (-f(\bar{x}(\tau), \bar{u}(\tau), -\tau))\Big] \\ &= -\bar{\lambda}_0 \frac{\partial}{\partial \bar{x}} L(\bar{x}(\tau), \bar{u}(\tau), -\tau) + \bar{\lambda}(\tau) \cdot \frac{\partial}{\partial \bar{x}} f(\bar{x}(\tau), \bar{u}(\tau), -\tau).\end{aligned}$$

Then
$$\dot{\bar{\lambda}}(\tau) = \frac{d\bar{\lambda}(\tau)}{d\tau} = -\frac{d\lambda(s)}{d\tau} = -\frac{d\lambda(s)}{ds}\frac{ds}{d\tau} = \dot{\lambda}(s)$$

and

$$\begin{aligned}&-\bar{\lambda}_0 \frac{\partial}{\partial \bar{x}} L(\bar{x}(\tau), \bar{u}(\tau), -\tau) + \bar{\lambda}(\tau) \cdot \frac{\partial}{\partial \bar{x}} f(\bar{x}(\tau), \bar{u}(\tau), -\tau) = \\ &\qquad -\lambda_0 \frac{\partial}{\partial x} L(x(s), u(s), s) - \lambda(s) \cdot \frac{\partial f}{\partial x}(x(s), u(s), s)\end{aligned}$$

such that

$$\begin{aligned}\dot{\lambda}(s) &= -\lambda_0 \frac{\partial}{\partial x} L(x(s), u(s), s) - \lambda(s) \cdot \frac{\partial}{\partial x} f(x(s), u(s), s) \\ &= \frac{\partial}{\partial x}\mathcal{H}(x(s), u(s), -\lambda_0, -\lambda(s), s),\end{aligned}$$

that is

$$\dot{\lambda}(s) = -\frac{\partial}{\partial x}\mathcal{H}(x(s), u(s), \lambda_0, \lambda(s), s).$$

This shows that (ii) is invariant to the given change of variable given λ_0 is the Hamiltonian function constant and $\lambda(s)$ is the costate vector obtained from the Maximum Principle.
For (iii), write

$$\begin{aligned}
&\mathcal{H}(\bar{x}^*(\tau), \bar{u}^*(\tau), \bar{\lambda}_0, \bar{\lambda}(\tau), -\tau) \leq \mathcal{H}(\bar{x}^*(\tau), \bar{u}(\tau), \bar{\lambda}_0, \bar{\lambda}(\tau), -\tau)\\
&\Leftrightarrow \overline{\lambda}_0 L(\bar{x}^*(\tau), \bar{u}^*(\tau), -\tau) + \overline{\lambda}(\tau)\cdot(-f(\bar{x}^*(\tau), \bar{u}^*(\tau), -\tau)) \leq\\
&\qquad\qquad \overline{\lambda}_0 L(\bar{x}^*(\tau), \bar{u}(\tau), -\tau) + \overline{\lambda}(\tau)\cdot(-f(\bar{x}^*(\tau), \bar{u}(\tau), -\tau))\\
&\Leftrightarrow \lambda_0 L(x^*(s), u^*(s), s) + \lambda(s)\cdot f(x^*(s), u^*(s), s) \leq \lambda_0 L(x^*(s), u(s), s) + \lambda(s)\cdot f(x^*(s), u(s), s)\\
&\Leftrightarrow \mathcal{H}(x^*(s), u^*(s), \lambda_0, \lambda(s), s) \leq \mathcal{H}(x^*(s), u(s), \lambda_0, \lambda(s), s).
\end{aligned}$$

Thus (iii) is also invariant to the change of variable given. Clearly, since $\lambda_0 = \bar{\lambda}_0$ and $\lambda(s) = -\bar{\lambda}(\tau)$, (i) translates to

$$(\lambda_0, \lambda(s)) \neq (0, 0, \cdots, 0)$$

as well.
Thus the required invariance result has been shown. ∎

It is now straightforward to prove

Theorem 2.14 *[Pontryagin Minimum Principle, full real axis]*
Consider the previously given generic optimal control problem on a doubly-infinite horizon (10), (11), (12) *and* (13).
Let $(x^*(s), u^*(s))$ be an admissible pair for the above problem: thus $u^*(s)$ is a piecewise continuous control defined on $I_S = (-\infty, +\infty)$ with associated optimal path $x^*(s)$. If $(x^*(s), u^*(s))$ solves the above problem i.e. is optimal, there exist a constant λ_0^* and a continuous and piecewise continuously differentiable costate vector function $\lambda^*(s) = (\lambda_1^*(s), \ldots, \lambda_m^*(s))$ such that, $\forall s \in (-\infty, +\infty)$,

$$(\lambda_0^*, \lambda_1^*(s), \ldots \lambda_m^*(s)) \neq (0, 0, \ldots 0) \text{ and} \tag{22}$$

$u^*(s)$ minimises the Hamiltonian $\mathcal{H}(x^*(s), u(s), \lambda_0^*, \lambda^*(s), s)$ for $u(s) \in U$, i.e.

$$\mathcal{H}(x^*(s), u^*(s), \lambda_0^*, \lambda^*(s), s) \leq \mathcal{H}(x^*(s), u(s), \lambda_0^*, \lambda^*(s), s)\ \forall u(s) \in U, s \in I_S. \tag{23}$$

The costates $\lambda_i^*(s)$, for $i \in \{1, 2, \ldots m\}$, verify

$$\dot{\lambda}_i^*(s) = -\frac{\partial \mathcal{H}}{\partial x_i} \tag{24}$$

except at any points of discontinuity for $u^*(s)$.
Finally, scaling always offers the possibility to normalise λ_0^* so either $\lambda_0^* = 1$ or $\lambda_0^* = 0$.

Proof Let $(x^*(s), u^*(s)) \in \mathcal{A}_{(-\infty,+\infty)}$ be an optimal pair for the generic problem on $(-\infty, +\infty)$ according to any of the optimality modes given in definition 2.6. The previous Bellman Optimality Principle result (theorem 2.12) entails that $(x^*(s), u^*(s))$ remains finitely optimal for the problem restrictions to both $I_- = (-\infty, \theta]$ and $I_+ = [\theta, +\infty)$, $\forall\theta \in \mathbb{R}$. Trivially, the restrictions of $x^*(s)$ to I_- and I_+ continue to be continuous and differentiable, with $\lim_{s\to\theta^-} x^*(s) = \lim_{s\to\theta^+} x^*(s)$.

Then, from theorem 2.10, the Pontryagin conditions hold separately for both I_- and I_+, with separate Hamiltonian constants λ_0^{I-} and λ_0^{I+} as well as separate costate vectors $\lambda_{I_-} : (-\infty, \theta] \to \mathbb{R}$ and $\lambda_{I_+} : [\theta, +\infty) \to \mathbb{R}$. From the theorem, both $\lambda_{I_-}(s)$ and $\lambda_{I_+}(s)$ are continuous piecewise differentiable for all θ.

Also, lemma 2.13 indicates that the costate differential equation and Hamiltonian maximisation conditions (23) and (24) of the PMP are separately true on both I_- and I_+. It remains to show that $\lambda_0^{I-} = \lambda_0^{I+}$ and that $\lambda_{I_-}(\theta) = \lambda_{I_+}(\theta)$ so as to verify the required continuity conditions.

Successive application of the Bellman Optimality Principle on doubly-infinite and infinite horizons (theorems 2.12 and 2.11) allows for the definition of three separate Hamiltonians, each of which verifies the Minimum Principle:

$$\mathcal{H}_{I_-} = \lambda_0^{I-} L(x^*(s), u^*(s), s) + \lambda_{I_-}(s) \cdot f(x^*(s), u^*(s), s) \text{ on } I_- = (-\infty, \theta]$$
$$\mathcal{H}_{I_+} = \lambda_0^{I+} L(x^*(s), u^*(s), s) + \lambda_{I_+}(s) \cdot f(x^*(s), u^*(s), s) \text{ on } I_+ = [\theta, +\infty)$$
$$\mathcal{H}_{[\theta-\alpha,\theta+\alpha]} = \lambda_0^{[\theta-\alpha,\theta+\alpha]} L(x^*(s), u^*(s), s) + \lambda_{[\theta-\alpha,\theta+\alpha]}(s) \cdot f(x^*(s), u^*(s), s) \text{ on } [\theta-\alpha, \theta+\alpha]$$

Thus, on $s \in [\theta, \theta + \alpha]$, the latter two Hamiltonians give

$$\begin{cases} \lambda_0^{I+} L(x^*(s), u^*(s), s) + \lambda_{I_+}(s) \cdot f(x^*(s), u^*(s), s) \leq \\ \qquad \lambda_0^{I+} L(x^*(s), u(s), s) + \lambda_{I_+}(s) \cdot f(x^*(s), u(s), s) \\ \lambda_0^{[\theta-\alpha,\theta+\alpha]} L(x^*(s), u^*(s), s) + \lambda_{[\theta-\alpha,\theta+\alpha]}(s) \cdot f(x^*(s), u^*(s), s) \leq \\ \qquad \lambda_0^{[\theta-\alpha,\theta+\alpha]} L(x^*(s), u(s), s) + \lambda_{[\theta-\alpha,\theta+\alpha]}(s) \cdot f(x^*(s), u(s), s) \end{cases}$$

$\forall s \in [\theta, \theta + \alpha], \forall u \in U$ and $\forall L \in \mathcal{C}^1(-\infty, +\infty)$, $\forall f \in (\mathcal{C}^1(-\infty, +\infty))^m$. Equivalently,

$$\begin{cases} \lambda_0^{I+} \Big(L(x^*(s), u^*(s), s) - L(x^*(s), u(s), s) \Big) \leq \\ \qquad \lambda_{I_+}(s) \cdot \Big(f(x^*(s), u(s), s) - f(x^*(s), u^*(s), s) \Big) \\ \lambda_0^{[\theta-\alpha,\theta+\alpha]} \Big(L(x^*(s), u^*(s), s) - L(x^*(s), u(s), s) \Big) \leq \\ \qquad \lambda_{[\theta-\alpha,\theta+\alpha]}(s) \cdot \Big(f(x^*(s), u(s), s) - f(x^*(s), u^*(s), s) \Big). \end{cases}$$

Given this needs to hold $\forall L \in \mathcal{C}^1(-\infty, +\infty)$ and $\forall f \in (\mathcal{C}^1(-\infty, +\infty))^m$, the above inequality shows that, necessarily, $\lambda_0^{I+} = \lambda_0^{[\theta-\alpha,\theta+\alpha]}$ and $\lambda_{I_+}(s) = \lambda_{[\theta-\alpha,\theta+\alpha]}(s)$ $\forall s \in [\theta, \theta + \alpha]$.

By similarly exploiting the overlap between $[\theta - \alpha, \theta + \alpha]$ and I_-, it is clear that $\lambda_0^{I-} = \lambda_0^{[\theta-\alpha,\theta+\alpha]}$ and $\lambda_{I_-}(s) = \lambda_{[\theta-\alpha,\theta+\alpha]}(s)$ $\forall s \in [\theta - \alpha, \theta]$.

Hence $\lambda_0^{I_-} = \lambda_0^{I_+}$ and $\lambda_{I_-}(\theta) = \lambda_{I_+}(\theta)$. Thus a single scalar $\lambda_0^* = \lambda_0^{I_-} = \lambda_0^{I_+}$ applies to the doubly-infinite problem. Also, the set of resulting costates

$$\lambda^*(s) = \begin{cases} \lambda_{I_-}(s) \text{ for } s \in (-\infty, \theta] \\ \lambda_{I_+}(s) \text{ for } s \in [\theta, +\infty) \end{cases}$$

is continuous and piecewise differentiable on $(-\infty, +\infty)$.
Thus all the assertions of theorem 2.14 have been proven. ∎

2.0.2 The Euler-Lagrange equations in an optimal control context.

By carrying out a derivation based on variational arguments but starting from a generalised problem of the optimal control type (3), (4), (5) and (6), an optimal control equivalent of the Euler-Lagrange equations is obtained: see e.g. [20] or [21] for a variational proof of the below theorem in its finite-horizon version. However, having established theorem 2.14, the below conditions can be more directly obtained as a subset of the conditions given by the Pontryagin Principle.

Theorem 2.15 *[First-order (Euler-Lagrange) and second-order (Legendre-Clebsch) necessary conditions, full real axis]*

Consider the previously given generic optimal control problem on a doubly-infinite horizon (10), (11) and (12). The admissible control domain is taken to be unconstrained in (13): thus $U = \mathbb{R}^n$.

Let $(x^*(s), u^*(s))$ be an admissible pair for the above problem, with $u^*(s)$ additionally continuous: thus $u^*(s)$ is a continuous control defined on $(-\infty, +\infty)$ with associated optimal path $x^*(s)$, and the pair $(x^*(s), u^*(s))$ satisfies (11) and (12). If $(x^*(s), u^*(s))$ solves the above problem i.e. is optimal, there exists a constant λ_0^* and a continuously differentiable costate vector function $\lambda^*(s) = (\lambda_1^*(s), \ldots \lambda_m^*(s))$ such that the quadruple $(x^*(s), u^*(s), \lambda_0^*, \lambda^*(s))$ satisfies the system (11), (12),

$$\dot{\lambda}^*(s) = -\mathcal{H}_u(x(s), u(s), \lambda_0^*, \lambda^*(s), s) \text{ and} \tag{25}$$

$$\mathcal{H}_u(x(s), u(s), \lambda_0^*, \lambda^*(s), s) = 0 \tag{26}$$

and, $\forall s \in (-\infty, +\infty)$,

$$(\lambda_0^*, \lambda_1^*(s), \ldots \lambda_m^*(s)) \neq (0, 0, \ldots 0).$$

Additionally, the matrix of second derivatives of the problem Hamiltonian with respect to the controls at the optimal pair remains semi-definite positive:

$$z'\Big(\mathcal{H}_{uu}(x^*(s), u^*(s), \lambda_0^*, \lambda^*(s), s)\Big)z \geq 0,$$

for all $s \in \mathbb{R}$ and for all $z \in \mathbb{R}^n$.

Proof Existence of costates meeting the necessary requirements directly follows from the Pontryagin Principle (theorem 2.14). If $u^*(s)$ is continuous (as required), the Principle specifies the costate equations which are now verified on the whole of $\mathbb{R}$ since discontinuous controls are excluded. Thus, on $\mathbb{R}$, the $\lambda_i^*(s)$ are solutions of the ordinary differential equation (25) and thus continuous and piecewise continuously differentiable: see e.g. [42].

Additionally, the Pontryagin condition

$$\mathcal{H}(x^*(s), u^*(s), \lambda_0^*, \lambda^*(s), s) \leq \mathcal{H}(x^*(s), u(s), \lambda_0^*, \lambda^*(s), s)$$

entails that the first-order condition

$$\frac{\partial \mathcal{H}}{\partial u} = 0$$

is verified as long as the optimal control $u^*(s)$ is in the interior of the control region U i.e. its components remain finite $\forall s \in \mathbb{R}$, which is the case for any non-degenerate problem for which there exists at least one feasible solution.

The Pontryagin condition given above also entails verification of the second-order condition given the Hamiltonian is minimised by the optimal pair.

Finally, note that the condition on $(\lambda_0^*, \lambda^*(s))$ not being null is inherited from the formulation of the Principle and carries over to the finite-horizon Euler-Lagrange conditions given in theorem 2.2.

Thus all the theorem relations are verified on the whole of $\mathbb{R}$. ∎

2.0.3 The Mangasarian sufficiency theorem.

The previous sections have extended well-known necessity results for optimal control problems to the doubly-infinite context. These results justify the type of problem formulation used in this paper, and may lead to a number of candidate solutions for a particular problem's optimal pair. However, they do not allow for the conclusion that any of the candidate pairs do in fact correspond to the optimal pair. To draw such a conclusion with certainty, a sufficiency result is necessary. In the literature, two classical sufficiency results are given by the Mangasarian and Arrow sufficiency theorems (see [43] and [44]); in what follows, the Mangasarian result will be generalised to the current doubly-infinite context.

The Mangasarian conditions are related to those seen in the Pontryagin Principle. Sufficiency is achieved largely through the additional requirement that the Hamiltonian $\mathcal{H}(x(s), u(s), \lambda(s), s)$ be convex, or concave if the problem requires maximisation of the cost/benefit function. In the finite horizon case, their formulation is as follows:.

Theorem 2.16 *[Mangasarian sufficiency conditions, finite horizon]*

Consider the previously given finite horizon optimal control problem (3), (4), (5) *and* (6).

Let $(x^*(s), u^*(s))$ be an admissible pair for the above problem: thus $u^*(s)$ is a measurable control defined on $I_S = [s^-, s^+]$ with associated optimal path $x^*(s)$. Additionally suppose that $u^*(s) \in U$ $\forall s \in I_S$, with U a convex subset of $\mathbb{R}^n$, and *that the partial derivatives of both $L(x(s), u(s), s)$ and the $f_i(x(s), u(s), s), i \in \{1, 2, \dots n\}$, with respect to both x and u exist and are continuous.*

Suppose there exists a vector of costate functions $\lambda^*(s) = (\lambda_1^*(s), \dots, \lambda_m^*(s))$ $\forall s \in I_S$ and let λ_0^* be a constant. Introduce the alternative notation $\mathcal{H}^* = \mathcal{H}(x^*(s), u^*(s), \lambda_0^*, \lambda^*(s), s)$ for simplicity. Suppose that the following conditions are satisfied $\forall s \in I_S$ with $\lambda_0^* = 1$:

$$\dot{\lambda}_i^*(s) = -\frac{\partial \mathcal{H}^*}{\partial x_i} \ \forall i \in \{1, 2, \dots n\}$$

$$\mathcal{H}(x(s), u(s), \lambda_0^*, \lambda^*(s), s) \text{ is jointly convex in } (x(s), u(s))$$

$$\sum_{j=1}^{n} \frac{\partial \mathcal{H}^*}{\partial u_j} \Big(u_j^*(s) - u_j(s) \Big) \leq 0 \ \forall u(s) \in U.$$

Then the pair $(x^*(s), u^*(s))$ solves the problem (3), (4), (5) and (6).

If $\mathcal{H}(x(s), u(s), \lambda_0^*, \lambda^*(s), s)$ is strictly convex in $x(s)$ and in $u(s)$, the pair $(x^*(s), u^*(s))$ constitutes the unique solution to the problem.

Proof A proof built using the maximisation version of the problem is given in [28]; this can easily be adapted by a judicious swapping of inequalities, and of concavity for convexity. ∎

Moving to problem domains with a doubly infinite horizon is then simple given only boundary conditions involving limit equalities are considered.

Theorem 2.17 *[Mangasarian sufficiency conditions, doubly infinite horizon]*

Consider the previously given generic optimal control problem on a doubly-infinite horizon (10), (11), (12) and (13).

Let $(x^*(s), u^*(s))$ be an admissible pair for the above problem: thus $u^*(s)$ is a measurable control defined on $I_S = \mathbb{R}$ with associated optimal path $x^*(s)$. Additionally suppose that $u^*(s) \in U \ \forall s \in [s^-, s^+]$, with U a convex subset of $\mathbb{R}^n$, and that the partial derivatives of both $L(x(s), u(s), s)$ and the $f_i(x(s), u(s), s), i \in \{1, 2, \dots n\}$, with respect to both x and u exist and are continuous.

Suppose there exists a vector of costate functions $\lambda^*(s) = (\lambda_1^*(s), \dots, \lambda_m^*(s))$ $\forall s \in [s^-, s^+]$ and let λ_0^* be a constant. Suppose that the following conditions are satisfied $\forall s \in \mathbb{R}$ with $\lambda_0^* = 1$:

$$\dot{\lambda}_i^*(s) = -\frac{\partial \mathcal{H}^*}{\partial x_i} \ \forall i \in \{1, 2, \dots n\}$$

$$\mathcal{H}(x(s), u(s), 1, \lambda^*(s), s) \text{ is jointly convex in } (x(s), u(s))$$

$$\sum_{j=1}^{n} \frac{\partial \mathcal{H}^*}{\partial u_j} \Big(u_j^*(s) - u_j(s) \Big) \leq 0 \ \forall u(s) \in U.$$

Then the pair $(x^*(s), u^*(s))$ solves the problem (10), (11), (12) and (13), with

1. *finite optimality if, for any admissible $x(s)$ and any real interval $I = [\theta, \psi]$ with $-\infty < \theta < \psi < +\infty$,*
$$\Delta(\theta, \psi) \leq 0;$$

2. *OT-optimality if, for any admissible $x(s)$, $\exists \theta, \psi \in \mathbb{R}$ such that $\theta < \psi$,*
$$\lambda^*(s) \cdot (x(s) - x^*(s)) \geq 0 \ \forall s \in (-\infty, \theta] \ \text{and} \tag{27}$$
$$\lambda^*(s) \cdot (x(s) - x^*(s)) \leq 0 \ \forall s \in [\psi, +\infty). \tag{28}$$

3. *strong optimality if, for any admissible pair $(x(s), u(s))$, the improper integral*
$$PV \int_{-\infty}^{+} \infty L(x(s), u(s)) ds \tag{29}$$
converges and each costate component is bounded, i.e. $\exists A_i \ \forall i \in \{1, \ldots n\}$ such that
$$|\lambda_i^*(s)| \leq A_i \ \forall s \in I_S. \tag{30}$$

If $\mathcal{H}(x(s), u(s), \lambda_0^*, \lambda^*(s), s)$ is strictly convex in $x(s)$ and in $u(s)$, the pair $(x^*(s), u^*(s))$ constitutes the unique solution to the problem.

Proof First note that, for any $s^-, s^+ \in \mathbb{R}$ such that $s^- < s^+$,
$$\Delta(s^-, s^+) \leq \lambda^*(s^+) \cdot \Big(x(s^+) - x^*(s^+)\Big) - \lambda^*(s^-) \cdot \Big(x(s^-) - x^*(s^-)\Big) \tag{31}$$
from a proof for the finite-horizon case as given in [28].

The given condition for *finite optimality* directly corresponds to the definition.

For *OT-optimality*, assume given θ and $\psi \in \mathbb{R}$ such that the relations (27) and (28) are verified. Choose any $s^- \leq \theta$ and $s^+ \geq \psi$. Then (31) holds, with both terms on the right-hand side of the inequality negative by assumption. Accordingly $\exists \theta, \psi \in \mathbb{R}$ such that $\Delta(s^-, s^+) \leq 0$ for any $s^- < \theta$ and any $s^+ > \psi$: thus $(x^*(s), u^*(s))$ is OT-optimal.

For *strong optimality*, assume that the relations (29) and (30) are verified. Then the double limit
$$\lim_{s \to +\infty} \Delta(-s, s) \leq 0$$
exists and is finite as a difference of convergent integrals. It is then possible to take limits in (31); this leads to
$$\lim_{s \to +\infty} \Delta(-s, s) \leq \lim_{s \to +\infty} \Big(\lambda^*(s) \cdot \Big(x(s) - x^*(s)\Big)\Big) + \lim_{s \to +\infty} \Big(- \lambda^*(-s) \cdot \Big(x(-s) - x^*(-s)\Big)\Big).$$
Now let $A = \max(A_i : i \leq n)$: then the previous expression becomes
$$\lim_{s \to +\infty} \Delta(-s, s) \leq A \Big[\lim_{s \to +\infty} \Big(x(s) - x^*(s)\Big) + \lim_{s \to +\infty} \operatorname{sgn}\Big(x(-s) - x^*(-s)\Big) \Big(x(-s) - x^*(-s)\Big)\Big],$$

where sgn represents the signum function.

But both $\lim_{s\to+\infty}\Big(x(s)-x^*(s)\Big)$ and $\lim_{s\to+\infty}\Big(x(-s)-x^*(-s)\Big)$ exist and are equal to zero given the equality boundary condition specified for the generic control problem (10), (11), (12) and (13); also, the modulus of sgn is bounded. Hence

$$\lim_{s\to+\infty}\Delta(-s,s)\leq 0,$$

which shows that the pair $(x^*(s), u^*(s))$ is strongly optimal. ∎

3 The mean-variance problem with signalling in the presence of a risk-free asset

This section initially uses the tools developed in the previous section to confirm Ferson and Siegel's results in an optimal control context. It then introduces an empirical illustration carried out to examine the performance of the conditioned optimisation portfolio strategy when applied to real-world data. First the data set used and the parameters of the backtesting experiment are described. Then ex ante and ex post strategy results are described, and the overall performance with respect to classical (Markowitz) strategies commented on.

3.1 The unconstrained case

The presentation now proceeds to recover the results from [1] using the optimal control tools discussed in the previous section. The version of the problem examined here is that in which a risk-free asset, with a known return of r_f, is available in the market.

In order to rewrite this problem in an optimal control context, it is initially imperative to decide how to map state and control variables, respectively. The given problem involves minimisation of the expected unconditional variance subject to a constraint on the expected return. Since both of these quantities correspond to expectations, the variables used need to correspond to the integrands involved in the corresponding expectation integrals. The obvious choice for the controls are the portfolio weights, while the expected value constraint can most simply be implemented as a state variable $x_1(s)$. Thus it becomes necessary to optimise the Lagrange cost functional[2]

$$J_{I_S}(x(s),u(s)) = \frac{1}{2}\int_{s^-}^{s^+} u'(s)\Big[(\mu(s)-r_f e)(\mu(s)-r_f e)' + \Sigma_\epsilon^2\Big]u(s)p_S(s)ds \quad (32)$$

over the signal support I_S and given the state trajectory

$$\dot{x}_1(s) = u'(s)(\mu(s)-r_f e)p_S(s) \quad (33)$$

[2]Note a cost functional factor of 1/2 has been added to simplify the following calculations.

with initial and final constraints $x_1(s^-) = 0$ and $x_1(s^+) = \mu_p - r_f$. Here $p_S(s)$ is the signal probability density function (pdf).

With this notation, there is no additional terminal (Mayer) cost term. Notice that the signal has in this case taken on the role normally reserved for time in optimal control problems. With respect to the discussion in section 2, the most important result in connection with the problem in this subsection is that the core relations implied by the Euler-Lagrange equations do not change, whether the problem horizon is finite or infinite.

Remark 1) For most distributions, the support corresponds to the full real axis and thus the expectation integral end points s^- and s^+ are equal to $-\infty$ and $+\infty$, respectively, with initial and final constraints expressed as limits at infinity. For full theoretical generality and as discussed in section 2, it may be imprudent to focus purely on the minimisation of an improper integral which need not converge, and the desired optimisation paradigm should be chosen from the set given in definition 2.6. Even so, all objective function integrals that appear using the present formulation will involve the probability density factor $p_S(s)$. A defining factor of any pdf is that it integrates to unity over its support I_S: thus $\int_{s^-}^{s^+} p_S(s)ds = 1$ and so any non-pathological case will present negligible probability mass at extreme values of the signal. In that case, the standard Cauchy criterion for convergence of $J_{I_S}(x(s), u(s))$ will be verified and, as a consequence, strong optimality will be the applicable optimality criterion in any practical context. Also, in what follows, it will be assumed that $s^- = -\infty$, $s^+ = +\infty$ and boundary conditions are $\lim_{s\to-\infty} x_1(s) = 0$ and $\lim_{s\to+\infty} x_1(s) = \mu_P - r_f$, respectively: finite domain intervals then constitute a simpler specific case of the problem examined.

2) Clearly, probability density functions $p_S(s)$ are not in general convex in s. Note, however, that this affects neither the convexity requirements of Mangasarian-like theorems (such as theorem 2.17), which require convexity of the Hamiltonian in both the state and the control, nor those of finite-dimensional optimisers used to numerically solve the discretised problem.

Now define $\Lambda^{-1}(s) = (\mu(s) - r_f e)(\mu(s) - r_f e)' + \Sigma_\epsilon^2$ for simplicity: thus

$$J(u) = 1/2 \int_{s-}^{s+} u'(s)\Lambda^{-1}u(s)p_S(s)ds.$$

The following result will be helpful.

Proposition 3.1 *Let $\Lambda^{-1}(s)$ be the $n \times n$ matrix as defined above. Then $\Lambda^{-1}(s)$ is symmetric, positive definite, invertible and diagonalisable.*

Proof Immediate. ∎

Then and assuming the problem is regular (so $\lambda_0 = 1$ can be chosen), the problem Hamiltonian function $\mathcal{H}$ corresponds to

$$\begin{aligned}
\mathcal{H}(x(s), u(s), \lambda_0, \lambda(s), s) &= \mathcal{H}(u(s), 1, \lambda(s), s) \\
&= L(u(s), s) + \lambda(s) \cdot f(u(s), s) \\
&= \frac{1}{2} u'(s)\Lambda^{-1} u(s) p_S(s) + \lambda(s) u'(s)(\mu(s) - r_f e) p_S(s),
\end{aligned} \tag{34}$$

where the costate $\lambda(s)$ is scalar given the state variable is one-dimensional.

This leads to the Euler-Lagrange equations verified $\forall s^- \leq s \leq s^+$:

$$\begin{cases}
\dot{x} = \dot{x}_1 = \frac{\partial \mathcal{H}}{\partial \lambda} = u'(s)(\mu(s) - r_f e) p_S(s) \\
\dot{\lambda} = -\frac{\partial \mathcal{H}}{\partial x_1} = 0 \\
\frac{\partial \mathcal{H}}{\partial u} = \Big[\Lambda^{-1}(s) u(s) + \lambda(s)(\mu(s) - r_f e)\Big] p_S(s) = 0.
\end{cases}$$

The solution for the costate equation $\dot{\lambda} = 0$ is $\lambda(s) = \lambda = \text{constant } \forall s \in I_S$.

As suggested in section 2, it is necessary to exclude the possibility that the problem is singular: this is easy at this stage. So assume that $\lambda_0 = 0$. Then the Hamiltonian is

$$\mathcal{H}(x(s), u(s), 0, \lambda(s), s) = \lambda(s) u'(s)(\mu(s) - r_f e) p_S(s)$$

and the first-order optimality condition is

$$\frac{\partial \mathcal{H}}{\partial u} = \lambda(\mu(s) - r_f e) p_S(s) = 0,$$

where it is noted that λ is still scalar. Since $(\mu(s) - r_f e) p_S(s) \neq 0$ in general and outside of null sets, this implies that $\lambda(s) = \lambda = 0$, which violates the PMP requirement that $(\lambda_0, \lambda_1, \ldots \lambda_m) \neq (0, \ldots 0)$. Thus the problem is regular and it can be assumed that $\lambda_0 = 1$.

However, the scalar λ is still unknown. The local optimality condition on $\mathcal{H}_u$ now yields

$$\begin{aligned}
&\Lambda^{-1}(s) u(s) + \lambda(\mu(s) - r_f e) = 0 \\
&\Leftrightarrow u(s) = -\lambda \Lambda(s)(\mu(s) - r_f e),
\end{aligned} \tag{35}$$

where the transition to the second line uses the fact that Λ^{-1} is invertible from proposition 3.1. The proposition additionally shows that Λ^{-1} is positive definite and, as a consequence and given that $p_S(s) \geq 0 \ \forall s \in I_S$,

$$\mathcal{H}_{uu} = \Lambda^{-1}(s) p_S(s) \text{ is positive semidefinite } \forall s.$$

Accordingly, the candidate value for $u(s)$ found above is indeed a candidate for a possible portfolio variance *minimum* as per the Legendre-Clebsch second-order condition (theorem 2.15). Next, replace (35) in the problem state equation (33) to get

$$\dot{x}_1(s) = -\lambda(\mu(s) - r_f e)' \Lambda(s)(\mu(s) - r_f e) p_S(s) :$$

clearly, $\Lambda(s) = \Lambda'(s)$ given the symmetry of $\Lambda(s)$.

For clarity, define $\dot{f}(s) = (\mu(s) - r_f e)'\Lambda(\mu(s) - r_f e)p_S(s)$ such that $\dot{x}_1(s) = -\lambda\dot{f}(s)$. Note that the initial condition $\lim_{s\to s^-} x_1(s) = 0 \Leftrightarrow \lim_{s\to s^-} f(s) = 0$ is automatically verified by construction. The terminal condition can then be examined to obtain a value for the Lagrange multiplier λ:

$$\begin{aligned}
&\lim_{s\to s^+} x_1(s) = \mu_p - r_f = -\lambda \lim_{s\to s^+} f(s) \\
&\Leftrightarrow \lambda = -\frac{\mu_p - r_f}{\lim_{s\to s^+} f(s)}.
\end{aligned} \tag{36}$$

Given that I_S represents the support of the signal probability density p_S, $\lim_{s\to s^+} f(s) = E[\dot{f}(s)]$, leading to

$$\begin{aligned}
u^*(s) &= -\Lambda(s)\left(-\frac{\mu_p - r_f}{E[\dot{f}(s)]}\right)(\mu(s) - r_f e) \\
&= (\mu_p - r_f)\Lambda(s)\left(\frac{\mu(s) - r_f e}{E[(\mu(s) - r_f e)'\Lambda(s)(\mu(s) - r_f e)]}\right),
\end{aligned} \tag{37}$$

which corresponds to the original Ferson-Siegel result given in [1] and quoted as (2).

3.2 The weight constrained case

In practice, additional constraints are often added to the portfolio optimisation problem discussed in the previous section. In particular, negative portfolio weights correspond to short sales of the underlying assets. It is generally unrealistic to assume that short positions can be entered at no extra cost; this is especially true for naked short positions which correspond to a short sale of a security without borrowing the asset at the same time. Additionally, a significant proportion of investors either want to avoid the unlimited downside risk associated with shorting or, as in the case of pension funds, may be prohibited by regulators from entering uncovered investments: see e.g. [45] for the situation in Luxembourg. Clearly, ignoring these restrictions is even less realistic when dealing with less liquid assets such as investment funds.

Another realistic constraint on invested weights would limit them to a certain interval often centred on zero, such as $[-a, a]$ for $a \in \mathbb{R}$, so as to avoid entering excessively large positions in particular assets and thus incurring excessive undiversified risk.

For the traditional Markowitz problem, the introduction of portfolio weights constraints of any kind means a closed-form solution is no longer available, and a numerical algorithm such as a quadratic programming solver (e.g. [46]) or the Markowitz critical line algorithm detailed in e.g [47] has to be used. The situation in the presence of conditioning information is comparable, but the optimal control formulation of the problem introduced in the previous section allows for a numerical solution using the full set of available methods.

As before, the chosen problem involves optimisation of the unconditional moments as observed by an uninformed investor. Thus, the full optimal control problem has to be solved: this then leads to optimal weights as functions of the signal value. Only once this full solution has been obtained, ensuring that the unconditional expected variance is minimised and the expected value constraint is respected over the full signal space, may the portfolio manager apply their privileged information by using the observed signal value with the postulated signal-return relationship when evaluating a portfolio strategy.

Introducing portfolio weight constraints leaves the same optimisation problem as described by (32) and (33), except that the control constraint set $U \subseteq \mathbb{R}^n$ is now considered a proper convex subset of $\mathbb{R}^n$ for all values of s, subsuming the previous formulation.

Thus the investor aims to optimise, over the signal support domain I_S,

$$J_{I_S}(x(s), u(s)) = \frac{1}{2}\int_{s^-}^{s^+} u'(s)\Lambda^{-1}(s)u(s)p_S(s)ds \tag{38}$$

given the state trajectory

$$\dot{x}_1(s) = u'(s)(\mu(s) - r_f e)p_S(s)$$

with initial and final constraints $\lim_{s\to s^-} x_1(s) = 0$ and $\lim_{s\to s^+} x_1(s) = \mu_p - r_f$ and such that $u(s) \in U \;\forall\; s \in I_S$. Note that the control constraint set U remains convex even if different per-control maxima and minima are specified: thus quite general constraints of the form

$$\forall\; i \in \{1, \dots n\}, u_i^- \leq u_i(s) \leq u_i^+ \;\forall s \in I_S$$

are supported.

The resulting problem is an infinite-horizon variation of a classical optimal control problem often called a *minimum-energy problem*. We now move on to an empirical illustration of the optimal control solution technique.

3.3 Empirical illustration

3.3.1 Data set

The data set used collects eleven years of daily returns data chosen to represent a market relevant to domestic EUR investors. This market is made up of ten different funds[3] chosen across both equity and fixed income markets as well as Morningstar style classifications. Here, funds rather than individual assets

[3]AXA L Fund Equity Europe (AXA), Credit Suisse Bond Fund Management Company Luxembourg Small (CSU), Dekalux Midcap TF (DEK), Dexia Luxpart C (DEX), DWS Euro Bonds Long (DWS), Fidelity Funds Euro Bond Fund A Global Certificate (FIB), Fortis L Fund Equity Socially Responsible Europe (FOB), Invesco Pan European Small Cap Equity Fund Lux (INV), KBC Money Euro Medium Cap (KBC) and Morgan Stanley European Currencies High Yield Bond (MSE). In every case the reinvesting variant of the fund was picked.

were chosen given they provide a level of built-in diversification and a ten-asset market composed of funds is thus seen as more realistic as an equivalent equity market; additionally, interest-rate exposure is easily achieved through funds. All funds involved provide EUR return quotes and manage at most a proportion of 30% in non-EUR assets so as to prevent currency risk from representing too important a factor in the analysis. The data covers business days from January 1999 to February 2010: in total, each series contains 2891 returns.

As a proxy to an idealised risk-free asset, the EURIBOR interbank rate with a 1 week tenor was chosen. The intention at this stage is to provide a numerical illustration of the previous theory and, consequently, the only signal used was the Kumar and Persaud currency-based GRAI ([17]) as described in section 1 - checks were made using GRAI indices built using forward rates both with 1 month and 3 month tenors; differences were found to be negligible and the results reported apply to a GRAI built using 3 month tenors. The currency pairs used involved the currencies AUD, CAD, CZK, EUR, HKD, JPY, NZD, NKR, PLN, SGD, ZAR, SKR, SFR and GBP with respect to USD; although this was not seen as an ideal choice with respect to a EUR-centric investment universe, this selection (based on available data) is still thought a reasonable approach to representing risk as visible throughout the major global currency markets.

3.3.2 Approach

The numerical illustration was set up as a backtesting experiment. Given that the obvious strategy benchmark for the conditioned optimal portfolio is set by its Markowitz equivalent, both strategies were executed side by side over the data period. Portfolios were rebalanced each business day, with unconditional moment, signal-return relationship and signal kernel density estimates carried out over a rolling window of sixty business days i.e. three months. The resulting out-of-sample backtesting run thus starts three months after the first available data point.

The simplest case of a linear form $\mu(s) = a + bs$ was assumed for the lagged relationship (1); as a consequence, $\mu(s)$ was estimated by a simple linear regression. For each business (rebalancing) day, a 20 point efficient frontier was then built, requiring the solution of 20 separate optimisation problems. In the presence of a risk-free asset, problems were solved using a *direct* numerical method. Direct methods to solve optimal control problems have evolved from the original Euler discretisation of variational calculus problems and the adjective (as opposed to *indirect*) is used for algorithms which directly discretise the problem instead of working with relations extracted from the Euler-Lagrange or Pontryagin conditions. The particular transcription scheme used discretises both state variable and controls: such schemes tend to require the state differential equations to be verified at specified *collocation* points and are thus often known as direct collocation methods, see e.g. [50]. The implementation realises a piecewise constant discretisation for the control variables and a first-order linear polynomial discretisation for the state variables, as discussed, for instance, in the PhD thesis of von Stryk ([51]). Note that this discretisation mode is the simplest that is meaningful as the state variables need to verify the state differential equations. A direct solver algorithm of this type, especially for a reasonably fine discretisation grid, is computationally quite heavyweight; however, it is also entirely general and can be used on further problem variants not amenable to closed-form or quasi-closed-form solutions.

Following the computation of an efficient frontier, quadratic utilities were then computed for various risk aversion coefficients over the entire spectrum from 0 to 10.[4] Here a value of 0 corresponds to risk neutrality, which is not helpful in a portfolio optimisation context as a risk neutral investor would ignore the risk-return tradeoff and pursue the greatest possible expected return that market rules and the investor's finances would allow them to pursue. At the other end, an upper coefficient boundary of 10 is thought sufficient to cover the realistic range. For each risk aversion coefficient covered, the optimal utility point on the frontier was then computed and the optimal position, obtained by computing the optimal weights functional values for the observed value of the signal, entered accordingly. At the end of each investment interval, the actual investment returns on all portfolios, obtained given the observed historical returns, were evaluated.

3.3.3 Results

Figure 1 shows optimal weight functionals for the ten risky assets obtained as the solution to a typical problem. Similarly, figure 2 shows optimal weight functionals obtained for the same problem but with the weights constrained to exclude the possibility of short sales.

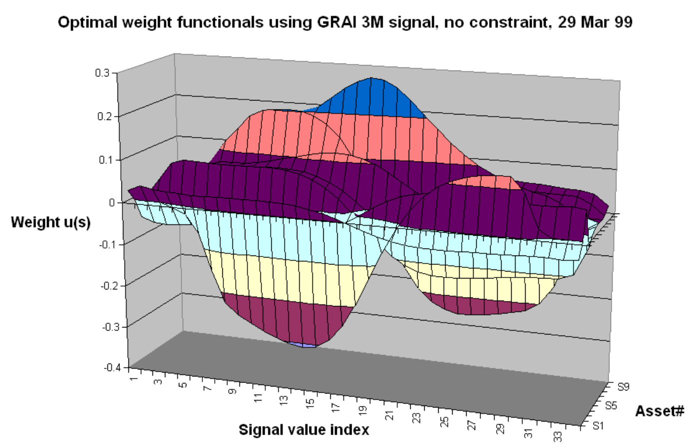

Figure 1: Typical unconstrained weight functionals.

[4]Note the use of quadratic utilities makes sense if investors are concerned only with the first two moments of returns, which any mean-variance optimisation process implicitly assumes.

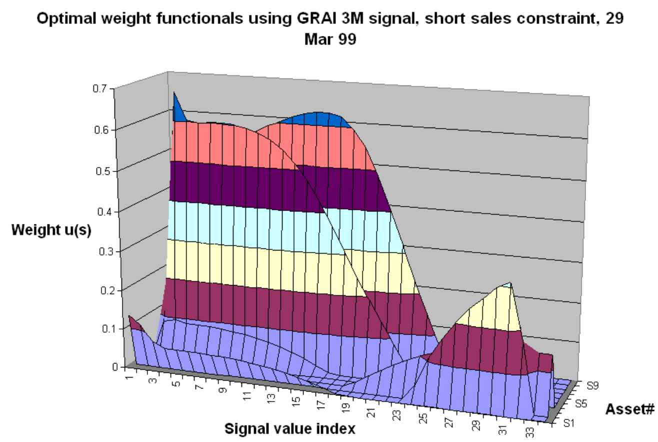

Figure 2: Typical weight functionals obtained with short-selling constraint.

It can be seen that the constrained solution does not simply constitute a truncated version of the unconstrained solution. This is the case even though the overall appearance of the functionals is not dissimilar - notice in particular that positive weight investments found in the absence of weights constraints tend to be accentuated when short sales are excluded. This observation also corresponds to theoretical discussions in the calculus of variations, an instance of which is [52].

To evaluate the effectiveness of the conditioned strategy with respect to classical Markowitz optimisation, the respective Sharpe ratios were compared, giving an a priori performance comparison accessible to investors without access to the conditioning information. The Sharpe ratio expression used was

$$SR_P = \frac{E[P] - r_f}{\sigma_P},$$

where $E[P]$ denotes the unconditional expected portfolio return, r_f the risk-free rate of return and σ_P the unconditional ex ante portfolio return standard deviation.

As regards portfolio weights, three different constraint types were set. The case where portfolio weights are unconstrained corresponds to the problem analysed by Ferson and Siegel. The remaining two constraint types correspond to a prohibition on short positions, frequently used by investment funds as discussed previously, and to a limitation on position size regardless of its sign, which is another plausible type of constraint designed to limit maximum portfolio risk.

A typical pair of unconditional efficient frontiers obtained is given in Figure 3. Note that it is legitimate to directly juxtapose efficient frontiers in this way given that unconditional moments are used in both cases. The improvement obtained through the use of conditioning information is substantial and this is, on average, the case over the entire data set, as Table 1 shows.

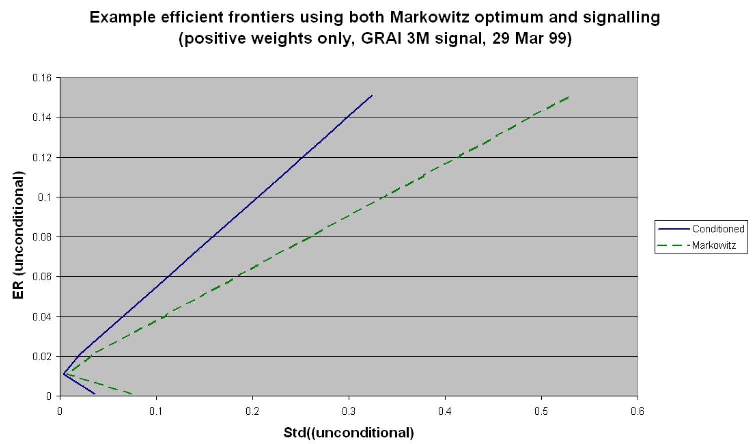

Figure 3: Example efficient frontiers for conditioned and Markowitz strategies with short-selling constraint.

	Unconstrained (Ferson-Siegel)	$[-1, 1]$	$[0, +\infty)$
Markowitz	0.563	0.524	0.326
Conditioned	0.721	0.694	0.426
Improvement	28.07%	32.38%	30.79%

Table 1: Sharpe ratios obtained for Markowitz and conditioned strategies for different weight constraints; relative Sharpe ratio improvement.

Clearly, the absolute performance seen degrades significantly for the short sales constraint given that this limits portfolio possibilities very significantly. The degradation seen for the limit in absolute position size is much smaller given that the allocated interval is sufficient to cover most of the efficient frontier regions computed. Importantly, it can be seen that the relative a priori improvement seen with respect to the Markowitz strategy is essentially constant regardless of the constraint mode chosen.

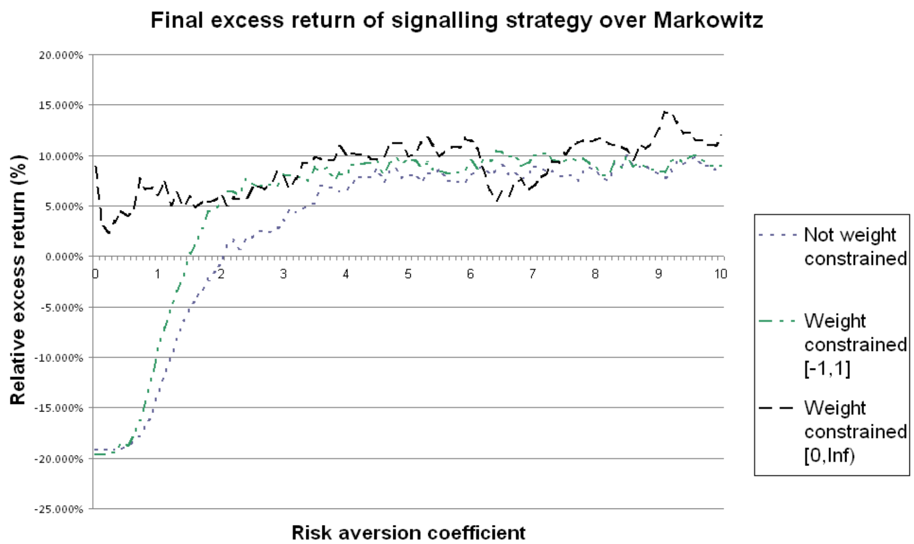

Figure 4: Relative additive excess returns of variously constrained investments with respect to their Markowitz equivalents.

Along with these ex ante considerations, the backtesting results were computed to examine the quite separate a posteriori question of the respective performances achieved. Figure 4 shows overall excess returns of conditioned optimisation strategies with respect to their Markowitz equivalents subject to the same portfolio weight constraints.

It is apparent that the improvements shown ex post (and detailed in table 2) by the conditioned strategies are somewhat disappointing with respect to the ex ante improvements seen in the respective Sharpe ratios. In particular, investors with strong risk appetites find that the classical Markowitz optimisation has outperformed conditioned optimisation for the given data set. For all other cases, including the important case of short sales prohibition over the entire range of risk aversion coefficients, a relative improvement of typically around 8% is obtained. However, considering the risk-return tradeoff provided in each case, it is still noticeable that the conditioned approach may be of interest in most cases. In particular, conditioned optimisation standard deviations are significantly lower than their Markowitz equivalents for low levels of risk aversion. This reflects the form of weight functionals optimal in the conditioned context, which decrease leverage for extreme signal values - see [1] for the unconstrained weight closed form function, whose properties appear to carry across to the numerical cases.

It is further noticeable that the standard deviation risk metric shows relatively higher risk levels for the conditioned strategy as risk aversion increases. Whether this is positive or negative for an individual investor depends on their attitude toward the (lower) absolute risk levels seen for those risk aversion categories. It is, however, true that, as for return improvements, the relative levels of risk seen are significantly more attractive from the conditioned viewpoint for the case where short sales are prohibited compared to the other two cases. Whether this is generally the case, or specific to the problem setup used, is a topic for further research.

4 Conclusion

The present paper has started by discussing the family of portfolio problems conditioned on signalling information. The corresponding setup is intuitively appealing and has generally yielded excellent results in previous empirical studies. The theory of Section 2 allows for the optimal control interpretation of conditioned optimisation problems. This is translated to the case of the mean-variance optimisation problem, with weights both constrained and unconstrained, in Section 3, which then gives a numerical illustration of the algorithm on real-world data. This shows that the approach gives measurable improvements in a somewhat realistic context involving any type of weights constraints used in practice, and that the set of available numerical methods for the resolution of optimal control problems can be applied. It is argued that the ex post improvements, although disappointing with respect to the ex ante increase in the Sharpe ratios, are nevertheless interesting for any investor accepting of the conventions of the Markowitz framework, and that larger improvements can be expected if either other signals more adapted to the market are used or a more accurate model of the signal to returns relationship is introduced. However, the purpose of the present application was merely to illustrate the functionality of the approach. As such, further investigations will be left to a future empirical paper.

We underline that the numerical problem solution described is largely generic when using a direct method. Thus the optimal control formulation given can be used to obtain solutions to a number of variations of the basic mean-variance problem in their conditioned guise: in this way it provides a framework for conditioned portfolio optimisation problems regardless of whether a closed-form solution for those problems exists. To our knowledge, such a framework was not available before, and the interest of this theoretical innovation will be explored in our future research.

λ	0	1	2	3	4	5	6	7	8	9	10
Markowitz return uncon.	264.66%	246.01%	195.63%	162.01%	137.08%	118.93%	106.36%	96.86%	89.80%	84.20%	79.22%
Ferson-Siegel return uncon.	213.81%	212.66%	194.70%	167.63%	146.18%	128.45%	114.88%	105.54%	97.38%	91.11%	85.92%
Relative return improvement uncon.	-19.21%	-13.56%	-0.48%	3.47%	6.64%	8.00%	8.01%	8.96%	8.44%	8.21%	8.45%
Markowitz std uncon.	0.552	0.488	0.375	0.301	0.249	0.199	0.169	0.145	0.126	0.113	0.105
Conditioned std uncon.	0.488	0.429	0.376	0.313	0.267	0.227	0.196	0.169	0.150	0.134	0.123
Std ratio uncon.	0.798	0.879	1.001	1.039	1.069	1.142	1.160	1.163	1.190	1.187	1.179
Markowitz return $[-1, 1]$	261.62%	230.95%	178.28%	147.08%	128.01%	113.15%	102.92%	94.39%	88.24%	82.88%	78.18%
Conditioned return $[-1, 1]$	210.14%	209.85%	187.54%	158.86%	138.22%	124.11%	112.60%	103.78%	96.03%	90.00%	85.14%
Relative return improvement $[-1, 1]$	-19.68%	-9.14%	5.19%	8.01%	7.98%	9.69%	9.41%	9.95%	8.82%	8.59%	8.90%
Markowitz std $[-1, 1]$	0.566	0.463	0.333	0.264	0.219	0.184	0.162	0.143	0.124	0.112	0.103
Conditioned std $[-1, 1]$	0.446	0.424	0.350	0.286	0.242	0.211	0.186	0.164	0.148	0.133	0.120
Std ratio $[-1, 1]$	0.788	0.916	1.059	1.083	1.103	1.146	1.149	1.151	1.188	1.188	1.164
Markowitz return $[0, +\infty)$	179.08%	172.96%	142.57%	120.54%	103.11%	89.53%	80.00%	75.18%	67.86%	63.61%	60.23%
Conditioned return $[0, +\infty)$	195.00%	183.54%	151.36%	130.10%	113.60%	98.41%	89.13%	80.49%	75.72%	71.63%	67.39%
Relative return improvement $[0, +\infty)$	8.89%	6.12%	6.17%	7.93%	10.17%	9.91%	11.42%	7.06%	11.57%	12.62%	11.88%
Markowitz std $[0, +\infty)$	0.646	0.395	0.274	0.220	0.186	0.154	0.128	0.110	0.098	0.087	0.079
Conditioned std $[0, +\infty)$	0.462	0.386	0.271	0.214	0.183	0.156	0.131	0.114	0.103	0.092	0.083
Std ratio $[0, +\infty)$	0.715	0.976	0.989	0.970	0.981	1.010	1.024	1.035	1.054	1.053	1.050

Table 2: Overall returns, relative improvements obtained using various strategies as a function of the quadratic risk aversion coefficient λ.

References

[1] Wayne E. Ferson and Andrew F. Siegel. The efficient use of conditioning information in portfolios. *The Journal of Finance*, 56(3):967–982, 2001. ISSN 00221082. URL `http://www.jstor.org/stable/222538`.

[2] Jr. Lucas, Robert E. Asset prices in an exchange economy. *Econometrica*, 46(6):pp. 1429–1445, 1978. ISSN 00129682. URL `http://www.jstor.org/stable/1913837`.

[3] Lars Peter Hansen and Scott F. Richard. The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica*, 55(3):587–613, 1987. ISSN 0012-9682. doi: 10.2307/1913601. URL `http://dx.doi.org/10.2307/1913601`.

[4] Stephen A. Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360, December 1976. URL `http://ideas.repec.org/a/eee/jetheo/v13y1976i3p341-360.html`.

[5] Harold J. Kushner. On stochastic extremum problems: Calculus. *J. Math. Anal. Appl.*, 10:354–367, 1965. ISSN 0022-247x.

[6] Harold J. Kushner. On the stochastic maximum principle: Fixed time of control. *J. Math. Anal. Appl.*, 11:78–92, 1965. ISSN 0022-247x.

[7] I-Hsuan Ethan Chiang. Modern Portfolio Management with Conditioning Information. *SSRN eLibrary*, 2008.

[8] Guofu Zhou. On the fundamental law of active portfolio management: How to make conditional investments unconditionally optimal. *Journal of Portfolio Management*, 35(1):12 – 21, 2008. ISSN 00954918.

[9] Richard C. Grinold. The fundamental law of active management. *Journal of Portfolio Management*, 15(3):30 – 37, 1989. ISSN 00954918.

[10] Devraj Basu, Chi-Hsiou Hung, Roel C. Oomen, and Alexander Stremme. When to Pick the Losers: Do Sentiment Indicators Improve Dynamic Asset Allocation? *SSRN eLibrary*, 2006.

[11] Devraj Basu, Roel C. Oomen, and Alexander Stremme. Exploiting the Informational Content of the Linkages between Spot and Derivatives Markets? *Warwick Business School Working Paper Series*, 2006.

[12] Robert E. Whaley. Understanding the vix. *Journal of Portfolio Management*, 35(3):98 – 105, 2009. ISSN 00954918.

[13] Jinhui Luo, Philip Saks, and Steve Satchell. Implementing risk appetite in the management of currency portfolios. *Journal of Asset Management*, 9:380–397(18), February 2009. URL `http://www.ingentaconnect.com/content/pal/jam/2009/00000009/00000006/art00003`.

[14] Prasanna Gai and Nicholas Vause. Measuring investors' risk appetite. *International Journal of Central Banking*, 2(1), March 2006. URL http://ideas.repec.org/a/ijc/ijcjou/y2006q1a5.html.

[15] Jonathan E. Ingersoll. *Theory of Financial Decision Making*. Studies in Financial Economics. Rowman Littlefield, April 1987. ISBN 0-8476-7359-6.

[16] Virginie Coudert and Mathieu Gex. Does risk aversion drive financial crises? testing the predictive power of empirical indicators. *Journal of Empirical Finance*, 15(2):167–184, March 2008. URL http://ideas.repec.org/a/eee/empfin/v15y2008i2p167-184.html.

[17] Manmohan S Kumar and Avinash Persaud. Pure contagion and investors' shifting risk appetite: Analytical issues and empirical evidence. *International Finance*, 5(3):401–36, Winter 2002. URL http://ideas.repec.org/a/bla/intfin/v5y2002i3p401-36.html.

[18] Miroslav Misina. What does the risk-appetite index measure? (03-23), 2003. URL http://ideas.repec.org/p/bca/bocawp/03-23.html.

[19] Miroslav Misina. Benchmark index of risk appetite. (06-16), 2006. URL http://ideas.repec.org/p/bca/bocawp/06-16.html.

[20] Michael Athans and Peter L. Falb. *Optimal control. An introduction to the theory and its applications*. McGraw-Hill Book Co., New York, 1966.

[21] Arthur E. Bryson, Jr. and Yu Chi Ho. *Applied optimal control*. Hemisphere Publishing Corp. Washington, D. C., 1975. Optimization, estimation, and control, Revised printing.

[22] Volker Michel. Singular optimal control - the state of the art. *Technical report, University of Kaiserslautern*, 169, 1996. URL http://kluedo.ub.uni-kl.de/volltexte/2000/609/.

[23] A. Leizarowitz and V. J. Mizel. One dimensional infinite-horizon variational problems arising in continuum mechanics. *Archive for Rational Mechanics and Analysis*, 106:161–194, June 1989. doi: 10.1007/BF00251430.

[24] Suresh P. Sethi. Dynamic optimal control models in advertising: A survey. *SIAM Review*, 19(4):685–725, 1977. ISSN 00361445. URL http://www.jstor.org/stable/2030250.

[25] Gustav Feichtinger, Richard F. Hartl, and Suresh P. Sethi. Dynamic optimal control models in advertising: Recent developments. *Management Science*, 40(2):195–226, 1994. ISSN 00251909. URL http://www.jstor.org/stable/2632762.

[26] F. P. Ramsey. A mathematical theory of saving. *The Economic Journal*, 38(152):543–559, 1928. ISSN 00130133. URL http://www.jstor.org/stable/2224098.

[27] Atle Seierstad and Knut Sydsaeter. Sufficient conditions in optimal control theory. *International Economic Review*, 18(2):367–391, 1977. ISSN 00206598. URL `http://www.jstor.org/stable/2525753`.

[28] Atle Seierstad and Knut Sydsæter. *Optimal control theory with economic applications*, volume 24 of *Advanced Textbooks in Economics*. North-Holland Publishing Co., Amsterdam, 1987. ISBN 0-444-87923-4.

[29] Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952. ISSN 00221082. URL `http://www.jstor.org/stable/2975974`.

[30] Gustavo Athayde and Renato Flores. Incorporating skewness and kurtosis in portfolio optimization: a multidimensional efficient set. In *Advances in Portfolio Construction and Implementation*, pages 243–257. Butterworth-Heinemann, Boston, EUA, 2003.

[31] Gustavo Athayde and Renato Flores. Finding a maximum skewness portfolio. (273), April 2001. URL `http://ideas.repec.org/p/sce/scecf1/273.html`.

[32] Vladimir A. Zorich. *Mathematical analysis. I.* Universitext. Springer-Verlag, Berlin, 2004. ISBN 3-540-40386-8. Translated from the 2002 fourth Russian edition by Roger Cooke.

[33] D. A. Carlson and A. Haurie. *Infinite horizon optimal control*, volume 290 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin, 1987. ISBN 3-540-17824-4. Theory and applications.

[34] Atle Seierstad. Necessary conditions for infinite horizon optimal control problems with state space constraints. *Proceedings of 25th Conference on Decision and Control, Athens, Greece*, 25:512–513, 1986.

[35] L S Pontryagin, V G Boltyanskii, R V Gamkrelidze, E F Mishchenko, and K N Trirogoff (translation). *The Mathematical Theory of Optimal Processes*. 1962.

[36] Philippe Michel. On the transversality condition in infinite horizon optimal problems. *Econometrica*, 50(4):975–985, 1982. ISSN 0012-9682. doi: 10.2307/1912772. URL `http://dx.doi.org/10.2307/1912772`.

[37] Hubert Halkin. Necessary conditions for optimal control problems with infinite horizons. *Econometrica*, 42(2):267–272, 1974. ISSN 00129682. URL `http://www.jstor.org/stable/1911976`.

[38] Sergei M. Aseev and Arkady V. Kryazhimskiy. The pontryagin maximum principle and transversality conditions for a class of optimal control problems with infinite time horizons. *SIAM J. Control Optim.*, 43(3):1094–1119, 2004. ISSN 0363-0129. doi: http://dx.doi.org/10.1137/S0363012903427518.

[39] Valeriano Antunes de Oliveira and Geraldo Nunes Silva. Optimality conditions for infinite horizon control problems with state constraints. *Nonlinear Analysis: Theory, Methods And Applications*, 71(12):e1788 – e1795, 2009. ISSN 0362-546X. doi: DOI: 10.1016/j.na.2009.02.052. URL http://www.sciencedirect.com/science/article/B6V0Y-4VM9K7S-1/2/74cf812ef71406e4bf4d6d63f

[40] Sergei Aseev, Arkadii Kryazhimskii, and Alexander Tarasyev. First order necessary optimality conditions for a class of infinite-horizon optimal control problems. In *PROCEEDINGS OF THE STEKLOV INSTITUTE OF MATHEMATICS*, pages 64–80, 2001.

[41] Richard Bellman. *Dynamic programming*. Dover Publications Inc., Mineola, NY, 2003. ISBN 0-486-42809-5. Reprint of the sixth (1972) edition, With an introduction by Eric V. Denardo.

[42] Carmen Chicone. *Ordinary differential equations with applications*, volume 34 of *Texts in Applied Mathematics*. Springer, New York, second edition, 2006. ISBN 978-0387-30769-5; 0-387-30769-9.

[43] O. L. Mangasarian. Sufficient conditions for the optimal control of nonlinear systems. *SIAM J. Control*, 4:139–152, 1966. ISSN 0363-0129.

[44] Kenneth J. Arrow. Applications of control theory to economic growth. In *Mathematics of the Decision Sciences, Part 2*, pages 85–119. American Mathematical Society, Providence, R.I., 1968.

[45] Commission de surveillance du secteur financier. Loi du 13 juillet 2005 relative aux institutions de retraite professionnelle sous forme de société d'épargne-pension. 2005.

[46] D. Goldfarb and A. Idnani. A numerically stable dual method for solving strictly convex quadratic programs. *Math. Programming*, 27 (1):1–33, 1983. ISSN 0025-5610. doi: 10.1007/BF02591962. URL http://dx.doi.org/10.1007/BF02591962.

[47] Harry M. Markowitz. *Mean-variance analysis in portfolio choice and capital markets*. Basil Blackwell, Oxford, 1987. ISBN 0-631-15381-0.

[48] Robert C. Merton. An analytic derivation of the efficient portfolio frontier. *Journal of Financial & Quantitative Analysis*, 7(4):1851 – 1872, 1972. ISSN 00221090.

[49] Jean-Luc Prigent. *Portfolio optimization and performance analysis*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2007. ISBN 978-1-58488-578-8; 1-58488-578-5.

[50] John T. Betts. *Practical methods for optimal control and estimation using nonlinear programming*, volume 19 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, first edition, 2001. ISBN 0-89871-488-5.

[51] O. von Stryk. *Numerische Lösung optimaler Steuerungsprobleme: Diskretisierung, Parameteroptimierung und Berechnung der adjungierten Variablen.* Number 441 in Fortschritt-Berichte VDI, Reihe 8: Meß-, Steuer- und Regelungstechnik. VDI Verlag, Düsseldorf, 1994.

[52] Morton I. Kamien and Nancy L. Schwartz. *Dynamic optimization*, volume 31 of *Advanced Textbooks in Economics*. North-Holland Publishing Co., Amsterdam, second edition, 1991. ISBN 0-444-01609-0. The calculus of variations and optimal control in economics and management.