

Duality and optimality conditions in stochastic optimization and mathematical finance

Sara Biagini* Teemu Pennanen† Ari-Pekka Perkkiö‡

April 25, 2015

Abstract

This article studies convex duality in stochastic optimization over finite discrete-time. The first part of the paper gives general conditions that yield explicit expressions for the dual objective in many applications in operations research and mathematical finance. The second part derives optimality conditions by combining general saddle-point conditions from convex duality with the dual representations obtained in the first part of the paper. Several applications to stochastic optimization and mathematical finance are given.

1 Introduction

Let (Ω, \mathcal{F}, P) be a complete probability space with a filtration $(\mathcal{F}_t)_{t=0}^T$ of complete sub σ -algebras of \mathcal{F} and consider the dynamic stochastic optimization problem

$$\text{minimize } Ef(x, u) := \int f(x(\omega), u(\omega), \omega) dP(\omega) \quad \text{over } x \in \mathcal{N} \quad (P_u)$$

parameterized by a measurable function $u \in L^0(\Omega, \mathcal{F}, P; \mathbb{R}^m)$. Here and in what follows,

$$\mathcal{N} := \{(x_t)_{t=0}^T \mid x_t \in L^0(\Omega, \mathcal{F}_t, P; \mathbb{R}^{n_t})\},$$

for given integers n_t and f is an extended real-valued $\mathcal{B}(\mathbb{R}^n \times \mathbb{R}^m) \otimes \mathcal{F}$ -measurable function on $\mathbb{R}^n \times \mathbb{R}^m \times \Omega$, where $n := n_0 + \dots + n_T$. The variable $x \in \mathcal{N}$ is interpreted as a decision strategy where x_t is the decision taken at time t . Throughout this paper, we define the expectation of a measurable function ϕ as $+\infty$ unless the positive part ϕ^+ is integrable (In particular, the sum of extended real numbers is defined as $+\infty$ if any of the terms equals $+\infty$). The function

*University of Pisa

†King's College London

‡Technische Universität Berlin. The author is grateful to the Einstein Foundation for the financial support.

Ef is thus well-defined extended real-valued function on $\mathcal{N} \times L^0(\Omega, \mathcal{F}, P; \mathbb{R}^m)$. We will assume throughout that the function $f(\cdot, \cdot, \omega)$ is *proper*, *lower semicontinuous* and *convex* for every $\omega \in \Omega$.

It was shown in [6] that, when applied to (P_u) , the conjugate duality framework of Rockafellar [14] allows for a unified treatment of many well known duality frameworks in stochastic optimization and mathematical finance. An important step in the analysis is to derive dual expressions for the optimal value function

$$\varphi(u) := \inf_{x \in \mathcal{N}} Ef(x, u)$$

over an appropriate subspace of L^0 . In this context, the absence of a duality gap is equivalent to the closedness of the value function. Pennanen and Perkkiö [9] and more recently Perkkiö [10] gave conditions that guarantee that φ is closed and that the optimum in (P_u) is attained for every $u \in L^0$. The given conditions provide far reaching generalizations of well-known no-arbitrage conditions used in financial mathematics.

The present paper makes two contributions to the duality theory for (P_u) . First, we extend the general duality framework of [6] by allowing more general dualizing parameters and by relaxing the time-separability property of the Lagrangian. We show that, under suitable conditions, the expression in [6, Theorem 2.2] is still valid in this extended setting. This also provides a correction to [6, Theorem 2.2] which omitted certain integrability conditions that are needed in general; see [7]. Second, we give optimality conditions for the optimal solutions of (P_u) . Again, we follow the general conjugate duality framework of [14] by specializing the saddle-point conditions to the present setting. The main difficulty here is that, in general, the space \mathcal{N} does not have a proper topological dual so we cannot write the generalized Karush-Kuhn-Tucker condition in terms of subgradients. Nevertheless, the dual representations obtained in the first part of the paper allow us to write the saddle-point conditions more explicitly in many interesting applications.

In the case of perfectly liquid financial markets, we recover well-known optimality conditions in terms of martingale measures. For Kabanov's currency market model with transaction costs [3], we obtain optimality conditions in terms of dual variables that extend the notion of a "consistent price system" to possibly nonconical market models. We treat problems of convex optimal control under the generalized framework of Bolza much as in [17]. Our formulation and its embedding in the conjugate duality framework of [14] is slightly different from that in [17], however, so direct comparisons are not possible. Our formulation is motivated by applications in mathematical finance. In particular, the optimality conditions for the currency market model are derived by specializing those obtained for the problem of Bolza.

2 Duality

From now on, we will assume that the parameter u belongs to a decomposable space $\mathcal{U} \subset L^0$ which is in separating duality with another decomposable space $\mathcal{Y} \subset L^0$ under the bilinear form

$$\langle u, y \rangle = E(u \cdot y).$$

Recall that \mathcal{U} is *decomposable* if

$$\mathbb{1}_A u + \mathbb{1}_{\Omega \setminus A} u' \in \mathcal{U}$$

whenever $A \in \mathcal{F}$, $u \in \mathcal{U}$ and $u' \in L^\infty$; see e.g. [15]. Examples of such dual pairs include the Lebesgue spaces $\mathcal{U} = L^p$ and $\mathcal{Y} = L^q$ and decomposable pairs of Orlicz spaces; see Section 3. The *conjugate* of $\varphi : \mathcal{U} \rightarrow \overline{\mathbb{R}}$ is the extended real-valued convex function on \mathcal{Y} defined by

$$\varphi^*(y) = \sup_{u \in \mathcal{U}} \{\langle u, y \rangle - \varphi(u)\}.$$

Here and in what follows, $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$. If φ is closed¹ with respect to the weak topology induced on \mathcal{U} by \mathcal{Y} , the biconjugate theorem (see e.g. [14, Theorem 5]) gives the *dual representation*

$$\varphi(u) = \sup_{y \in \mathcal{Y}} \{\langle u, y \rangle - \varphi^*(y)\}.$$

This simple identity is behind many well known duality relations in operations research and mathematical finance. It was shown in [9] that appropriate generalizations of the *no-arbitrage* condition from mathematical finance guarantee the closedness of φ . Recently, the conditions were extended in [10] to allow for more general objectives.

In general, it may be difficult to derive more explicit expressions for φ^* . Following [14], one can always write the conjugate as

$$\varphi^*(y) = - \inf_{x \in \mathcal{N}} L(x, y),$$

where the *Lagrangian* $L : \mathcal{N} \times \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ is defined by

$$L(x, y) = \inf_{u \in \mathcal{U}} \{Ef(x, u) - \langle u, y \rangle\}.$$

We will show that, under appropriate conditions, the second infimum in

$$\varphi^*(y) = - \inf_{x \in \mathcal{N}} \inf_{u \in \mathcal{U}} \{Ef(x, u) - \langle u, y \rangle\}$$

can be taken scenariowise while the first infimum may be restricted to the space \mathcal{N}^∞ of essentially bounded strategies. Both infima are easily calculated in many interesting applications; see [6] and the examples below.

¹A convex function is *closed* if it is lower semicontinuous and either proper or a constant. A function is *proper* if it never takes the value $-\infty$ and it is finite at some point.

Accordingly, we define the *Lagrangian integrand* on $\mathbb{R}^n \times \mathbb{R}^m \times \Omega$ by

$$l(x, y, \omega) := \inf_{u \in \mathbb{R}^m} \{f(x, u, \omega) - u \cdot y\}.$$

We will also need the pointwise conjugate of f :

$$\begin{aligned} f^*(v, y, \omega) &:= \sup_{x \in \mathbb{R}^n, u \in \mathbb{R}^m} \{x \cdot v + u \cdot y - f(x, u, \omega)\} \\ &= \sup_{x \in \mathbb{R}^n} \{x \cdot v - l(x, y, \omega)\}. \end{aligned}$$

By [18, Theorem 14.50], the pointwise conjugate of a normal integrand is also a normal integrand. Clearly, l is upper semicontinuous in the second argument since it is the pointwise infimum of continuous functions of y . Similarly,

$$\underline{l}(x, y, \omega) := \sup_{v \in \mathbb{R}^n} \{x \cdot v - f^*(v, y, \omega)\}$$

is lower semicontinuous in the first argument. In fact, $\underline{l}(\cdot, y, \omega)$ is the biconjugate of $l(\cdot, y, \omega)$ while $-l(x, \cdot, \omega)$ is the biconjugate of $-\underline{l}(x, \cdot, \omega)$; see [13, Theorem 34.2]. The function $(x, \omega) \mapsto \underline{l}(x, y(\omega), \omega)$ is a normal integrand for any $y \in \mathcal{Y}$ while $(y, \omega) \mapsto -l(x(\omega), y, \omega)$ is a normal integrand for any $x \in \mathcal{N}$ and thus, the integral functionals $E\underline{l}$ and El are well-defined on $\mathcal{N} \times \mathcal{Y}$ (recall our convention of defining an integral as $+\infty$ unless the positive part of the integrand is integrable). Indeed, we have $l(x(\omega), y, \omega) = -h^*(y, \omega)$ for $h(u, \omega) := f(x(\omega), u, \omega)$, where h is a normal integrand, by [18, Proposition 14.45(c)]. Similarly for \underline{l} .

Restricting strategies to the space $\mathcal{N}^\infty \subset \mathcal{N}$ of essentially bounded strategies gives rise to the auxiliary value function

$$\tilde{\varphi}(u) = \inf_{x \in \mathcal{N}^\infty} Ef(x, u).$$

Under the conditions of Theorem 2 below, the conjugates of $\tilde{\varphi}$ and φ coincide, or in other words, closures of $\tilde{\varphi}$ and φ are equal. The following lemma from [10] will play an important role. We denote

$$\mathcal{N}^\perp := \{v \in L^1(\Omega, \mathcal{F}, \mathbb{R}^n) \mid E(x \cdot v) = 0 \ \forall x \in \mathcal{N}^\infty\}.$$

Lemma 1. *Let $x \in \mathcal{N}$ and $v \in \mathcal{N}^\perp$. If $E[x \cdot v]^+ \in L^1$, then $E(x \cdot v) = 0$.*

We will use the notation

$$\text{dom}_1 Ef := \{x \in \mathcal{N} \mid \exists u \in \mathcal{U} : Ef(x, u) < +\infty\}.$$

Recall that *algebraic closure*, $\text{acl } C$, of a set C is the set of points x such that $(x, z] \subset C$ for some $z \in C$. Clearly, $C \subseteq \text{acl } C$ while in a topological vector space, $\text{acl } C \subseteq \text{cl } C$.

Theorem 2. *If $\text{dom } El(\cdot, y) \cap \mathcal{N}^\infty \subseteq \text{acl}(\text{dom}_1 Ef \cap \mathcal{N}^\infty)$, then*

$$\tilde{\varphi}^*(y) = - \inf_{x \in \mathcal{N}^\infty} El(x, y).$$

If, for every $x \in \mathcal{N}^\infty$ with $x \in \text{cl dom}_1 f$ almost surely, there exists $\bar{x} \in \mathcal{N}^\infty$ with $\bar{x} \in \text{rint dom}_1 f$ almost surely and $(\bar{x}, x) \in \text{dom}_1 Ef$, then

$$\tilde{\varphi}^*(y) = - \inf_{x \in \mathcal{N}^\infty} El(x, y).$$

We always have

$$\tilde{\varphi}^*(y) \leq \varphi^*(y) \leq \inf_{v \in \mathcal{N}^\perp} Ef^*(v, y).$$

In particular, if there exists $v \in \mathcal{N}^\perp$ such that $\tilde{\varphi}^(y) = Ef^*(v, y)$, then*

$$\tilde{\varphi}^*(y) = \varphi^*(y).$$

Proof. By the interchange rule ([18, Theorem 14.60]), $L(x, y) = El(x, y)$ for $x \in \text{dom}_1 Ef$ and thus,

$$\begin{aligned} -\tilde{\varphi}^*(y) &= \inf_{x \in \mathcal{N}^\infty} L(x, y) = \inf_{x \in \mathcal{N}^\infty \cap \text{dom}_1 Ef} L(x, y) \\ &= \inf_{x \in \mathcal{N}^\infty \cap \text{dom}_1 Ef} El(x, y) \geq \inf_{x \in \mathcal{N}^\infty} El(x, y). \end{aligned}$$

The converse holds trivially if $\text{dom } El(\cdot, y) \cap \mathcal{N}^\infty \neq \emptyset$. Otherwise, let $a \in \mathbb{R}$ and $\tilde{x} \in \mathcal{N}^\infty$ be such that $El(\tilde{x}, y) < a$. By the first assumption, there exists an $x' \in \text{dom}_1 Ef \cap \mathcal{N}^\infty$ such that $(1 - \lambda)\tilde{x} + \lambda x' \in \text{dom}_1 Ef$ for all $\lambda \in (0, 1]$. By convexity, $L((1 - \lambda)\tilde{x} + \lambda x', y) = El((1 - \lambda)\tilde{x} + \lambda x', y) < a$ for λ small enough and thus

$$-\tilde{\varphi}^*(y) = \inf_{x \in \mathcal{N}^\infty} El(x, y).$$

To prove the second claim, let $y \in \mathcal{Y}$ and $x \in \text{dom } El(\cdot, y) \cap \mathcal{N}^\infty$. By [13, Theorem 34.3], $x \in \text{cl dom}_1 f$ almost surely so, by assumption, there exists $\bar{x} \in \mathcal{N}^\infty$ with $\bar{x} \in \text{rint dom}_1 f$ almost surely and $(x, \bar{x}) \in \text{dom}_1 Ef$. Thus, $x \in \text{acl dom}_1 Ef$ so the first domain condition is satisfied. By [13, Theorem 6.1], $(x, \bar{x}) \subset \text{rint dom}_1 f$, and thus by the same line segment argument as above, the infimum in

$$-\tilde{\varphi}^*(y) = \inf_{x \in \mathcal{N}^\infty} El(x, y)$$

can be restricted to those x for which $x \in \text{rint dom}_1 f$. Then, by [13, Theorem 34.2], we may replace l by \underline{l} without affecting the infimum.

As to the last claim, the Fenchel inequality gives

$$f(x, u) + f^*(v, y) \geq x \cdot v + u \cdot y.$$

Therefore, for $(x, u) \in \text{dom } Ef$ and $v \in \mathcal{N}^\perp$ with $Ef^*(v, y) < \infty$, we have $E(x \cdot v) = 0$ by Lemma 1, so we get

$$\varphi^*(y) = \sup_{x \in \mathcal{N}, u \in \mathcal{U}} E[u \cdot y - f(x, u)] \leq \inf_{v \in \mathcal{N}^\perp} Ef^*(v, y).$$

Since $\tilde{\varphi} \geq \varphi$, have $\tilde{\varphi}^* \leq \varphi^*$. This completes the proof of the inequalities. The last statement concerning the equality clearly follows from the inequalities. \square

Theorem 2 gives conditions under which the conjugate of the value function of (P_u) can be expressed as

$$\varphi^*(y) = \inf_{x \in \mathcal{N}^\infty} El(x, y).$$

The second part gives conditions that allow one to replace l by \underline{l} in the above expression. One can then resort to the theory of normal integrands when calculating the infimum. It was shown in [6] that this yields many well-known dual expressions in operations research and mathematical finance. Unfortunately, the proof of the above expression in [6, Theorem 2.2] was incorrect and the given conditions are not sufficient in general; see [7]. The following example illustrates what can go wrong if the condition on the domains is omitted. Here and in what follows, δ_C denotes the indicator function of a set C : $\delta_C(x)$ equals 0 or $+\infty$ depending on whether $x \in C$ or not.

Example 3. Let \mathcal{F}_0 be trivial, $n_0 = 1$, $\mathcal{U} = L^\infty$, $\mathcal{Y} = L^1$, $\beta \in L^2$ be such that β^+ and β^- are unbounded, and let

$$f(x, u, \omega) = \delta_{\mathbb{R}_-}(\beta(\omega)x_0 + u)$$

so that $\text{dom } Ef = \{x \in \mathcal{N} \mid x_0 = 0\} \times L^\infty$, $\tilde{\varphi}^* = \delta_{L^1_+}$ and

$$l(x, y, \omega) = y\beta(\omega)x_0 - \delta_{\mathbb{R}_+}(y).$$

For $y = \beta^+$, we get $\inf_{x \in \mathcal{N}^\infty} El(x, y) = -\infty$ while $\tilde{\varphi}^*(y) = 0$. Here $\text{dom } El(\cdot, y) = \mathcal{N}$, so the first condition in Theorem 2 is violated.

The following example shows how the equality $\varphi^* = \tilde{\varphi}^*$ may fail to hold even when the first condition of Theorem 2 is satisfied.

Example 4. Let $T = 1$, $n = 2$, \mathcal{F}_0 is the completed trivial σ -algebra, and

$$f(x, u, \omega) = |x_0 - 1| + \delta_{\mathbb{R}_-}(\alpha(\omega)|x_0| - x_1) + \frac{1}{2}|u|^2,$$

where $\alpha \in L^0(\mathcal{F}_1)$ is positive and unbounded. It is easily checked that with $\mathcal{U} = \mathcal{Y} = L^2$ the first condition of Theorem 2 holds but $\tilde{\varphi}(u) = 1 + \frac{1}{2}\|u\|^2$ and $\varphi(u) = \frac{1}{2}\|u\|^2$.

At the moment, it is an open question whether the algebraic closure in the domain condition could be replaced by a topological closure. This can be done, however, e.g. if $El(\cdot, y)$ is upper semicontinuous on the closure of $\text{dom}_1 Ef \cap \mathcal{N}^\infty$. Also, in the deterministic case (where P is integer-valued on \mathcal{F}), the first condition is automatically satisfied since $\text{dom } l(\cdot, y) \subset \text{cl } \text{dom}_1 f$ for all y , by [13, Theorem 34.3].

The following describes a general situation where the assumptions of Theorem 2 are satisfied.

Example 5. Consider the problem

$$\begin{aligned} & \text{minimize} && Ef_0(x) \quad \text{over } x \in \mathcal{N} \\ & \text{subject to} && f_j(x) \leq 0 \text{ } P\text{-a.s.}, \quad j = 1, \dots, m, \end{aligned}$$

where f_j are normal integrands. The problem fits the general framework with

$$f(x, u, \omega) = \begin{cases} f_0(x, \omega) & \text{if } f_j(x, \omega) + u_j \leq 0 \text{ for } j = 1, \dots, m, \\ +\infty & \text{otherwise.} \end{cases}$$

This model was studied in Rockafellar and Wets [16] who gave optimality conditions in terms of dual variables. We will return to optimality conditions in the next section. For now, we note that if $f_j(x) \in \mathcal{U}$ for all $x \in \mathcal{N}^\infty$ and $j = 0, \dots, m$, then the assumptions of Theorem 2 are satisfied.

Indeed, $\text{dom}_1 Ef \cap \mathcal{N}^\infty = \mathcal{N}^\infty$ so the first two conditions in Theorem 2 hold. The lagrangian integrand can now be written as

$$l(x, y, \omega) = f_0(x, \omega) + \sum_{j=1}^m y_j f_j(x, \omega).$$

By [14, Theorem 22], the function $El(\cdot, y)$ is Mackey-continuous (with respect to the usual pairing of L^∞ and L^1) at the origin of L^∞ for every $y \in \mathcal{Y}$ so, by convexity, the function

$$\phi_y(\tilde{x}) := \inf_{x \in \mathcal{N}^\infty} El(x + \tilde{x}, y),$$

is Mackey-continuous as well; see e.g. [14, Theorem 8]. By [14, Theorem 11], this implies that ϕ_y has a subgradient at the origin, i.e. a $v \in L^1$ such that

$$El(x + \tilde{x}, y) \geq \phi_y(0) + E(\tilde{x} \cdot v) \quad \forall \tilde{x} \in L^\infty, x \in \mathcal{N}^\infty$$

or equivalently

$$El(\tilde{x}, y) \geq \phi_y(0) + E((\tilde{x} - x) \cdot v) \quad \forall \tilde{x} \in L^\infty, x \in \mathcal{N}^\infty.$$

This implies $v \in \mathcal{N}^\perp$ and $\inf_{\tilde{x} \in L^\infty} E[l(\tilde{x}, y) - \tilde{x} \cdot v] = \phi_y(0)$. By the interchange rule ([18, Theorem 14.60]), this can be written as $-Ef^*(v, y) = \phi_y(0) = -\tilde{\varphi}^*(y)$, so the last condition of Theorem 2 holds.

Section 4 of [5] studied financial models for pricing and hedging of portfolio-valued contingent claims (claims with physical delivery) along the lines of Kabanov [3]; see Example 12 below. The following example is concerned with the more classical financial model with claims with cash-delivery.

Example 6. Consider the problem

$$\text{minimize} \quad EV \left(u - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1} \right) \quad \text{over } x \in \mathcal{N}, \quad (\text{ALM})$$

where V convex normal integrand on $\mathbb{R} \times \Omega$ such that $V(\cdot, \omega)$ is nondecreasing nonconstant and $V(0, \omega) = 0$. This models the optimal investment problem of an agent with financial liabilities $u \in \mathcal{U}$ and a “disutility function” V . The \mathcal{F}_t -measurable vector s_t gives the unit prices of “risky” assets at time t and the vector x_t the units held over $(t, t + 1]$; see e.g. Rásonyi and Stettner [11] and the references therein.

Assume that, for every $x \in \mathcal{N}^\infty$, there is a $u \in \mathcal{U}$ such that $EV(u - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1}) < \infty$, then the closure of the value function φ of (ALM) has the dual representation

$$(\text{cl } \varphi)(u) = \sup_{y \in \mathcal{Q}} E[uy - V^*(y)],$$

where \mathcal{Q} is the set of positive multiples of martingale densities $y \in \mathcal{Y}$, i.e. densities dQ/dP of probability measures $Q \ll P$ under which the price process s is a martingale.

Indeed, (ALM) fits the general model (P_u) with

$$\begin{aligned} f(x, u, \omega) &= V \left(u - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1}(\omega), \omega \right), \\ l(x, y, \omega) &= -V^*(y, \omega) - y \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1}(\omega), \\ f^*(v, y, \omega) &= \begin{cases} V^*(y, \omega) & \text{if } v_t = -y \Delta s_{t+1}(\omega) \text{ for } t < T \text{ and } v_T = 0, \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

The integrability condition implies $\text{dom}_1 Ef \cap \mathcal{N}^\infty = \mathcal{N}^\infty$ so the first two conditions of Theorem 2 hold. Thus,

$$-\tilde{\varphi}^*(y) = \inf_{x \in \mathcal{N}^\infty} El(x, y) = -EV^*(y) + \inf_{x \in \mathcal{N}^\infty} E \left[- \sum_{t=0}^{T-1} x_t \cdot (y \Delta s_{t+1}) \right].$$

By Fenchel inequality,

$$uy - \sum_{t=0}^{T-1} x_t \cdot (y \Delta s_{t+1}) \leq V \left(u - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1} \right) + V^*(y) \quad P\text{-a.s.}$$

where for every $x \in \mathcal{N}^\infty$ and $y \in \text{dom } EV^*$, the right side is integrable for some $u \in \mathcal{U}$. Thus, $-\sum_{t=0}^{T-1} x_t \cdot (y \Delta s_{t+1})$ is integrable for every $x \in \mathcal{N}^\infty$ and $y \in \text{dom } EV^*$. Therefore the last infimum equals $-\infty$ unless $E_t(y \Delta s_{t+1}) = 0$ for every t , i.e. unless $y \in \mathcal{Q}$. Moreover, if $y \in \text{dom } \tilde{\varphi}^*$ then $Ef^*(v, y) = \tilde{\varphi}^*(y)$ holds with $v_t = y \Delta s_{t+1}$ for $t < T$ and $v_T = 0$, so the last condition of Theorem 2 holds.

The following example addresses a parameterized version of the generalized problem of Bolza studied in Rockafellar and Wets [17].

Example 7. Consider the problem

$$\underset{x \in \mathcal{N}}{\text{minimize}} \quad E \sum_{t=0}^T K_t(x_t, \Delta x_t + u_t), \quad (1)$$

where $n_t = d$, $\Delta x_t := x_t - x_{t-1}$, $x_{-1} := 0$ and each K_t is an \mathcal{F}_t -measurable normal integrand on $\mathbb{R}^d \times \mathbb{R}^d \times \Omega$.

We assume that, for every $x \in \mathcal{N}^\infty$ with $x_t \in \text{cl dom}_1 K_t$ for all t , there exists $\bar{x} \in \mathcal{N}^\infty$ with $\bar{x}_t \in \text{rint dom}_1 K_t$ for all t and $(\bar{x}, x) \in \text{dom}_1 Ef$, and that ${}^a u \in \mathcal{U}$ and ${}^a y \in \mathcal{Y}$ for every $u \in \mathcal{U}$ and $y \in \mathcal{Y}$, where ${}^a u_t = E_t u_t$. Then the closure of the value function φ of (1) has the dual representation

$$(\text{cl } \varphi)(u) = \sup_{y \in \mathcal{Y} \cap \mathcal{N}} E \sum_{t=0}^T [u_t \cdot y_t - K_t^*(E_t \Delta y_{t+1}, y_t)]$$

for every adapted $u \in \mathcal{U}$.

Indeed, the problem fits our general framework with

$$\begin{aligned} f(x, u, \omega) &= \sum_{t=0}^T K_t(x_t, \Delta x_t + u_t, \omega), \\ l(x, y, \omega) &= \sum_{t=0}^T [-x_t \cdot \Delta y_{t+1} + H_t(x_t, y_t, \omega)], \\ f^*(v, y, \omega) &= \sum_{t=0}^T K_t^*(v_t + \Delta y_{t+1}, y_t), \end{aligned}$$

where $u = (u_0, \dots, u_T)$ with $u_t \in \mathbb{R}^d$ and

$$H_t(x_t, y_t, \omega) := \inf_{u_t \in \mathbb{R}^d} \{K_t(x_t, u_t, \omega) - u_t \cdot y_t\}$$

is the associated Hamiltonian.

The domain condition in Theorem 2 is satisfied, so

$$-\tilde{\varphi}^*(y) = \inf_{x \in \mathcal{N}^\infty} El(x, y) = \inf_{x \in \mathcal{N}^\infty} E \sum_{t=0}^T [-x_t \cdot \Delta y_{t+1} + \underline{H}_t(x_t, y_t)],$$

where $\underline{H}_t(x_t, y_t, \omega) = \sup_{v_t} \{v_t \cdot x_t - K_t^*(v_t, y_t, \omega)\}$. Thus, by the interchange rule [18, Theorem 14.50],

$$\begin{aligned} -\tilde{\varphi}^*(y) &= - \inf_{x \in \mathcal{N}^\infty} E \sum_{t=0}^T [-x_t \cdot E_t \Delta y_{t+1} + \underline{H}_t(x_t, y_t, \omega)] \\ &= -E \sum_{t=0}^T K_t^*(E_t \Delta y_{t+1}, y_t) \end{aligned}$$

for adapted y . Moreover, with $v_t := E_t \Delta y_{t+1} - \Delta y_{t+1}$ we get $\tilde{\varphi}^*(y) = Ef^*(v, y)$, where $v \in \mathcal{N}^\perp$. The last condition in Theorem 2 thus holds. For any y , Jensen's inequality gives

$$\begin{aligned} -\tilde{\varphi}^*(y) &= \inf_{x \in \mathcal{N}^\infty} El(x, y) \\ &\leq \inf_{x \in \mathcal{N}^\infty} E \sum_{t=0}^T [-x_t \cdot E_t \Delta y_{t+1} + H_t(x_t, E_t y_t)] \\ &= -\tilde{\varphi}^*({}^a y). \end{aligned}$$

Therefore, for adapted u , we get

$$\text{cl } \varphi(u) = \sup_{y \in \mathcal{Y} \cap \mathcal{N}} E \sum_{t=0}^T [u_t \cdot y_t - K_t^*(E_t \Delta y_{t+1}, y_t)].$$

The dual representation of the value function in Example 7 was claimed to hold in [5] under the assumption that the Hamiltonian is lsc in x . The claim is, however, false in general since it omitted the domain condition in Theorem 2. The integrability condition posed in Example 7 not only provides a sufficient condition for that, but it also makes the lower semicontinuity of the Hamiltonian a redundant assumption.

3 Optimality conditions

The previous section as well as the articles [6, 9] were concerned with dual representations of the value function φ . Continuing in the general conjugate duality framework of Rockafellar [14], this section derives optimality conditions for (P_u) by assuming the existence of a subgradient of φ at u . Besides optimality conditions, this assumption implies the lower semicontinuity of φ at u (with respect to the weak topology induced on \mathcal{U} by \mathcal{Y}) and thus, the absence of a duality gap as well. Whereas in the above reference, the topology of convergence in measure in \mathcal{N} played an important role, below, topologies on \mathcal{N} are irrelevant.

Recall that a $y \in \mathcal{Y}$ is a *subgradient* of φ at $u \in \mathcal{U}$ if

$$\varphi(u') \geq \varphi(u) + \langle u' - u, y \rangle \quad \forall u' \in \mathcal{U}.$$

The set of all such y is called the *subdifferential* of φ at u and it is denoted by $\partial\varphi(u)$. If $\partial\varphi(u) \neq \emptyset$, then φ is lower semicontinuous at u and

$$\varphi(u) = \langle u, y \rangle - \varphi^*(y)$$

for every $y \in \partial\varphi(u)$. By [14, Theorem 11], $\partial\varphi(u) \neq \emptyset$, in particular, when φ is continuous at u .

We assume from now on that Ef is closed in u and that φ is proper.

Theorem 8. Assume that $\partial\varphi(u) \neq \emptyset$ and that for every $y \in \partial\varphi(u)$ there exists $v \in \mathcal{N}^\perp$ such that $\varphi^*(y) = Ef^*(v, y)$. Then an $x \in \mathcal{N}$ solves (P_u) if and only if it is feasible and there exist $y \in \mathcal{Y}$ and $v \in \mathcal{N}^\perp$ such that

$$(v, y) \in \partial f(x, u)$$

P -almost surely, or equivalently, if

$$v \in \partial_x l(x, y) \quad \text{and} \quad u \in \partial_y [-l](x, y)$$

P -almost surely.

Proof. Note first that if $y \in \partial\varphi(u)$ and $v \in \mathcal{N}^\perp$ such that $\varphi^*(y) = Ef^*(v, y)$, then $\varphi(u) + \varphi^*(y) = \langle u, y \rangle$ and

$$\inf_{x \in \mathcal{N}} Ef(x, u) = \langle u, y \rangle - \varphi^*(y) = E[u \cdot y - f^*(y, v)].$$

Thus, x solves (P_u) if and only if

$$E[f(x, u) + f^*(v, y)] = E[u \cdot y].$$

By Fenchel's inequality,

$$f(x, u) + f^*(v, y) \geq x \cdot v + u \cdot y,$$

so, by Lemma 1, $E[x \cdot v] = 0$ for every feasible x . Thus, x solves (P_u) if and only if x is feasible and the above inequality holds as an equality P -almost surely, that is, if

$$(v, y) \in \partial f(x, u)$$

P -almost surely. By [13, Theorem 37.5], this is equivalent to

$$v \in \partial_x l(x, y) \quad \text{and} \quad u \in \partial_y [-l](x, y)$$

P -almost surely. □

Example 9. In Example 5, the optimality conditions of Theorem 8 mean that

$$\begin{aligned} f_j(x) + u_j &\leq 0, \\ x &\in \operatorname{argmin}_{z \in \mathbb{R}^n} \left\{ f_0(z) + \sum_{j=1}^m y_j f_j(z) - z \cdot v \right\}, \\ y_j f_j(z) &= 0 \quad j = 1, \dots, m, \\ y_j &\geq 0 \end{aligned}$$

P -almost surely. These are the optimality conditions derived in [16], where sufficient conditions were given for the existence of an optimal $x \in \mathcal{N}^\infty$ and the corresponding dual variables $y \in \mathcal{Y}$ and $v \in \mathcal{N}^\perp$ (in our notation); see [16,

Theorem 1]. The conditions of [16] imply the continuity of the optimum value at the origin with respect to the L^∞ -norm. This yields the existence of dual variables in the norm dual $(L^\infty)^$. They then used the condition of “relatively complete recourse” to show that the projections of the dual variables to the subspace $L^1 \subset (L^\infty)^*$ satisfy the optimality conditions as well.*

We will now describe another general setup which covers many interesting applications and where the subdifferentiability condition $\partial\varphi(u) \neq \emptyset$ in Theorem 8 is satisfied. This framework is motivated by Biagini [1], where similar arguments were applied to optimal investment in the continuous-time setting. The idea is simply to require stronger continuity properties on φ in order to get the existence of dual variables directly in \mathcal{Y} (without going through the more exotic space $(L^\infty)^*$ first as in [16]).

A topological vector space is said to be *barreled* if every closed convex absorbing set is a neighborhood of the origin. By [14, Corollary 8B], a lower semicontinuous convex function on a barreled space is continuous throughout the algebraic interior (core) of its domain. On the other hand, by [14, Theorem 11], continuity implies subdifferentiability. Fréchet spaces and, in particular, Banach spaces are barreled. In the following, we say that \mathcal{U} is *barreled* if it is barreled with respect to a topology compatible with the pairing with \mathcal{Y} .

The following applies Theorem 8 to optimal investment in perfectly liquid financial markets.

Example 10. *Consider Example 6 and assume that \mathcal{U} is barreled, EV is finite on \mathcal{U} and EV^* is proper in \mathcal{Y} . Then an $x \in \mathcal{N}$ solves (ALM) if and only if it is feasible and there exists $y \in \mathcal{Q}$ such that*

$$y \in \partial V\left(u - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1}\right) \quad P\text{-a.s.}$$

Proof. By [14, Theorem 21], EV is lower semicontinuous so by [14, Corollary 8B], it is continuous. Since

$$\varphi(u) \leq Ef(0, u) = EV(u),$$

[14, Theorem 8] implies that φ is continuous and thus subdifferentiable throughout \mathcal{U} , by [14, Theorem 11]. Moreover, $\varphi^*(y) = Ef^*(v, y)$ for every $y \in \text{dom } \varphi^*$ and $v \in \mathcal{N}^\perp$ given by $v_t = y \Delta s_{t+1}$. The assumptions of Theorem 8 are thus satisfied. The subdifferential conditions for the Lagrangian integrand l can now be written as

$$\begin{aligned} v_t &= -y \Delta s_{t+1} \quad P\text{-a.s. for } t < T \text{ and } v_T = 0, \\ u - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1} &\in \partial V^*(y) \quad P\text{-a.s.} \end{aligned}$$

Since $v \in \mathcal{N}^\perp$, the former means that $y \in \mathcal{Q}$ while the latter can be written in terms of ∂V as stated; see [14, Theorem 12]. \square

The optimality conditions in Example 10 are classical in financial mathematics; for continuous-time models, see e.g. Schachermayer [19] or Biagini and Frittelli [2] and the references therein.

The next example applies Theorem 8 to the problem of Bolza from Example 7. It constructs a subgradient $y \in \partial\varphi(u)$ using Jensen's inequality.

Example 11. *Consider Example 7 and assume that \mathcal{U} is barreled and that there exists a normal integrand θ such that $E\theta$ is finite on \mathcal{U} , $E\theta^*$ is proper on \mathcal{Y} , and, for every $u \in \mathcal{U} \cap \mathcal{N}$,*

$$\sum_{t=0}^T K_t(x_t, \Delta x_t + u_t) \leq \theta(u) \quad P\text{-a.s.}$$

for some $x \in \mathcal{N}$. Then an $x \in \mathcal{N}$ solves

$$\underset{x \in \mathcal{N}}{\text{minimize}} \quad E \sum_{t=0}^T K_t(x_t, \Delta x_t + u_t)$$

for a $u \in \mathcal{U} \cap \mathcal{N}$ if and only if x is feasible and there exists $y \in \mathcal{Y} \cap \mathcal{N}$ such that

$$(E_t \Delta y_{t+1}, y_t) \in \partial K_t(x_t, \Delta x_t + u_t)$$

P -almost surely for all t , or equivalently, if

$$\begin{aligned} E_t \Delta y_{t+1} &\in \partial_x H_t(x_t, y_t), \\ u_t + \Delta x_t &\in \partial_y [-H_t](x_t, y_t) \end{aligned}$$

P -almost surely for all t .

Proof. The space $\mathcal{U} \cap \mathcal{N}$ is a barreled space and its continuous dual may be identified with $\mathcal{Y} \cap \mathcal{N}$. Indeed, by Hahn–Banach, any $y \in (\mathcal{U} \cap \mathcal{N})^*$ may be extended to a $\bar{y} \in \mathcal{Y}$ for which ${}^a \bar{y} \in \mathcal{Y} \cap \mathcal{N}$ coincides with y on $\mathcal{U} \cap \mathcal{N}$. Moreover, for any closed convex absorbing set B in $\mathcal{U} \cap \mathcal{N}$, we have that $\hat{B} = \{u \in \mathcal{U} \mid {}^a u \in B\}$ is a closed convex absorbing set in \mathcal{U} , so it is a neighborhood of the origin of \mathcal{U} which implies that $\hat{B} \cap \mathcal{N} \subset B$ is a neighborhood of the origin of $\mathcal{U} \cap \mathcal{N}$.

By assumption, for every $u \in \mathcal{U} \cap \mathcal{N}$ there is an $x \in \mathcal{N}$ such that

$$\varphi(u) \leq E \sum_{t=0}^T K_t(x_t, \Delta x_t + u_t) \leq E\theta(u).$$

By [14, Theorem 21], $E\theta$ is lower semicontinuous so by [14, Corollary 8B], it is continuous. Thus, φ is continuous and in particular, subdifferentiable on $\mathcal{U} \cap \mathcal{N}$ (see [14, Theorem 11]) so for every $u \in \mathcal{U} \cap \mathcal{N}$ there is a $y \in \mathcal{Y} \cap \mathcal{N}$ such that

$$\varphi(u') \geq \varphi(u) + \langle u' - u, y \rangle \quad \forall u' \in \mathcal{U} \cap \mathcal{N}.$$

Since each K_t is \mathcal{F}_t -measurable, Jensen's inequality gives

$$\varphi(u') \geq \varphi({}^a u') \geq \varphi(u) + \langle {}^a u' - u, y \rangle = \varphi(u) + \langle u' - u, y \rangle \quad \forall u' \in \mathcal{U},$$

so $y \in \partial\varphi(u)$ as well. Moreover, as observed in Example 7, we have $Ef^*(v, y) = \varphi^*(y)$ for $v_t = E_t \Delta y_{t+1} - \Delta y_{t+1}$. The subdifferential conditions in Theorem 8 for the Lagrangian integrand l become

$$\begin{aligned} v_t + \Delta y_{t+1} &\in \partial_x H_t(x_t, y_t) \quad P\text{-a.s.} \quad \forall t, \\ u_t + \Delta x_t &\in \partial_y [-H_t](x_t, y_t) \quad P\text{-a.s.} \quad \forall t. \end{aligned}$$

By [13, Theorem 37.5], these are equivalent to the conditions in terms of K_t . \square

The optimality condition in terms of K_t in Example 11 can be viewed as a *stochastic Euler-Lagrange condition* in discrete-time much like that in [17, Theorem 4]. There is a difference, however, in that [17] studied the problem of minimizing $E \sum K_t(x_{t-1}, \Delta x_t)$ and, accordingly, the measurability conditions posed on the dual variables were different as well. The condition in terms of H_t can be viewed as a *stochastic Hamiltonian system* in discrete-time; see [12, Section 9] for deterministic models in continuous-time. Our assumptions on the problem also differ from those made in [17]. Whereas the assumptions of [17] and the line of argument follows that in [16] (see Example 9), our assumption implies the continuity of φ on the adapted subspace $\mathcal{U} \cap \mathcal{N}$.

It is essential that the growth condition in Example 11 is required only for adapted $u \in \mathcal{U}$. Indeed, since x is adapted, it would often be too much to ask the upper bound for nonadapted u . This is the case e.g. in the following example which is concerned with the financial model studied in [8, 9]. The model is an extension of the currency market model introduced by Kabanov [3]; see also Kabanov and Safarian [4].

Example 12. *Consider the optimal investment-consumption problem*

$$\begin{aligned} \text{minimize} \quad & E \sum_{t=0}^T V_t(-k_t) \quad \text{over } (z, k) \in \mathcal{N} & \text{(OCP)} \\ \text{subject to} \quad & \Delta z_t + u_t + k_t \in C_t, \quad z_T = 0 \quad P\text{-a.s. } t = 0, \dots, T, \end{aligned}$$

where $z_{-1} := 0$, and for each t , $C_t : \Omega \rightrightarrows \mathbb{R}^d$ is \mathcal{F}_t -measurable² and closed convex-valued with $0 \in C_t$ and V_t is an \mathcal{F}_t -measurable nondecreasing (wrt \mathbb{R}_+^d) convex normal integrand with $V_t(0) = 0$.

We will assume as in Example 7 that adapted projections of the elements of \mathcal{U} and \mathcal{Y} still belong to \mathcal{U} and \mathcal{Y} , respectively. We also assume that \mathcal{U} is barreled and that there exists a normal integrand θ such that $E\theta$ is finite on \mathcal{U} , $E\theta^*$ is proper on \mathcal{Y} , and that

$$V_T \left(\sum_{t=0}^T u_t \right) \leq \theta(u) \quad \forall u \in \mathbb{R}^{(T+1)d}.$$

² C_t is \mathcal{F}_t -measurable if $\{\omega \in \Omega \mid C_t(\omega) \cap O \neq \emptyset\} \in \mathcal{F}_t$ for every open O .

Then $(z, k) \in \mathcal{N}$ solves (OCP) for $u \in \mathcal{U} \cap \mathcal{N}$ if and only if (z, k) is feasible and there exists a martingale $y \in \mathcal{Y}$ such that³

$$\begin{aligned} -k_t &\in \partial V_t^*(y_t), \\ \Delta z_t + u_t + k_t &\in \partial \sigma_{C_t}(y_t) \end{aligned}$$

P -almost surely for $t = 0, \dots, T$. When C_t are conical, the last condition means that

$$\Delta z_t + u_t + k_t \in C_t, \quad y_t \in C_t^*, \quad (\Delta z_t + u_t + k_t) \cdot y_t = 0$$

P -almost surely for $t = 0, \dots, T$. Thus, y generalizes the notion of a “consistent price system” introduced in [3] to nonconical market models; see [9].

Proof. Problem (OCP) can be expressed in the format of Example 11 with $x = (z, k)$, $u = (u^z, u^k)$ and

$$\begin{aligned} K_t(x_t, u_t, \omega) &= V_t(-k_t, \omega) + \delta_{C_t(\omega)}(u_t^z + k_t) \quad t = 0, \dots, T-1, \\ K_T(x_T, u_T, \omega) &= V_T(-k_T, \omega) + \delta_{C_T(\omega)}(u_T^z + k_T) + \delta_{\{0\}}(z_T). \end{aligned}$$

The Hamiltonian becomes

$$\begin{aligned} H_t(x_t, y_t, \omega) &= \delta_{\{0\}}(y_t^k) + k_t \cdot y_t^z + V_t(-k_t, \omega) - \sigma_{C_t(\omega)}(y_t^z) \quad t = 0, \dots, T-1, \\ H_T(x_T, y_T, \omega) &= \delta_{\{0\}}(y_T^k) + k_T \cdot y_T^z + V_T(-k_T, \omega) - \sigma_{C_T(\omega)}(y_T) + \delta_{\{0\}}(z_T). \end{aligned}$$

We are in the setting of Example 7, since the domain condition is satisfied by $\bar{x} = (0, \bar{k})$ for any $x \in \mathcal{N}^\infty$, where $\bar{k}_t = 1$ for all t .

Given $u \in \mathcal{U} \cap \mathcal{N}$, we have

$$\sum_{t=0}^T K_t(x_t^u, \Delta x_t + u_t, \omega) = V_T \left(\sum_{t=0}^T u_t \right) \leq \theta(u),$$

where $x^u = (z^u, k^u) \in \mathcal{N}$ is defined by $\Delta z_t^u = -u_t$, $k_t^u = 0$ for $t < T$ and $z_T^u = 0$ and $k_T^u = -\sum_{t=0}^T u_t$. Thus, the assumptions of Example 11 are satisfied with $y = 0$ and $\beta = 0$. The given optimality conditions are just a special case of the optimality conditions in Example 11. Indeed, the subdifferential conditions become

$$\begin{aligned} z_T &= 0 \text{ and } E_t \Delta y_{t+1}^z = 0 \quad \forall t = 0, \dots, T-1, \\ E_t \Delta y_{t+1}^k &\in y_t^z - \partial V_t(-k_t) \quad \forall t = 0, \dots, T, \\ u_t^z + \Delta z_t &\in -k_t + \partial \sigma_{C_t}(y_t^z) \quad \forall t = 0, \dots, T, \\ y_t^k &= 0 \quad \forall t = 0, \dots, T. \end{aligned}$$

The first condition means that y^z is a martingale and $z_T = 0$. The second and the last condition mean that $y^k = 0$ and $y_t^z \in \partial V_t(-k_t)$. By [14, Theorem 12], the subdifferential condition can be written as $-k_t \in \partial V_t^*(y_t)$. \square

³Here, $\sigma_{C_t}(y_t) := \sup\{z \cdot y_t \mid z \in C_t\}$.

References

- [1] S. Biagini. An Orlicz spaces duality for utility maximization in incomplete markets. In *Seminar on Stochastic Analysis, Random Fields and Applications V*, volume 59 of *Progr. Probab.*, pages 445–455. Birkhäuser, Basel, 2008.
- [2] S. Biagini and M. Frittelli. A unified framework for utility maximization problems: An Orlicz spaces approach. *The Annals of Applied Probability*, 18(3):929–966, 2008.
- [3] Y. M. Kabanov. Hedging and liquidation under transaction costs in currency markets. *Finance and Stochastics*, 3(2):237–248, 1999.
- [4] Y. M. Kabanov and M. Safarian. *Markets with transaction costs*. Springer Finance. Springer-Verlag, Berlin, 2009. Mathematical theory.
- [5] T. Pennanen. Arbitrage and deflators in illiquid markets. *Finance and Stochastics*, 15(1):57–83, 2011.
- [6] T. Pennanen. Convex duality in stochastic optimization and mathematical finance. *Mathematics of Operations Research*, 36(2):340–362, 2011.
- [7] T. Pennanen. Correctional note: "Convex duality in stochastic optimization and mathematical finance". *manuscript*, 2015.
- [8] T. Pennanen and I. Penner. Hedging of claims with physical delivery under convex transaction costs. *SIAM Journal on Financial Mathematics*, 1:158–178, 2010.
- [9] T. Pennanen and A.-P. Perkkiö. Stochastic programs without duality gaps. *Mathematical Programming*, 136(1):91–110, 2012.
- [10] A.-P. Perkkiö. Stochastic programs without duality gaps for objectives without a lower bound. *manuscript*, 2014.
- [11] M. Rásonyi and L. Stettner. On utility maximization in discrete-time financial market models. *Ann. Appl. Probab.*, 15(2):1367–1395, 2005.
- [12] R. T. Rockafellar. Conjugate convex functions in optimal control and the calculus of variations. *J. Math. Anal. Appl.*, 32:174–222, 1970.
- [13] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
- [14] R. T. Rockafellar. *Conjugate duality and optimization*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974.
- [15] R. T. Rockafellar. Integral functionals, normal integrands and measurable selections. In *Nonlinear operators and the calculus of variations (Summer School, Univ. Libre Bruxelles, Brussels, 1975)*, pages 157–207. Lecture Notes in Math., Vol. 543. Springer, Berlin, 1976.

- [16] R. T. Rockafellar and R. J.-B. Wets. The optimal recourse problem in discrete time: L^1 -multipliers for inequality constraints. *SIAM J. Control Optimization*, 16(1):16–36, 1978.
- [17] R. T. Rockafellar and R. J.-B. Wets. Deterministic and stochastic optimization problems of Bolza type in discrete time. *Stochastics*, 10(3-4):273–312, 1983.
- [18] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.
- [19] W. Schachermayer. Optimal investment in incomplete markets when wealth may become negative. *Ann. Appl. Probab.*, 11(3):694–734, 2001.