MATHEMATICS OF OPERATIONS RESEARCH
Vol. 00, No. 0, Xxxxxx 20xx, pp. xxx–xxx
ISSN 0364-765X | EISSN 1526-5471 |xx|0000|0xxx

DOI 10.1287/moor.xxxx.xxxx


# Convex duality in stochastic optimization and mathematical finance

Teemu Pennanen
Department of Mathematics and Statistics, University of Jyväskylä
email: teemu.pennanen@jyu.fi http://math.tkk.fi/~teemu/

This paper proposes a general duality framework for the problem of minimizing a convex integral functional over a space of stochastic processes adapted to a given filtration. The framework unifies many well-known duality frameworks from operations research and mathematical finance. The unification allows the extension of some useful techniques from these two fields to a much wider class of problems. In particular, combining certain finite-dimensional techniques from convex analysis with measure theoretic techniques from mathematical finance, we are able to close the duality gap in some situations where traditional topological arguments fail.

*Key words*: Convex duality, stochastic programming, mathematical finance

*MSC2000 Subject Classification*: Primary: 52A41, 90C15, 46A20; Secondary: 91B25, 49J53

*OR/MS subject classification*: Primary: Programming/stochastic, Mathematics/convexity; Secondary: Finance/asset pricing

*History*: Received: Xxxx xx, xxxx; Revised: Yyyyyy yy, yyyy and Zzzzzz zz, zzzz.

**1. Introduction** Let $(\Omega, \mathcal{F}, P)$ be a probability space with a filtration $(\mathcal{F}_t)_{t=0}^T$ (an increasing sequence of sub-sigma-algebras of $\mathcal{F}$) and consider the problem

$$\text{minimize} \quad Ef(x(\omega), u(\omega), \omega) \quad \text{over } x \in \mathcal{N} \tag{1}$$

where $f$ is an extended real-valued function, $\mathcal{N}$ is a space of $(\mathcal{F}_t)_{t=0}^T$-adapted stochastic processes and $u$ is a measurable function (exact definitions will be given below). The variable $u$ represents parameters or perturbations of a dynamic decision making problem where the objective is to minimize the expectation over *decision strategies* $x$ adapted to the information available to the decision maker over time. This paper derives dual expressions for the optimal value of (1) by incorporating some measure theoretic techniques from mathematical finance into the general conjugate duality framework of Rockafellar [41].

Problem (1) covers many important optimization models in operations research and mathematical finance. Specific instances of stochastic optimization problems can often be put in the above format by appropriately specifying the integrand $f$. Allowing the integrand $f$ to take on the value $+\infty$, we can represent various pointwise (almost sure) constraints by infinite penalties. Some of the earliest examples can be found in Danzig [12] and Beale [4]. Problem (1) provides a very general framework also for various optimization and pricing problems in mathematical finance. Certain classes of stochastic control problems can also be put the above form; see [45, Section 6]. In some applications, the parameter $u$ is introduced into a given problem in order to derive information (such as optimality conditions or bounds on the optimal value) about it. This is the point of view taken e.g. in [41]. In other applications, the parameter $u$ has a natural interpretation in the original formulation itself. Examples include financial applications where $u$ may represent the payouts of a financial instrument such as an option and one is trying to minimize the initial cost of a hedging portfolio.

Convex duality has widespread applications in operations research, calculus of variations and mechanics. Besides in deriving optimality conditions, duality is used in numerical optimization and bounding techniques. The essence of convex duality is beautifully summarized by the conjugate duality framework of [41] which subsumes more special duality frameworks such as Lagrangian (and in particular linear programming) and Fenchel duality; see also Ekeland and Temam [18]. Several duality results, including optimality conditions for certain instances of (1) have been derived from the conjugate duality framework in Rockafellar and Wets [43, 44, 45, 46].

Convex duality has long been an integral part also of mathematical finance but there, duality results are often derived ad hoc instead of embedding a given problem in a general optimization framework; see however Pliska [37] and King [28], where finite-dimensional financial models are treated in the classical linear programming duality framework. Attempts to derive financial duality results in general probability spaces from known optimization frameworks are often hindered by two features. First, general duality frameworks are often formulated in locally convex topological vector spaces while in financial problems the decision strategies are usually chosen from a space that lacks an appropriate locally convex topology. Second, general duality results are often geared towards attainment of the dual optimum which requires conditions that often fail to hold in financial applications; see however Korf [31] and Tian and Wets [54] where financial applications are treated in an optimization framework without dual attainment.

The main contribution of this paper is to propose a general duality framework that covers several problems both in operations research as well as in mathematical finance. Our framework, to be rigorously specified in Section 2, is an extension of the stochastic programming duality frameworks proposed in [43, 45]. In our framework the parameters $u$ enter the model in a more general manner and we do not restrict the decision strategies $x$ to be bounded or integrable a priori. Allowing strategies to be general adapted processes has turned out to be useful in deriving various duality results for financial models; see e.g. Schachermayer and Delbaen [16], Kabanov and Safarian [25] and their references. This paper extends such techniques to a much more general class of models. We obtain dual representations for the optimal value of (1) but not necessarily the dual attainment as opposed to the strong duality results in [43, 44, 45, 46]. Consequently, we cannot claim the necessity of various optimality conditions involving dual variables. Nevertheless, the mere absence of duality gap is useful in many situations e.g. in mathematical finance where the "constraint qualifications" required for classical duality results often fail to hold. For example, various dual representations of hedging costs correspond to the absence of the duality gap while the dual optimum might not be attained. As an application, we extend certain results on superhedging and optimal consumption to a general market model with nonlinear illiquidity effects and convex portfolio constraints. This will be done by extending the elegant (currency) market model of Kabanov [24] where all assets are treated symmetrically. More traditional market models are then covered as special cases. The absence of duality gap is useful also in deriving certain simulation-based numerical techniques for bounding the optimum value of (1) as e.g. those proposed in Rogers [48] and Haugh and Kogan [21] in the case of optimal stopping problems. We extend such techniques for a more general class of problems.

The rest of this paper is organized as follows. Section 2 presents the general duality framework for problem (1) based on the conjugate duality framework of [41]. Sections 3 and 4 give some well-known examples and extensions of duality frameworks from operations research and mathematical finance, respectively. Section 5 extends some classical closedness criteria from finite-dimensional spaces to the present infinite-dimensional stochastic setting.

**2. Conjugate duality** We study (1) in the *conjugate duality* framework of Rockafellar [41]. However, we deviate from [41] in that the space $\mathcal{N}$ of decision variables need not be a locally convex topological vector space paired with another one. This precludes the completely symmetric duality in [41] but in some situations it yields more regularity for the optimal value than what can be obtained e.g. with integrable strategies.

For given integers $n_t$, we set

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

where $L^0(\Omega, \mathcal{F}_t, P; \mathbb{R}^{n_t})$ denotes the space of equivalence classes of $\mathcal{F}_t$-measurable $\mathbb{R}^{n_t}$-valued functions that coincide $P$-almost surely. Each $x_t$ is interpreted as a decision that is made after observing all available information at time $t$. In applications, the filtration $(\mathcal{F}_t)_{t=0}^T$ is often generated by a finite-dimensional stochastic process whose values are observed at discrete points in time. If $\mathcal{F}_0$ is the trivial sigma algebra $\{\emptyset, \Omega\}$ then the first component $x_0$ is deterministic.

The function $f$ is assumed to be an *extended real-valued convex normal integrand* on $\mathbb{R}^n \times \mathbb{R}^m \times \Omega$ where $n = n_0 + \ldots + n_T$ and $m$ is a given integer. This means that the set-valued mapping $\omega \mapsto \{(x, u, \alpha) \,|\, f(x, u, \omega) \le \alpha\}$ is $\mathcal{F}$-measurable and it has closed and convex values (so $(x, u) \mapsto f(x, u, \omega)$ is convex and lower semicontinuous for every $\omega$); see e.g. [47, Chapter 14]. This implies that $f$ is $\mathcal{B}(\mathbb{R}^n \times \mathbb{R}^m) \otimes \mathcal{F}$-measurable and that the function $(x, u) \mapsto f(x, u, \omega)$ is lower semicontinuous and convex for every $\omega$. It follows that $\omega \mapsto f(x(\omega), u(\omega), \omega)$ is $\mathcal{F}$-measurable for every $x \in L^0(\Omega, \mathcal{F}, P; \mathbb{R}^n)$

and $u \in L^0(\Omega, \mathcal{F}, P; \mathbb{R}^m)$. Throughout this paper, the expectation of an extended real-valued measurable function is defined as $+\infty$ unless the positive part is integrable. The *integral functional*

$$I_f(x,u) := Ef(x(\omega), u(\omega), \omega)$$

in the objective of (1) is then well-defined extended real-valued convex function on $L^0(\Omega, \mathcal{F}, P; \mathbb{R}^n) \times L^0(\Omega, \mathcal{F}, P; \mathbb{R}^m)$. Normal integrands possess many useful properties and they arise quite naturally in many optimization problems in practice. Examples will be given in the following sections. We refer the reader to [42], [8] or [47, Chapter 14] for general treatments of normal integrands.

For each $u \in L^0(\Omega, \mathcal{F}, P; \mathbb{R}^m)$, the optimal value of (1) is given by the *value function*

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

By [41, Theorem 1], $\varphi$ is convex. We will derive dual expressions for $\varphi$ on the space $L^p := L^p(\Omega, \mathcal{F}, P; \mathbb{R}^m)$ using the conjugate duality framework of Rockafellar [41]. To this end, we pair $L^p$ with $L^q$, where $q \in [1, \infty]$ is such that $1/p + 1/q = 1$. The bilinear form

$$\langle u, y \rangle = E[u(\omega) \cdot y(\omega)]$$

puts $L^p$ and $L^q$ in separating duality. The weakest and the strongest locally convex topologies on $L^p$ compatible with the pairing will be denoted by $\sigma(L^p, L^q)$ and $\tau(L^p, L^q)$, respectively (similarly for $L^q$). By the classical separation argument, a convex function is lower semicontinuous with respect to $\sigma(L^p, L^q)$ if it is merely lower semicontinuous with respect to $\tau(L^p, L^q)$.

REMARK 2.1 *For $p \in [1, \infty)$, $\tau(L^p, L^q)$ is the norm topology and $\sigma(L^p, L^q)$ is the weak-topology that $L^p$ has as a Banach space with the usual $L^p$-norm. For $p = \infty$, $\sigma(L^p, L^q)$ is the weak*-topology that $L^p$ has as the Banach dual of $L^q$ while $\tau(L^p, L^q)$ is, in general, weaker than the norm topology. It follows from the Mackey-Arens and Dunford-Pettis theorems, that a sequence in $L^\infty$ converges with respect to $\tau(L^\infty, L^1)$ if and only if it is norm-bounded and converges in measure; see Grothendieck [20, Part 4] for the case of locally compact measure spaces. In mathematical finance, a convex function on $L^\infty$ is sometimes said to have the "Fatou property" if it is sequentially lower-semicontinuous with respect to $\tau(L^\infty, L^1)$.*

REMARK 2.2 *Instead of $L^p$ and $L^q$, we could take an arbitrary pair of spaces of measurable $\mathbb{R}^m$-valued functions which are in separating duality under the bilinear form $\langle u, y \rangle = E[u(\omega) \cdot y(\omega)]$. Examples include Orlicz spaces which have recently been used in a financial context by Biagini and Frittelli [7].*

The *conjugate* of a function $\varphi$ on $L^p$ is the convex function on $L^q$ defined by

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

The conjugate of a function on $L^q$ is defined similarly. It is a fundamental result in convex duality that $\varphi^{**} = \operatorname{cl} \varphi$ where

$$\operatorname{cl} \varphi = \begin{cases} \operatorname{lsc} \varphi & \text{if } (\operatorname{lsc} \varphi)(u) > -\infty \ \forall u \in L^p, \\ -\infty & \text{otherwise} \end{cases}$$

is the *closure* of $\varphi$; see e.g. [41, Theorem 5]. Here $\operatorname{lsc} \varphi$ denotes the *lower semicontinuous hull* of $\varphi$. If $\operatorname{lsc} \varphi$ has a finite value at some point then $\operatorname{lsc} \varphi$ is proper and $\operatorname{lsc} \varphi = \operatorname{cl} \varphi$; see [41, Theorem 4].

The *Lagrangian* associated with (1) is the extended real-valued function on $\mathcal{N} \times L^q$ defined by

$$L(x,y) = \inf_{u \in L^p} \{I_f(x,u) - \langle u, y \rangle\}.$$

The Lagrangian is convex in $x$ and concave in $y$. The *dual objective* is the extended real-valued function on $L^q$ defined by

$$g(y) = \inf_{x \in \mathcal{N}} L(x,y).$$

Since $g$ is the pointwise infimum of concave functions, it is concave. The basic duality result [41, Theorem 7] says, in particular, that

$$g = -\varphi^*.$$

This follows directly from the above definitions and does not rely on topological properties of $\mathcal{N}$. The biconjugate theorem then gives the dual representation

$$(\operatorname{cl}\varphi)(u) = \sup\{\langle u, y\rangle + g(y)\}. \tag{2}$$

In many applications, the parameter $u$ has practical significance, and the dual representation (2) may yield valuable information about the function $\varphi$. On the other hand, in some situations, one is faced with a fixed optimization problem and the parameter $u$ is introduced in order to derive information about the original problem. This is the perspective taken in [41], where the minimization problem

$$\text{minimize} \quad I_f(x,0) \quad \text{over} \quad x \in \mathcal{N} \tag{3}$$

would be called the *primal problem* and

$$\text{maximize} \quad g(y) \quad \text{over} \quad y \in L^q \tag{4}$$

the *dual problem.* By (2), the optimum values of (3) and (4) are equal exactly when $(\operatorname{cl}\varphi)(0) = \varphi(0)$. An important topic which is studied in [41] but not in the present paper is derivatives of the value function $\varphi$ and the associated optimality conditions. In this paper, we concentrate on the more general property of lower semicontinuity of $\varphi$; see Section 5. The lower semicontinuity already yields many interesting results in operations research and mathematical finance. Moreover, lower semicontinuity is useful for proving the continuity of $\varphi$ for $p < \infty$ since a lower semicontinuous convex function on a barreled space is continuous throughout the interior of its domain; see e.g. [41, Corollary 8B].

Remark 2.3 *As long as the integral functional $(x,u) \mapsto I_f(x,u)$ is closed in u (which holds under quite general conditions given e.g. in Rockafellar [38]), the biconjugate theorem gives*

$$I_f(x,u) = \sup_{y\in L^q}\{L(x,y) + \langle u, y\rangle\}$$

*and, in particular, $I_f(x,0) = \sup_y L(x,y)$ so that $\varphi(0) = \inf_{x\in\mathcal{N}} \sup_{y\in L^q} L(x,y)$. On the other hand, (2) gives $(\operatorname{cl}\varphi)(0) = \sup_{y\in L^q} \inf_{x\in\mathcal{N}} L(x,y)$ so that the condition $(\operatorname{cl}\varphi)(0) = \varphi(0)$ can be expressed as*

$$\inf_{x\in\mathcal{N}} \sup_{y\in L^q} L(x,y) = \sup_{y\in L^q} \inf_{x\in\mathcal{N}} L(x,y).$$

*In other words, the function L has a saddle-value iff $\varphi$ is closed at the origin. Along with the general duality theory for convex minimization, the conjugate duality framework of [41] addresses general convex-concave minimax problems.*

The following interchange rule will be useful in deriving more explicit expressions for the dual objective $g$. It is a special case of [47, Theorem 14.60] and it uses the fact that for an $\mathcal{F}$-measurable normal integrand $h$, the function $\omega \mapsto \inf_u h(u,\omega)$ is $\mathcal{F}$-measurable; see [47, Theorem 14.37].

Theorem 2.1 (Interchange rule) *Given an $\mathcal{F}$-measurable normal integrand h on $\mathbb{R}^k \times \Omega$, we have*

$$\inf_{u\in L^p} Eh(u(\omega),\omega) = E \inf_{u\in\mathbb{R}^k} h(u,\omega)$$

*as long as the left side is less than $+\infty$.*

Theorem 2.1 yields a simple proof of Jensen's inequality. Throughout this paper, the *conditional expectation* of a random variable $x$ with respect to $\mathcal{F}_t$ will be denoted by $E_t x$; see e.g. Shiryaev [53, II.7].

Corollary 2.1 (Jensen's inequality) *Let h be an $\mathcal{F}_t$-measurable convex normal integrand on $\mathbb{R}^k \times \Omega$ such that $Eh^*(v(\omega),\omega) < \infty$ for some $v \in L^q(\Omega, \mathcal{F}_t, P; \mathbb{R}^k)$. Then*

$$Eh((E_t x)(\omega),\omega) \le Eh(x(\omega),\omega)$$

*for every $x \in L^p(\Omega, \mathcal{F}, P; \mathbb{R}^k)$.*

Proof. Applying Theorem 2.1 twice, we get

$$\begin{aligned} I_h(E_t x) &= E \sup_v \{v \cdot (E_t x)(\omega) - h^*(v,\omega))\} \\ &= \sup_{v \in L^q(\mathcal{F}_t)} E\{v(\omega) \cdot (E_t x)(\omega) - h^*(v(\omega),\omega)\} \\ &= \sup_{v \in L^q(\mathcal{F}_t)} E\{v(\omega) \cdot x(\omega) - h^*(v(\omega),\omega)\} \\ &\le \sup_{v \in L^q(\mathcal{F})} E\{v(\omega) \cdot x(\omega) - h^*(v(\omega),\omega)\} \\ &= E \sup_v \{v \cdot x(\omega) - h^*(v,\omega)\} \\ &= Eh(x(\omega),\omega), \end{aligned}$$

where the third equality comes from the law of iterated expectations; see e.g. [53, Section II.7]. □

Going back to (1), we define

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

This is an extended real-valued function on $\mathbb{R}^n \times \mathbb{R}^m \times \Omega$, convex in $x$ and concave in $y$. Various dual expressions in stochastic optimization and in mathematical finance can be derived from the following result which expresses the dual objective in terms of $l$. In many situations, the expression can be written concretely in terms of problem data; see Sections 3 and 4. Given an $r \in [1,\infty]$, we let

$$\mathcal{N}^r := \mathcal{N} \cap L^r(\Omega, P, \mathcal{F}; \mathbb{R}^n).$$

Theorem 2.2 *The function $\omega \mapsto l(x(\omega), y(\omega), \omega)$ is measurable for any $x \in \mathcal{N}$ and $y \in L^q$ so the integral functional $I_l(x,y) = El(x(\omega), y(\omega), \omega)$ is well-defined on $\mathcal{N} \times L^q$. As long as $I_f \not\equiv +\infty$, we have*

$$g(y) = \inf_{x \in \mathcal{N}} I_l(x,y).$$

*If, in addition, $l$ is of the form*[1]

$$l(x,y,\omega) = \sum_{t=0}^T l_t(x_t, y, \omega)$$

*for some $\mathcal{B}(\mathbb{R}^{n_t}) \otimes \mathcal{B}(\mathbb{R}^m) \otimes \mathcal{F}$-measurable extended real-valued functions $l_t$ on $\mathbb{R}^{n_t} \times \mathbb{R}^m \times \Omega$ then*

$$g(y) = \inf_{x \in \mathcal{N}^r} I_l(x,y)$$

*as long as the right side is less than $+\infty$.*

Proof. We have $l(x(\omega), y(\omega), \omega) = -h^*(y(\omega), \omega)$, where $h(u,\omega) := f(x(\omega), u, \omega)$. To prove the measurability it suffices to show that $h^*$ is a normal integrand on $\mathbb{R}^m \times \Omega$. This follows from Proposition 14.45(c) and Theorem 14.50 of [47].

If $I_f \not\equiv +\infty$, then there exists an $x \in \mathcal{N}$ such that $L(x,y) < \infty$ for every $y \in L^q$. We can thus assume that $L(x,y) < \infty$ in the expression for $g$ so that

$$L(x,y) = E \inf_{u \in \mathbb{R}^m} \{f(x(\omega), u, \omega) - u \cdot y(\omega)\} = I_l(x,y),$$

by Theorem 2.1. Here we apply the interchange rule to the function $(u,\omega) \mapsto f(x(\omega), u, \omega) - u \cdot y(\omega)$ which is a normal integrand, by [47, Proposition 14.45(c)].

Fix a $y \in L^q$ and let $x \in \mathcal{N}^r$ be such that $I_l(x,y) < \infty$. Let $\alpha > g(y)$ be arbitrary and let $x' \in \mathcal{N}$ be such that $El(x'(\omega), y(\omega), \omega) \le \alpha$. Defining $x^\nu_t = x'_t \chi_{A^\nu_t} + x_t \chi_{\bar{\Omega} \setminus A^\nu_t}$, where $A^\nu_t = \{\omega \,|\, |x'_t(\omega)| \le \nu\}$, we have that the strategy $x^\nu = (x^\nu_t)_{t=0}^T$ is in $\mathcal{N}^r$ and that $x^\nu_t \to x'_t$ almost surely for every $t = 0, \dots, T$ as $\nu \nearrow \infty$. Since the functions $\omega \mapsto l_t(x^\nu_t(\omega), y(\omega), \omega)$ are dominated by the integrable function

$$\omega \mapsto \max\{l_t(x'_t(\omega), y(\omega), \omega), l_t(x_t(\omega), y(\omega), \omega), 0\},$$

[1] Throughout this paper, we define $\infty - \infty = +\infty$.

Fatou's lemma (applied in the product measure space $\Omega \times \{0,\dots,T\}$ obtained by equipping $\{0,\dots,T\}$ with the counting measure) gives

$$\begin{aligned}\limsup E\sum_{t=0}^{T} l_t(x^\nu_t(\omega),y(\omega),\omega) &\le E\sum_{t=0}^{T}\limsup l_t(x^\nu_t(\omega),y(\omega),\omega)\\ &= E\sum_{t=0}^{T} l_t(x'_t(\omega),y(\omega),\omega)\le\alpha.\end{aligned}$$

Since $\alpha > g(y)$ was arbitrary and $x^\nu \in \mathcal{N}^r$, the claim follows. □

The main content of the first part of Theorem 2.2 is that the infimum in the definition of the Lagrangian can be reduced to scenariowise minimization. This can sometimes be done even analytically. The last part of the above result shows that, while integrability of $x$ may be restrictive in the original problem, it may be harmless in the expression for the dual objective $g$. A simple example will be given at the end of Example 3.1 below. In some applications, the integrability can be used to derive more convenient expressions for $g$.

**3. Examples from operations research** This section reviews some well-known duality frameworks from operations research and shows how they can be derived from the abstract framework above. Many of the examples are from Rockafellar and Wets [43, 45] where they were formulated for bounded strategies. We will also point out some connections with more recent developments in finance and stochastics. A recent account of techniques and models of stochastic programming can be found in Shapiro, Dentcheva and Ruszczyski [52].

The best known duality frameworks involve functional constraints and Lagrange multipliers. The most classical example is linear programming duality. These frameworks are deterministic special cases of the following stochastic programming framework from [45], where sufficient conditions were given for the attainment of the dual optimum.

Example 3.1 (Inequality constraints) Let

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

where $f_j$ are convex normal integrands. To verify that $f$ is a normal integrand, we write it as $f = f_0 + \sum_{j=1}^m \delta_{C_j}$, where

$$\delta_{C_j}(x,u,\omega)=\begin{cases}0 & \text{if } (x,u)\in C_j(\omega),\\ +\infty & \text{otherwise}\end{cases}$$

and $C_j(\omega) = \{(x,u)\,|\, f_j(x,\omega)+u_j \le 0\}$. By [47, Proposition 14.33], the sets $C_j$ are measurable so the functions $\delta_{C_j}$ are normal integrands by [47, Example 14.32] and then $f$ is a normal integrand by [47, Proposition 14.44(c)]. The integral functional $I_f$ is thus well-defined and equals

$$I_f(x,u)=\begin{cases} Ef_0(x(\omega),\omega) & \text{if } f_j(x(\omega),\omega)+u(\omega)\le 0 \ P\text{-a.s. } j=1,\dots,m,\\ +\infty & \text{otherwise.}\end{cases}$$

The primal problem (3) can be written as

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

This is the classical formulation of a nonlinear stochastic optimization problem. It is a stochastic extension of classical mathematical programming models such as linear programming.

The Lagrangian integrand becomes

$$\begin{aligned} l(x,y,\omega) &= \inf_{u\in\mathbb{R}^m}\{f(x,u,\omega)-u\cdot y\}\\ &= \begin{cases} +\infty & \text{if } f_j(x,\omega)=\infty \text{ for some } j,\\ f_0(x,\omega)+y\cdot F(x,\omega) & \text{if } f_j(x,\omega)<\infty \text{ and } y\ge 0,\\ -\infty & \text{otherwise,}\end{cases}\end{aligned}$$

where $F(x,\omega)=(f_1(x,\omega),\dots,f_m(x,\omega))$. The expression $g(y)=\inf_{x\in\mathcal{N}} I_l(x,y)$ holds under the general condition of Theorem 2.2, but to get more explicit expressions for the dual objective $g$ one needs more structure on $f$; see the examples below.

To illustrate how the choice of the strategy space may affect the lower semicontinuity of $\varphi$, consider the case $n=m=1$, $f_0=0$ and

$$f_1(x,u,\omega)=a(\omega)x+u,$$

for some strictly positive $a$ such that $1/a\notin L^1$. We get $\varphi(u)=0$ for every $u\in L^p$ but there is no $x\in\mathcal{N}^1$ which satisfies the pointwise constraint when $\operatorname{ess\,inf} u>0$. However,

$$l(x,y,\omega)=\begin{cases} ya(\omega)x & \text{if } y\ge 0,\\ -\infty & \text{otherwise,}\end{cases}$$

so, by the second part of Theorem 2.2, the strategies can be taken even bounded when calculating $g$.

It was observed in [45, Section 3A] that the dual objective in Example 3.1 can be written in a more concrete form when the functions $f_j$ have a time-separable form.

Example 3.2 Consider Example 3.1 in the case

$$f_j(x,\omega)=\sum_{t=0}^T f_{j,t}(x_t,\omega),$$

where each $f_{j,t}$ is an $\mathcal{F}_t$-measurable normal integrand. Defining $F_t(x_t,\omega)=(f_{1,t}(x_t,\omega),\dots,f_{m,t}(x_t,\omega))$ and using the convention $\infty-\infty=+\infty$, we can write

$$l(x,y,\omega)=\sum_{t=0}^T l_t(x_t,y,\omega),$$

where

$$l_t(x_t,y,\omega)=\begin{cases} +\infty & \text{if } f_{j,t}(x_t,\omega)=\infty \text{ for some } j,\\ f_{0,t}(x_t,\omega)+y\cdot F_t(x_t,\omega) & \text{if } f_{j,t}(x_t,\omega)<\infty \text{ and } y\ge 0,\\ -\infty & \text{otherwise.}\end{cases}$$

Assume now that $F_t(x_t,\cdot)\in L^p$ for every $t$ and $x_t\in\mathbb{R}^{n_t}$ and that there is a $v\in\mathcal{N}^p$ and a $p$-integrable random variable $w$ such that $f_{j,t}(x,\omega)\ge v_t(\omega)\cdot x-w(\omega)$. It follows that $F(x(\cdot),\cdot)\in L^p$ for every $x\in\mathcal{N}^\infty$; see e.g. [42, Theorem 3K][2]. If there is an $x\in\mathcal{N}^\infty$ such that $\omega\mapsto f_{0,t}(x_t(\omega),\omega)$ are integrable then, by the second part of Theorem 2.2,

$$g(y)=\inf_{x\in\mathcal{N}^\infty} E\sum_{t=0}^T l_t(x_t(\omega),y(\omega),\omega).$$

Using the properties of conditional expectation (see e.g. [53, Section II.7]), we get

$$\begin{aligned} g(y)&=\inf_{x\in\mathcal{N}^\infty} E\sum_{t=0}^T E_t l_t(x_t(\omega),y(\omega),\omega)\\ &=\inf_{x\in\mathcal{N}^\infty} E\sum_{t=0}^T l_t(x_t(\omega),(E_t y)(\omega),\omega).\end{aligned}$$

Applying Theorem 2.1 for $t=0,\dots,T$, we can express the dual objective as

$$g(y)=E\sum_{t=0}^T g_t((E_t y)(\omega),\omega),$$

where

$$g_t(y,\omega)=\inf_{x_t\in\mathbb{R}^{n_t}} l_t(x_t,y,\omega).$$

[2]If $\|x\|_{L^\infty}\le r$, there is a finite set of points $x^i\in\mathbb{R}^J$ $i=1,\dots,n$ whose convex combination contains the ball $r\mathbb{B}$. By convexity, $f_{j,t}(z(\omega),\omega)\le\sup_{i=1,\dots,n} f_{j,t}(x^i,\omega)$, where the right hand side is $p$-integrable by assumption. Combined with the lower bound, we then have $F_t(x(\cdot),\cdot)\in L^p$ as claimed.

The dual problem can thus be written as

$$\underset{y\in\mathcal{M}^p}{\text{maximize}} \quad \sum_{t=0}^T g_t(y_t(\omega),\omega),$$

where $\mathcal{M}^q$ is the set of $\mathbb{R}^m$-valued $q$-integrable martingales.

In the linear case, considered already in Danzig [12], the dual problem in Example 3.2 can be written as another linear optimization problem.

Example 3.3 (Linear programming) Consider Example 3.2 in the case where

$$f_{0,t}(x_t,\omega) = \begin{cases} a_{0,t}(\omega)\cdot x_t & \text{if } x\in\mathbb{R}^{n_t}_+,\\ +\infty & \text{otherwise}\end{cases}$$

and $f_{j,t}(x_t,\omega) = a_{j,t}(\omega)\cdot x_t + b_{j,t}(\omega)$ for $\mathcal{F}_t$-measurable $p$-integrable $n_t$-dimensional vectors $a_{j,t}$ and $\mathcal{F}_t$-measurable integrable scalars $b_{j,t}$. The primal problem can then be written as

$$\begin{aligned} &\text{minimize} && E\sum_{t=0}^T a_{0,t}(\omega)\cdot x_t(\omega) \quad \text{over } x\in\mathcal{N}_+\\ &\text{subject to} && \sum_{t=0}^T [A_t(\omega)x_t(\omega)+b_t(\omega)]\le 0 \quad P\text{-a.s.},\end{aligned}$$

where $A_t(\omega)$ is the matrix with rows $a_{j,t}(\omega)$ and $b_t(\omega) = (b_{j,t}(\omega))_{j=1}^m$. We get

$$\begin{aligned} l_t(x_t,y,\omega) &= \begin{cases} +\infty & \text{if } x_t\not\geq 0,\\ a_{0,t}(\omega)\cdot x_t + y\cdot[A_t(\omega)x_t+b_t(\omega)] & \text{if } x_t\ge 0,\ y\ge 0,\\ -\infty & \text{otherwise}\end{cases}\\ &= \begin{cases} +\infty & \text{if } x_t\not\geq 0,\\ [A_t^*(\omega)y+a_{0,t}(\omega)]\cdot x_t + y\cdot b_t(\omega) & \text{if } x_t\ge 0,\ y\ge 0,\\ -\infty & \text{otherwise,}\end{cases}\end{aligned}$$

where $A_t^*(\omega)$ is the transpose of $A_t(\omega)$. It follows that

$$g_t(y,\omega) = \begin{cases} y\cdot b_t(\omega) & \text{if } y\ge 0 \text{ and } A_t^*(\omega)y + a_{0,t}(\omega)\ge 0,\\ -\infty & \text{otherwise}\end{cases}$$

and the dual problem can be written as

$$\begin{aligned} &\text{maximize} && E\sum_{t=0}^T b_t(\omega)\cdot y_t(\omega) \quad \text{over } y\in\mathcal{M}^q_+\\ &\text{subject to} && A_t^*(\omega)y_t(\omega)+a_{0,t}(\omega)\ge 0 \quad P\text{-a.s. } t=0,\dots,T,\end{aligned}$$

where $\mathcal{M}^q_+$ is the set of nonnegative $q$-integrable martingales. When $T=0$, we recover the classical linear programming duality framework.

The famous problem of optimal stopping is a one-dimensional special case of Example 3.3.

Example 3.4 (Optimal stopping) The optimal stopping problem with an adapted integrable nonnegative scalar process $Z$ can be formulated as

$$\underset{x\in\mathcal{N}_+}{\text{maximize}} \quad E\sum_{t=0}^T x_t Z_t \quad \text{subject to} \quad \sum_{t=0}^T x_t\le 1,\ x_t\in\{0,1\}\ P\text{-a.s.}$$

The feasible strategies $x$ are related to stopping times through $\tau(\omega)=\inf\{t\,|\,x_t(\omega)=1\}$. The optimal value is not affected if we relax the constraint $x_t\in\{0,1\}$ (see below). The relaxed problem fits the framework of Example 3.3 with $n_t = m = 1$, $p = \infty$, $a_{0,t}(\omega) = -Z_t(\omega)$, $a_{1,t}(\omega) = 1$ and $b_{1,t}(\omega) = -1/(T+1)$. The dual problem becomes

$$\underset{y\in\mathcal{M}^1}{\text{minimize}} \quad Ey_0 \quad \text{subject to} \quad y \geq Z \quad P\text{-a.s.},$$

where $\mathcal{M}^1$ is the space of martingales.

To justify the convex relaxation, we first note that the feasible set of the relaxed problem is contained in the space $\mathcal{N}^\infty$ of bounded strategies. Since $Z \in \mathcal{N}^1$ by assumption, it suffices (by the Krein-Millman theorem) to show that the feasible set of the relaxed problem equals the $\sigma(\mathcal{N}^\infty, \mathcal{N}^1)$-closed convex hull of the feasible set of the original problem. Let $x$ be feasible in the relaxed problem. For $\nu = 1, 2, \ldots$, define the stopping times

$$\tau^{\nu,i}(\omega) = \inf\{s \,|\, X_s(\omega) \geq i/\nu\} \quad i = 1\ldots,\nu,$$

where $X_s(\omega) = \sum_{t=0}^s x_t(\omega)$. The strategies

$$x_t^{\nu,i}(\omega) = \begin{cases} 1 & \text{if } \tau^{\nu,i}(\omega) = t, \\ 0 & \text{otherwise} \end{cases}$$

are feasible in the original problem. It suffices to show that the convex combinations

$$x^\nu(\omega) = \sum_{i=1}^{\nu} \frac{1}{\nu} x^{\nu,i}(\omega)$$

converge to $x$ in the weak topology. By construction,

$$X_s^\nu(\omega) := \sum_{t=0}^{s} x_t^\nu(\omega) = \sup\{i \,|\, i/\nu \leq X_s(\omega)\} \in [X_t(\omega) - 1/\nu, X_t(\omega)],$$

so that $X_t^\nu \to X_t$ and thus $x_t^\nu \to x_t$ in the $L^\infty$-norm.

Remark 3.1 *The above duality frameworks suggest computational techniques for estimating the optimal value of the primal problem. The dual objective in Example 3.2 is dominated for every $y \in \mathcal{M}^q_+$ by*

$$\tilde{g}(y) := E \inf_{x\in\mathbb{R}^n} \left\{ \sum_{t=0}^{T} [f_t(x_t,\omega) + y_t(\omega)\cdot F_t(x_t,\omega)] \,\middle|\, \sum_{t=0}^{T} F_t(x_t,\omega) \leq 0 \right\}. \tag{5}$$

*If $x' \in \mathcal{N}$ is feasible in the primal problem, we get for every $y \in \mathcal{M}^q_+$*

$$\begin{aligned} \tilde{g}(y) &\leq E\sum_{t=0}^{T}[f_t(x'_t(\omega),\omega) + y_t(\omega)\cdot F_t(x'_t(\omega),\omega)] \\ &= E\left\{\sum_{t=0}^{T} f_t(x'_t(\omega),\omega) + y_T(\omega)\cdot\sum_{t=0}^{T} F_t(x'_t(\omega),\omega)\right\} \\ &\leq E\sum_{t=0}^{T} f_t(x'_t(\omega),\omega). \end{aligned}$$

*Minimizing over all feasible strategies $x' \in \mathcal{N}$ shows that (5) lies between $g(y)$ and the optimum primal value $\varphi(0)$. When $\varphi$ is closed, we thus get that $\varphi(0) = \sup_{y\in\mathcal{M}^q_+} \tilde{g}(y)$. The problem of finding the infimum in (5) can be seen as a deterministic version of the primal problem augmented by a penalty term in the objective.*

*In the case of Example 3.4, (5) can be written for every $y \in \mathcal{M}^q_+$ as*

$$\begin{aligned} \tilde{g}(y) &= E\inf_{x\in\mathbb{R}^n}\left\{\sum_{t=0}^{T}[-Z_t x_t + y_t(x_t - 1/(T+1))] \,\middle|\, \sum_{t=0}^{T} x_t \leq 1\right\} \\ &= E\inf_{x\in\mathbb{R}^n}\left\{\sum_{t=0}^{T}[(y_t - Z_t)x_t - y_0] \,\middle|\, \sum_{t=0}^{T} x_t \leq 1\right\} \\ &= E\min_{t=0,\ldots,T}(y_t - y_0 - Z_t). \end{aligned}$$

*This is the dual representation for optimal stopping obtained by Davis and Karatzas [15]. This was used by Rogers [48] (see also Haugh and Kogan [21]) in a simulation based technique for computing upper bounds for the value of American options in complete market models. The technique is readily extended to the more general problem class of Example 3.2. The technique can be further extended using the following.*

The cost of the nonanticipativity constraint on the strategies has been studied in a number of papers; see e.g. Rockafellar and Wets [43] for a general discrete finite time framework as well as Wets [55], Back and Pliska [2], Davis [13] and Davis and Burnstein [14] on continuous-time models. The cost can be described in terms of dual variables representing the value of information. The following derives a dual representation in the framework of Section 2.

Example 3.5 (Shadow price of information) Let $h$ be a convex normal integrand and consider the problem

$$\operatorname*{minimize}_{x\in\mathcal{N}} \quad I_h(x). \tag{6}$$

This can be seen as the primal problem associated with the normal integrand

$$f(x,u,\omega)=h(x+u,\omega).$$

The value function $\varphi(u)$ corresponds to changing the information structure in (6) by adding a general $\mathcal{F}_T$-measurable vector $u_t$ to each $x_t$. We get

$$\begin{aligned}
l(x,y,\omega) &= \inf_{u\in\mathbb{R}^m}\{h(x+u,\omega)-u\cdot y\}\\
&= \inf_{w\in\mathbb{R}^m}\{h(w,\omega)-\sum_{t=0}^T(w_t-x_t)\cdot y_t\}\\
&= \sum_{t=0}^T x_t\cdot y_t - \sup_{w\in\mathbb{R}^m}\{\sum_{t=0}^T w_t\cdot y_t - h(w,\omega)\}\\
&= \sum_{t=0}^T x_t\cdot y_t - h^*(y,\omega).
\end{aligned}$$

As long as $I_h\not\equiv+\infty$, the conditions of Theorem 2.2 are satisfied with $r=p$. Indeed, we get $I_{h^*}(y)\ge I_h^*(y)>-\infty$ for every $y\in L^q$ so that $\inf_{x\in\mathcal{N}^p} I_l(x,y)\le I_l(0,y)=-I_{h^*}(y)<\infty$. Theorem 2.2 thus gives

$$\begin{aligned}
g(y) &= \inf_{x\in\mathcal{N}^p} E\{\sum_{t=0}^T x_t(\omega)\cdot y_t(\omega)-h^*(y(\omega),\omega)\}\\
&= \begin{cases}-I_{h^*}(y) & \text{if } y\perp\mathcal{N}^p,\\ -\infty & \text{otherwise,}\end{cases}
\end{aligned}$$

which is essentially the dual objective from [43, Section 4]. The dual representation (2) then gives

$$\begin{aligned}
(\operatorname{cl}\varphi)(0) &= \sup_{y\perp\mathcal{N}^p} -I_{h^*}(y)\\
&= \sup_{y\perp\mathcal{N}^p} E\inf_{x\in\mathbb{R}^n}\{h(x,\omega)-\sum_{t=0}^T x_t\cdot y_t(\omega)\}.
\end{aligned}$$

The infimum in the last expression differs from the original problem in that the information constraint has been replaced by a linear term. This gives the dual variable $y$ the interpretation as a "shadow price of information"; see [55, 43, 2, 13, 14] for further discussion.

The above expression for $(\operatorname{cl}\varphi)(0)$ can be used to compute lower bounds for the optimal value using simulation much like in Rogers [48] and Haugh and Kogan [21] in the case of optimal stopping problems; see Remark 3.1. Rockafellar and Wets [43] gave sufficient conditions for the existence of a $y\in\mathcal{M}^1$ such that $\varphi(0)=g(y)$ in the case of bounded strategies; see also Back and Pliska [2] for a continuous-time framework with a special class of objective functions.

The following problem format is adapted from Rockafellar and Wets [46]. It has its roots in calculus of variations and optimal control; see Rockafellar [40].

Example 3.6 (Problems of Bolza type) Let $n_t = d$ and consider the problem

$$\underset{x\in\mathcal{N}}{\text{minimize}} \quad E\sum_{t=0}^{T} L_t(x_t(\omega), \Delta x_t(\omega), \omega), \tag{7}$$

where $\Delta x_t := x_t - x_{t-1}$, $x_{-1} := 0$ and each $L_t$ is an $\mathcal{F}_t$-measurable normal integrand on $\mathbb{R}^d \times \mathbb{R}^d \times \Omega$. This fits our general framework with

$$f(x,u,\omega) = \sum_{t=0}^{T} L_t(x_t, \Delta x_t + u_t, \omega),$$

where $x_{-1} := 0$ and $u = (u_0, \dots, u_T)$ with $u_t \in \mathbb{R}^d$. Indeed, (7) is (1) with $u = 0$. We get

$$\begin{aligned} l(x,y,\omega) &= \inf_{u\in\mathbb{R}^m} \sum_{t=0}^{T} [L_t(x_t, \Delta x_t + u_t, \omega) - u_t \cdot y_t] \\ &= \inf_{u'\in\mathbb{R}^m} \sum_{t=0}^{T} [L_t(x_t, u'_t, \omega) - (u'_t - \Delta x_t)\cdot y_t] \\ &= \sum_{t=0}^{T} [\Delta x_t \cdot y_t - H_t(x_t, y_t, \omega)] \\ &= \sum_{t=0}^{T} [-x_t \cdot \Delta y_{t+1} - H_t(x_t, y_t, \omega)], \end{aligned}$$

where $y_{T+1} := 0$ and $H_t$ is the *Hamiltonian* defined by

$$H_t(x_t, y_t, \omega) = \sup_{u_t\in\mathbb{R}^d} \{u_t \cdot y_t - L_t(x_t, u_t, \omega)\}.$$

Thus

$$g(y) = \inf_{x\in\mathcal{N}} E\sum_{t=0}^{T} \left[\Delta x_t \cdot y_t - H_t(x_t, y_t)\right].$$

By Jensen's inequality, $g(y) \le g(\pi y)$ where $\pi$ denotes the projection $(y_t)_{t=0}^T \mapsto (E_t y_t)_{t=0}^T$. Consequently, when maximizing $g$, we do not loose anything if we restrict $y$ to the space $\mathcal{N}^q$ of adapted $q$-integrable processes. Moreover, if $u \in \mathcal{N}^p$ we have $\langle u, y\rangle = \langle u, \pi y\rangle$ and thus

$$(\operatorname{cl}\varphi)(u) = \sup_{y\in L^q} \{\langle u,y\rangle + g(y)\} = \sup_{y\in\mathcal{N}^q} \{\langle u,y\rangle + g(y)\}.$$

Assume now that there is an $x \in \mathcal{N}^p$ such that $EH_t(x_t, y_t) < \infty$ and that $(x_t,\omega) \mapsto -H_t(x_t, y_t(\omega), \omega)$ are $\mathcal{F}_t$-measurable normal integrands for every $y \in \mathcal{N}^q$. We then get from Theorem 2.2, the law of iterated expectations (see e.g. [53, Section II.7]) and Theorem 2.1 that for every $y \in \mathcal{N}^q$

$$\begin{aligned} g(y) &= \inf_{x\in\mathcal{N}^p} E\sum_{t=0}^{T} [-x_t \cdot \Delta y_{t+1} - H_t(x_t, y_t)] \\ &= \inf_{x\in\mathcal{N}^p} E\sum_{t=0}^{T} [-x_t \cdot E_t[\Delta y_{t+1}] - H_t(x_t, y_t)] \\ &= E\sum_{t=0}^{T} \inf_{x_t\in\mathbb{R}^{n_t}} [-x_t \cdot E_t[\Delta y_{t+1}] - H_t(x_t, y_t)] \\ &= -E\sum_{t=0}^{T} \sup_{x_t\in\mathbb{R}^{n_t}} \sup_{u_t\in\mathbb{R}^d} [x_t \cdot E_t[\Delta y_{t+1}] + u_t \cdot y_t - L_t(x_t, u_t)] \\ &= -E\sum_{t=0}^{T} L_t^*(E_t[\Delta y_{t+1}], y_t). \end{aligned}$$

The dual problem thus looks much like the primal except that the (forward) difference term enters the integral functional through the conditional expectation.

The above formulation of the Bolza problem was inspired by its continuous-time analogs. In the present discrete-time setting, the primal objective can be written as

$$\underset{x\in\mathcal{N}}{\text{minimize}}\quad E\sum_{t=0}^{T}\tilde{L}_t(x_t(\omega),x_{t-1}(\omega),\omega)$$

for the normal integrands $\tilde{L}_t(x_t,x_{t-i},\omega):=L(x_t,x_t-x_{t-1},\omega)$. This format covers the stochastic extensions of the von Neumann-Gale model studied e.g. in Dempster, Evstigneev and Taksar [17].

**4. Examples from mathematical finance** Convex duality has long been an integral part of mathematical finance. The case of American options was already discussed in Remark 3.1 above. Perhaps the most famous instance is the "fundamental theorem of asset pricing" which, in perfectly liquid market models, relates the existence of an arbitrage opportunity with that of an equivalent martingale measure for the underlying price process; see Delbaen and Schachermayer [16] for a comprehensive treatment of the perfectly liquid case and Kabanov and Safarian [25] for extensions to markets with proportional transaction costs. Other instances of convex duality can be found in problems of portfolio optimization or optimal consumption; see e.g. Cvitanik and Karatzas [10], Kramkov and Schachermayer [32] or Karatzas and Žitković [27]. Biagini [6] reviews utility maximization in perfectly liquid market models. Klein and Rogers [29] propose an abstract duality framework that unifies several earlier ones on optimal investment and consumption under market frictions. Several instances of convex duality in the financial context can be found in Föllmer and Schied [19] who give a comprehensive treatment of the classical perfectly liquid market model in finite discrete time; see Example 4.1 below.

We will show that, in finite discrete time, many duality frameworks in mathematical finance are instances of the abstract duality framework of Section 2. Moreover, our framework allows for various generalizations of existing financial models. We will study financial problems by following Kabanov [24] in that none of the assets is given the special role of a numeraire. Instead, all traded securities are treated symmetrically and contingent claims, consumption etc. take their values in the space of portfolios. This setting covers more traditional models where trading costs and claims are measured in cash; see Example 4.1 below.

Consider a market where $d$ securities are traded over finite discrete time $t=0,\dots,T$. At each time $t$ and state $\omega\in\Omega$, the market is described by two closed convex sets, $C_t(\omega)\subset\mathbb{R}^J$ and $D_t(\omega)\subset\mathbb{R}^J$ both of which contain the origin. The set $C_t(\omega)$ consists of the portfolios that are freely available in the market at time $t$ and $D_t(\omega)$ consists of the portfolios that the investor is allowed to hold over the period $[t,t+1)$. For each $t$, the sets $C_t$ and $D_t$ are assumed to be $\mathcal{F}_t$-measurable. If $C_t(\omega)$ are polyhedral cones and $D_t(\omega)\equiv\mathbb{R}^d$ (no portfolio constraints), we recover the model of [24]. In many applications, it is natural to assume that $\mathbb{R}_-\subseteq C_t(\omega)$ but this is not necessary for now.

A *contingent claim process (with physical delivery)* is a financial contract specified by an adapted $\mathbb{R}^J$-valued process $u=(u_t)_{t=0}^T$. At each $t=0,\dots,T$, the seller of the claim delivers a (possibly state dependent) portfolio $u_t$ to the buyer. Traditionally, financial mathematics has studied contingent claims that have only one payout date. This corresponds to $u_t=0$ for $t<T$. In real markets with portfolio constraints, it is important to distinguish between payments that occur at different points in time. We refer the reader to [34, 35] for further discussion of the topic in the case of claims with cash delivery.

A trading strategy $x\in\mathcal{N}$ *superhedges* a claim process $u\in\mathcal{N}$ if

$$\Delta x_t+u_t\in C_t,\ x_t\in D_t,\ t=0,\dots,T,\ x_T=0 \tag{8}$$

almost surely. Here and in what follows, we always set $x_{-1}=0$. Superhedging is the basis of many results in financial mathematics. Even though superhedging is not quite feasible in many practical situations, it turns out to be a useful notion in studying more realistic approaches based on risk preferences.

Example 4.1 (cash delivery) Most contingent claims in practice give payments in cash. If cash is represented by the asset indexed 0, a claim with cash delivery has $u_t=(u_t^0,0,\dots,0)$ for a scalar process $u^0=(u_t^0)_{t=0}^T$. In this case, it is convenient to specify the market model by

$$C_t(\omega)=\{(z^0,z)\in\mathbb{R}^d\,|\,z^0+S_t(z,\omega)\le 0\},$$

where the function $S_t(z,\omega)$ represents the cost in cash of buying a portfolio $z\in\mathbb{R}^{d-1}$ at time $t$ in state $\omega$. Such models are quite natural when studying securities markets where trades are settled in cash; see

[34, 35]. The set $C_t$ is $\mathcal{F}_t$-measurable as soon as $S_t$ is an $\mathcal{F}_t$-measurable normal integrand on $\mathbb{R}^{d-1} \times \Omega$. The budget constraint $\Delta x_t + u_t \in C_t$ can now be written as

$$\Delta z_t^0 + S_t(\Delta z_t) + u_t^0 \le 0.$$

If there are no constraints on $z^0$ (the position on the cash account), we can substitute out the variables $z_t^0$ for $t = 1, \dots, T$ to write the superhedging condition as

$$\sum_{t=0}^{T} u_t^0 + \sum_{t=0}^{T} S_t(\Delta z_t) \le z_0^0; \tag{9}$$

see [34, Example 3.1]. In this setting, there is no need to discriminate between payments at different points in time. If, moreover, the functions $S_t(\cdot, \omega)$ are linear so that $S_t(z, \omega) = s_t(\omega) \cdot z$ for an adapted price process $s = (s_t)$, we can rearrange terms in (9) to write it in terms of a "stochastic integral" as

$$\sum_{t=0}^{T} u_t^0 \le z_0^0 + \sum_{t=0}^{T} z_{t-1} \cdot \Delta s_t.$$

This is the traditional formulation of the superhedging problem; see e.g. [19] or [16]. It is based on the assumptions that instantaneous portfolio updates are costless and that the contingent claim gives payments in terms of an asset that can be held at unlimited positive as well as negative quantities. In practice, however, neither of these assumptions hold.

The following example gives a dual characterization of the set of claims that can be superhedged at zero cost.

Example 4.2 (Consistent price systems) The superhedging condition (8) can be studied in our general duality framework with $n_t = d$, $m = (T+1)d$ and

$$f(x, u, \omega) = \begin{cases} 0 & \text{if } \Delta x_t + u_t \in C_t(\omega),\ x_t \in D_t(\omega),\ x_T = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Indeed, we then get $\varphi = \delta_{\mathcal{C}}$, where

$$\mathcal{C} = \{u \in L^p \mid \exists x \in \mathcal{N} :\ \Delta x_t + u_t \in C_t,\ x_t \in D_t,\ x_T = 0\}$$

is the set of (not necessarily adapted) claim processes that can be superhedged at zero cost. This fits the framework of Example 3.6 with

$$L_t(x, u, \omega) = \begin{cases} 0 & \text{if } x \in D_t(\omega) \text{ and } u \in C_t(\omega), \\ +\infty & \text{otherwise,} \end{cases}$$

where $D_T(\omega) := \{0\}$. We get

$$\begin{aligned} H_t(x_t, y_t, \omega) &= \sup_{u_t \in \mathbb{R}^d} \{u_t \cdot y_t - L_t(x_t, u_t, \omega)\} \\ &= \begin{cases} \sigma_{C_t(\omega)}(y_t) & \text{if } x_t \notin D_t(\omega), \\ -\infty & \text{otherwise,} \end{cases} \end{aligned}$$

so the assumptions of Example 3.6 are satisfied. Since

$$L_t^*(v, y, \omega) = \sigma_{D_t(\omega)}(v) + \sigma_{C_t(\omega)}(y),$$

we get for every $y \in \mathcal{N}^q$

$$g(y) = -E \sum_{t=0}^{T} [\sigma_{D_t}(E_t \Delta y_{t+1}) + \sigma_{C_t}(y_t)],$$

where $\sigma_{D_T} = 0$. In the unconstrained case where $D_t = \mathbb{R}^d$ for $t = 0, \dots, T-1$, we have $\sigma_{D_t} = \delta_{\{0\}}$ so that $g(y) = -\infty$ unless $y$ is a martingale and thus, we recover [36, Lemma 4.3]. When $C$ and $D$ are conical, we have $g = -\delta_{\mathcal{D}}$, where

$$\mathcal{D} = \{y \in \mathcal{N}^q \mid E_t \Delta y_{t+1} \in D_t^*,\ y_t \in C_t^*\}$$

and $C_t^*(\omega)$ and $D_t^*(\omega)$ are the polar cones of $C_t(\omega)$ and $D_t(\omega)$ respectively. The elements of $\mathcal{D}$ are called *consistent price systems* for the market model $(C, D)$. The notion of a consistent price system was introduced in Kabanov [24] for the case of a polyhedral conical $C$ and $D \equiv \mathbb{R}^d$.

Much of trading in financial markets consists of exchanging sequences of cash-flows. In a typical situation, one exchanges a claim process $u \in \mathcal{N}^p$ for a multiple of another claim process $p \in \mathcal{N}^p$ – the *premium process*. Traditionally, financial mathematics has been mainly concerned with the special case where $p_t = 0$ for $t > 0$ and $u_t = 0$ for $t < T$. The best known application of this special setting is the pricing of European options. Due to portfolio constraints, however, premiums as well as claims are often paid over multiple points in time. Examples include swap contracts as well as various insurance contracts where premium payments are made throughout the life of the contract.

The *superhedging cost* of a claim process $u \in \mathcal{N}^p$ in terms of a premium process $p \in \mathcal{N}^p$ is defined as

$$\varphi(u) = \inf\{\alpha \,|\, u - \alpha p \in \mathcal{C}\},$$

where $\mathcal{C}$ is the set of claim processes that can be superhedged with zero cost; see Example 4.2. The special case where claims and premiums are paid in cash has been studied in [35]. The following addresses the general case of "physical delivery".

Example 4.3 (Pricing by superhedging) The superhedging cost is the value function in our general framework with

$$f(x,u,\omega) = \begin{cases} \alpha & \text{if } \Delta z_t + u_t - \alpha p_t \in C_t(\omega),\ z_t \in D_t(\omega),\ z_T = 0 \\ +\infty & \text{otherwise,} \end{cases}$$

where $x_0 = (z_0, \alpha)$ and $x_t = z_t$ for $t = 1, \dots, T$. We have assumed for simplicity that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ so that $z_0$ is deterministic. Alternatively, we could introduce a new decision stage at time $t = -1$ with $x_{-1} = \alpha$ and $\mathcal{F}_{-1} = \{\emptyset, \Omega\}$. We get

$$\begin{aligned} l(x,y,\omega) &= \inf_{u \in \mathbb{R}^m} \{f(x,u,\omega) - u \cdot y\} \\ &= \inf_{u \in \mathbb{R}^m} \{\alpha - u \cdot y \,|\, \Delta z_t + u_t - \alpha p_t \in C_t(\omega),\ z_t \in D_t(\omega),\ z_T = 0\} \\ &= \inf_{z \in \mathbb{R}^m} \left\{ \alpha + \sum_{t=0}^T (\Delta z_t - w_t - \alpha p_t) \cdot y_t \,\middle|\, w_t \in C_t(\omega),\ z_t \in D_t(\omega),\ z_T = 0 \right\} \\ &= \begin{cases} \alpha + \sum_{t=0}^T (\Delta z_t - \alpha p_t) \cdot y_t - \sum_{t=0}^T \sigma_{C_t(\omega)}(y_t) & \text{if } z_t \in D_t(\omega) \text{ and } z_T = 0, \\ +\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} \alpha(1 - \sum_{t=0}^T p_t \cdot y_t) - \sum_{t=0}^{T-1} z_t \cdot \Delta y_{t+1} - \sum_{t=0}^T \sigma_{C_t(\omega)}(y_t) & \text{if } z_t \in D_t(\omega), \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

This satisfies the assumptions of Theorem 2.2 with $r = \infty$ so

$$\begin{aligned} g(y) &= \inf_{x \in \mathcal{N}^\infty} \left\{ E\left[ \alpha(1 - \sum_{t=0}^T p_t \cdot y_t) - \sum_{t=0}^{T-1} z_t \cdot \Delta y_{t+1} - \sum_{t=0}^T \sigma_{C_t(\omega)}(y_t) \right] \,\middle|\, z_t \in D_t \right\} \\ &= \begin{cases} \inf_{z \in \mathcal{N}^\infty} \left\{ -E\left[ \sum_{t=0}^{T-1} z_t \cdot E_t \Delta y_{t+1} + \sum_{t=0}^T \sigma_{C_t}(y_t) \right] \,\middle|\, z_t \in D_t \right\} & \text{if } E \sum_{t=0}^T p_t y_t = 1 \\ -\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} -E\left[ \sum_{t=0}^{T-1} \sigma_{D_t(\omega)}(E_t \Delta y_{t+1}) + \sum_{t=0}^T \sigma_{C_t(\omega)}(y_t) \right] & \text{if } E \sum_{t=0}^T p_t \cdot y_t = 1, \\ -\infty & \text{otherwise.} \end{cases} \end{aligned}$$

where the last equality comes from the interchange rule in Theorem 2.1. This expression corresponds to [34, Lemma 7.1] which addressed contingent claim processes with cash delivery. By Jensen's inequality,

$$E_t \sigma_{C_t(\omega)}(y_t) \geq \sigma_{C_t(\omega)}(E_t y_t),$$

so that $g(y) \leq g(\pi y)$, where $\pi$ denotes the "projection" $(y_t)_{t=0}^T \mapsto (E_t y_t)_{t=0}^T$. This implies that, for adapted claims $u \in \mathcal{N}^p$,

$$\begin{aligned} (\operatorname{cl} \varphi)(u) &= \sup_{y \in L^q} \{\langle u, y \rangle + g(y)\} \\ &= \sup_{y \in \mathcal{N}^q} \{\langle u, y \rangle + g(y)\}, \end{aligned}$$

where $\mathcal{N}^q$ denotes the set of *adapted* $p$-integrable processes $y = (y_t)_{t=0}^T$. This corresponds to [35, Theorem 10] on claims with cash delivery. Closedness conditions will be given in Corollary 6.1 below. If $C$ and $D$ are conical, the above formula can be written as

$$(\mathrm{cl}\,\varphi)(u) = \sup_{y\in\mathcal{D}} \left\{ E\sum_{t=0}^T u_t\cdot y_t \;\middle|\; E\sum_{t=0}^T p_t\cdot y_t = 1 \right\}.$$

where $\mathcal{D}$ is the set of consistent price systems defined in Example 4.2.

We end this section with a model of optimal consumption problems in the general illiquid market model.

Example 4.4 (Optimal consumption) Consider the problem

$$\begin{array}{ll} \underset{x,c\in\mathcal{N}}{\text{maximize}} & E\sum_{t=0}^T U_t(c_t) \\ \text{subject to} & \Delta x_t + c_t \in C_t, \quad x_t\in D_t \quad t=0,\dots,T, \end{array}$$

where $D_T := \{0\}$ and $U_t$ is an $\mathcal{F}_t$-measurable concave normal integrand on $\mathbb{R}^d\times\Omega$. This represents a problem of optimal consumption where possibly all traded assets can be directly consumed. To model situations where some of the assets cannot be consumed, one can set $U_t(c,\omega) = -\infty$ for $c$ outside of the feasible consumption set. Defining $\mathcal{C}$ as in Example 4.2, we can write the problem concisely as

$$\text{maximize} \quad E\sum_{t=0}^T U_t(c_t) \quad \text{over } c\in\mathcal{C}.$$

This is the primal problem of Example 3.6 in the case

$$L_t(x,v,\omega) = \begin{cases} \inf_{c_t\in\mathbb{R}^d}\{-U_t(c_t,\omega)\,|\,v_t + c_t\in C_t(\omega)\} & \text{if } x_t\in D_t(\omega), \\ +\infty & \text{otherwise.} \end{cases}$$

By [47, Proposition 14.47], $L_t$ is an $\mathcal{F}_t$-measurable normal integrand as soon as it is lower semicontinuous in $(x,v)$. Conditions for lower semicontinuity, in turn, can be obtained by pointwise application of [39, Theorem 9.2]. It is easily checked that

$$L_t^*(v,y,\omega) = \sigma_{D_t(\omega)}(v) + \sigma_{C_t(\omega)}(y) - U_t^*(y,\omega),$$

where

$$U_t^*(y,\omega) = \inf_{c\in\mathbb{R}^d}\{c\cdot y - U_t(c,\omega)\}$$

is the conjugate of $U_t$ in the concave sense. If $C$ and $D$ are conical, we get

$$g(y) = \begin{cases} E\sum_{t=0}^T U_t^*(y_t) & \text{if } y\in\mathcal{D}, \\ -\infty & \text{otherwise,} \end{cases}$$

where $\mathcal{D}$ is the set of consistent price systems defined in Example 4.2. The dual problem can then be written in the symmetric form

$$\text{maximize} \quad E\sum_{t=0}^T U_t^*(y_t) \quad \text{over } y\in\mathcal{D}.$$

The dual pair of optimization problems above can be seen as a generalization (in discrete time) of the optimal consumption duality framework of Karatzas and Žitković [27] where the numeraire asset was consumed in a perfectly liquid market model in continuous time.

**5. Some closedness criteria** Much of duality theory in convex analysis has been concerned with optimality conditions and the attainment of dual optimum. Dual attainment is equivalent to the subdifferentiability of the value function $\varphi$ at the origin, which in turn is implied by continuity; see [41, Section 7]. In operations research, several "constraint qualifications" have bee proposed to guarantee the continuity of $\varphi$ at the origin. Unfortunately, such conditions fail in many infinite dimensional applications. In order to get the mere absence of a duality gap, it is sufficient (as well as necessary) that $\varphi$ be proper and lower semicontinuous at the origin. In this section, we outline techniques for establishing the lower semicontinuity of $\varphi$ and the attainment of the primal optimum.

The traditional technique for achieving lower semicontinuity of $\varphi$ would be to introduce a topology on (an appropriate subspace of) $\mathcal{N}$, to show that $I_f$ is lower semicontinuous and to impose inf-compactness conditions on $I_f$ with respect to $x$; see e.g. [47, Theorem 1.17]. This is essentially the "direct method" in calculus of variations for verifying the existence of a solution to a minimization problem; see e.g. [1]. As long as the topology is strong enough to imply almost sure convergence of converging sequences, the lower semicontinuity of $I_f$ often follows from Fatou's lemma and pointwise lower semicontinuity of normal integrands. The inf-compactness property, on the other hand, is often obtained with Alaoglu-type arguments provided the topology is weak enough. In particular, $\varphi$ is $\sigma(L^\infty, L^1)$-closed when the feasible set is $L^\infty$-bounded locally uniformly in $u$. This result applies already to many problems arising in practice and, in particular, to the optimal stopping problem in Example 3.4.

In some applications, the compactness condition does not hold. In the convex setting, the following version of a theorem of Komlós' [30] can often be used as a substitute.

Lemma 5.1 (Komlós' theorem) *Let $(x^\nu)_{\nu=1}^\infty$ be a sequence in $L^0(\Omega, \mathcal{F}, P; \mathbb{R}^n)$ which is almost surely bounded in the sense that*

$$\sup_\nu |x^\nu(\omega)| < \infty \quad P\text{-a.s.}$$

*Then there is a sequence of convex combinations $\bar{x}^\nu \in \operatorname{co}\{x^\mu \mid \mu \geq \nu\}$ that converges almost surely to an $\mathbb{R}^n$-valued function.*

Proof. See Delbaen and Schachermayer [16] or Kabanov and Safarian [25]. □

Different versions of Komlós' theorem have long been used in calculus of variations; see e.g. Balder [3]. The results of this section can be seen as extensions of the closedness results of Schachermayer [50] who used Komlós' theorem to prove the fundamental theorem of asset pricing for the classical perfectly liquid market model of Example 4.1.

The almost sure boundedness in Lemma 5.1 can sometimes be obtained by pointwise application of classical finite-dimensional boundedness conditions on directions of recession. Given a convex set $C$, we will denote its *recession cone* by

$$C^\infty = \{z \mid x + \alpha z \in C,\ \forall x \in C,\ \alpha > 0\}.$$

By [39, Theorem 8.4], a closed convex set $C$ in a finite-dimensional space is bounded if and only if $C^\infty = \{0\}$. The proof of this result is based on the classical Bolzano-Weierstrass theorem on converging subsequences in finite-dimensional spaces. The following simple modification of [26, Lemma 2] generalizes the finite-dimensional Bolzano-Weierstrass theorem to the present stochastic setting.

Lemma 5.2 *For an almost surely bounded sequence $(x^\nu)_{\nu=1}^\infty$ in $\mathcal{N}$ there exists a strictly increasing sequence of $\mathcal{F}_T$-measurable integer-valued functions $(\tau^\nu)$ and an $x \in \mathcal{N}$ such that*

$$x^{\tau^\nu} \to x$$

*almost surely.*

Proof. Applying [26, Lemma 2] to $(x_0^\nu)_{\nu=1}^\infty$ we get an $\mathcal{F}_0$-measurable random subsequence $\tau_0^\nu$ such that $x_0^{\tau_0^\nu} \to x_0$ for an $x_0 \in L^0(\Omega, \mathcal{F}_0, P; \mathbb{R}^{n_0})$. Applying [26, Lemma 2] next to $(x_1^{\tau_0^\nu})_{\nu=1}^\infty$ we get an $\mathcal{F}_1$-measurable subsequence $\tau_1^\nu$ of $\tau_0^\nu$ such that $x_1^{\tau_1^\nu} \to x_1$ for an $x_1 \in L^0(\Omega, \mathcal{F}_1, P; \mathbb{R}^{n_1})$. Since $x_0^{\tau_0^\nu} \to x_0$ we also have $x_0^{\tau_1^\nu} \to x_0$. Extracting further subsequences similarly for $t = 2, \ldots, T$ we arrive at the conclusion. □

Sequences of the form $(x^{\tau^\nu})_{\nu=1}^\infty$ in the above lemma are called *random subsequences* of the original sequence $(x^\nu)_{\nu=1}^\infty$.

If $C : \Omega \rightrightarrows \mathbb{R}^n$ is a closed convex-valued $\mathcal{F}$-measurable mapping, then $C^\infty(\omega) := C(\omega)^\infty$ defines an $\mathcal{F}$-measurable mapping whose values are closed convex cones; see [47, Exercise 14.21]. The following result generalizes [39, Theorem 8.4] to stochastic models in finite discrete time. The proof follows the inductive argument in the proof of [36, Theorem 3.3] with some simplifications. Theorem 3.3 of [36] deals with Example 8 in the case $D_t \equiv \mathbb{R}^d$ and its proof builds on earlier techniques developed for conical models of financial markets; see e.g. [51] or [25].

THEOREM 5.1 *Let $C : \Omega \rightrightarrows \mathbb{R}^n$ be closed convex-valued and $\mathcal{F}$-measurable. Every sequence in the set $\mathcal{C} = \{x \in \mathcal{N} \,|\, x \in C \text{ a.s.}\}$ is almost surely bounded if and only if $\{x \in \mathcal{N} \,|\, x \in C^\infty \text{ a.s.}\} = \{0\}$.*

PROOF. If the recession condition fails, then $\mathcal{C}$ contains a half-line so it cannot be a.s. bounded. To prove the converse, we may assume that $0 \in C$ almost surely. Indeed, if $\mathcal{C}$ is empty, there is nothing to prove. Otherwise, we take any $x \in \mathcal{C}$, set $C(\omega) := C(\omega) - x(\omega)$ and note that the translation does not affect the recession cone of $C$ or the almost sure boundedness of $\mathcal{C}$.

We use induction on $T$. Let $T > 0$ and assume first that the claim holds for every $(T-1)$-period model. Let $(x^\nu)_{\nu=1}^\infty \subset \mathcal{C}$ and consider the following two complementary cases.

Case 1: $\rho(\omega) := \sup |x_0^\nu(\omega)| < \infty$ almost surely. Let

$$\mathcal{N}_1 := \{(x_1, \dots, x_T) \,|\, x_t \in L^0(\Omega, \mathcal{F}_t, P; \mathbb{R}^{n_t})\},$$
$$C_1(\omega) := \{(x_1, \dots, x_T) \,|\, \exists x_0 \in \rho(\omega)\mathbb{B} : \ (x_0, \dots, x_T) \in C(\omega)\}.$$

By Proposition 14.11(a) and Proposition 14.13(a) of [47], $C_1$ is $\mathcal{F}$-measurable and, by [39, Theorem 9.1], it is closed and convex-valued with

$$C_1^\infty(\omega) = \{(x_1, \dots, x_T) \,|\, (0, x_1, \dots, x_T) \in C^\infty(\omega)\}.$$

Our assumption thus implies that $\{x \in \mathcal{N}_1 \,|\, x \in C_1^\infty \ P\text{-a.s.}\} = \{0\}$ so the sequence $(x_1^\nu, \dots, x_T^\nu)$ is almost surely bounded by the induction hypothesis.

Case 2: the set $A = \{\omega \in \Omega \,|\, \sup |x_0^\nu(\omega)| = \infty\}$ has positive probability. Let $\alpha^\nu = \chi_A / \max\{|x_0^\nu|, 1\}$ and $\bar{x}^\nu = \alpha^\nu x^\nu$. Passing to an $\mathcal{F}_0$-measurable random subsequence if necessary, we may assume that $\alpha^\nu \searrow 0$ almost surely. Since $\alpha^\nu$ are $\mathcal{F}_0$-measurable, $\bar{x}^\nu \in \mathcal{N}$. We also have that

$$\bar{x}^\nu \in \alpha^\nu C$$

and $|\bar{x}_0^\nu| \le 1$ almost surely. Since $\alpha^\nu \le 1$ and $0 \in C$ we get $\alpha^\nu C \subset C$, by convexity. We are thus in the same situation as in case 1, so $(\bar{x}^\nu)_{\nu=1}^\infty$ is almost surely bounded. By Lemma 5.2, there is an $\mathcal{F}_T$-measurable subsequence $\tau^\nu$ such that $(\bar{x}^{\tau^\nu})_{\nu=1}^\infty$ converges almost surely to an $\bar{x} \in \mathcal{N}$. By [39, Theorem 8.2],

$$\bar{x} \in C^\infty$$

almost surely, so $\bar{x} = 0$ by the assumption. This contradicts the positivity of $P(A)$ since on $A$, we have $|\bar{x}_0^\nu(\omega)| \nearrow 1$ so that $|\bar{x}_0(\omega)| = 1$.

It remains to prove the claim for $T = 0$. This can be done as in Case 2 above except that now we do not need to refer to Case 1 for the boundedness of $(\bar{x}^\nu)_{\nu=1}^\infty$ (This is essentially the finite-dimensional argument in [39, Theorem 8.4]). □

With Theorem 5.1, we can generalize finite-dimensional closedness results to the present stochastic setting much like [39, Theorem 8.4] was used in Section 9 of [39]. Orthogonal projections, which are central in the arguments of [39, Section 9 ], are not well-defined in the space $\mathcal{N}$ but the following lemma (whose origins can be traced back to Schachermayer [50]) can be used instead.

LEMMA 5.3 *Let $N : \Omega \rightrightarrows \mathbb{R}^n$ be an $\mathcal{F}$-measurable mapping whose values are linear. For each $t = 0, \dots, T$, there is an $\mathcal{F}_t$-measurable linear-valued mapping $N_t : \Omega \rightrightarrows \mathbb{R}^{n_t}$ such that*

$$\{x_t \in L^0(\Omega, \mathcal{F}_t, P; \mathbb{R}^{n_t}) \,|\, \exists z \in \mathcal{N} : \ (0, \dots, 0, x_t, z_{t+1} \dots, z_T) \in N\}$$
$$= \{x_t \in L^0(\Omega, \mathcal{F}_t, P; \mathbb{R}^{n_t}) \,|\, x_t \in N_t \ P\text{-a.s.}\}.$$

PROOF. It suffices to prove the claim for $t = 0$ since otherwise we can replace $N$ by the set-valued mapping $\bar{N}(\omega) = \{(x_t, \dots, x_T) \,|\, (0, \dots, 0, x_t, \dots, x_T) \in N(\omega)\}$ which is also closed-valued and

$\mathcal{F}$-measurable, by Proposition 14.11(a) and Proposition 14.13(a) of [47]. For any $\mathcal{F}$-measurable closed-valued mapping $S:\Omega\rightrightarrows\mathbb{R}^k$, there is an $\mathcal{F}_t$-measurable mapping $\Gamma_t S$ whose $\mathcal{F}_t$-measurable selectors coincide with those of $S$.[3] We define $\tilde N_t$ recursively by $\tilde N_T=\Gamma_T N$ and

$$\tilde N_t=\Gamma_t P_t\tilde N_{t+1},$$

where $(P_t\tilde N_{t+1})(\omega):=\{(x_0,\dots,x_t)\,|\,\exists x_{t+1}\in\mathbb{R}^{n_{t+1}}:\ (x_0,\dots,x_{t+1})\in\tilde N_{t+1}(\omega)\}$. The mapping $P_t\tilde N_{t+1}$ is $\mathcal{F}_{t+1}$-measurable (see [47, Proposition 14.13(a)]) and linear-valued. The mapping $N_0=\tilde N_0$ has the claimed properties. Indeed, it is clear that if $x_0$ belongs to the set on the left then $x_0\in N_0$ almost surely. The reverse direction follows from repeated application of the theorem on measurable selections [47, Corollary 14.5]. □

The following can be seen as a generalization of [39, Theorem 9.1] to our stochastic setting.

THEOREM 5.2 *Let $C:\Omega\rightrightarrows\mathbb{R}^n\times\mathbb{R}^m$ be a closed convex-valued $\mathcal{F}$-measurable mapping such that $\{x\in\mathcal{N}\,|\,(x,0)\in C^\infty$ a.s.$\}$ is a linear space. Then the set*

$$\mathcal{C}=\{u\in L^p\,|\,\exists x\in\mathcal{N}:\ (x,u)\in C\ a.s.\}$$

*is $\sigma(L^p,L^q)$-closed.*

PROOF. Let $N=\{x\in\mathbb{R}^n\,|\,(x,0)\in C^\infty(\omega)\cap(-C^\infty(\omega))\}$ and define $N_t$ as in Lemma 5.3. We may assume that $x_t\in N_t$ almost surely in the definition of $\mathcal{C}$ so that

$$\mathcal{C}=\{u\in L^p\,|\,\exists x\in\mathcal{N}:\ x_t\in N_t^\perp,\ (x,u)\in C\ a.s.\}.$$

Indeed, let $x_0^0(\omega)$ be the pointwise orthogonal projection of $x_0(\omega)$ to $N_0(\omega)$. By definition of $N_t$, there is an extension $x^0\in\mathcal{N}$ of $x_0^0$ such that $(x^0,0)\in C^0$ almost surely. Defining $\tilde x^0=x-x^0$, we have $\tilde x_0^0\in N_0^\perp$ and $(\tilde x^0,u)\in C$ almost surely. Repeating the procedure for $t=1,\dots,T$ we arrive at an $\tilde x^T\in\mathcal{N}$ with $\tilde x_t^T\in N_t^\perp$ and $(\tilde x^T,u)\in C$ almost surely.

Assume first that $p=1$. Since $\mathcal{C}$ is convex it suffices to prove $\tau(L^1,L^\infty)$-closedness. Since $\tau(L^1,L^\infty)$ is the norm topology, it suffices to verify sequential closedness. So assume that $(u^\nu)_{\nu=1}^\infty\subset\mathcal{C}$ converges to a $u$ in norm and let $x^\nu\in\mathcal{N}$ be such that $x_t^\nu\in N_t^\perp$ and $(x^\nu,u^\nu)\in C$. Passing to a subsequence, we may assume that $u^\nu\to u$ almost surely, so that the measurable function $\rho(\omega):=\sup_\nu|u^\nu(\omega)|$ is almost surely finite. Each $(x^\nu,u^\nu)$ thus belongs to the set

$$\mathcal{C}_\rho=\{(x,u)\in\mathcal{N}\times L^0\,|\,(x,u)\in C_\rho\text{ a.s.}\},$$

where $C_\rho(\omega)=\{(x,u)\,|\,x_t\in N_t^\perp(\omega),\ u\in\rho(\omega)\mathbb{B},\ (x,u)\in C(\omega)\}$. By [39, Corollary 8.3.3],

$$C_\rho^\infty(\omega)=\{(x,0)\,|\,x_t\in N_t^\perp(\omega),\ (x,0)\in C^\infty(\omega)\},$$

so, by the linearity assumption,

$$\begin{aligned}\{(x,u)\in\mathcal{N}\times L^0\,|\,&(x,u)\in C_\rho^\infty\text{ a.s.}\}\\&=\{(x,0)\in\mathcal{N}\times L^0\,|\,x_t\in N_t^\perp,\ (x,0)\in C^\infty\cap(-C^\infty)\text{ a.s.}\}\\&=\{x\in\mathcal{N}\,|\,x_t\in N_t^\perp,\ x\in N\text{ a.s.}\}\times\{0\},\end{aligned}$$

which equals $\{0,0\}$. Indeed, by the definition of $N_0$ in Lemma 5.3, $x\in\mathcal{N}$ and $x\in N$ a.s. imply $x_0\in N_0$, which together with $x_0\in N_0^\perp$ gives $x_0=0$. Repeating the argument for $t=1,\dots,T$ gives $x=0$. By Theorem 5.1, the sequence $(x^\nu,u^\nu)_{\nu=1}^\infty$ is then almost surely bounded so, by Lemma 5.1, there is a sequence of convex combinations $(\bar x^\nu,\bar u^\nu)_{\nu=1}^\infty$ that converges almost surely to a point $(\bar x,\bar u)$. We have $\bar u\in\mathcal{C}$ since $C$ is convex and closed-valued and $\bar u=u$ since the original sequence $(u^\nu)_{\nu=1}^\infty$ was convergent to $u$.

Now let $p\in[1,\infty]$ be arbitrary. We have $\mathcal{C}=\{u\in L^p\,|\,Au\in\mathcal{C}^1\}$, where $\mathcal{C}^1$ denotes the set $\mathcal{C}$ in the case $p=1$ considered above and $A:(L^p,\sigma(L^p,L^q))\to(L^1,\sigma(L^1,L^\infty))$ is the natural injection. Since $A$ is continuous, the $\sigma(L^p,L^q)$-closedness of $\mathcal{C}$ follows from the $\sigma(L^1,L^\infty)$-closedness of $\mathcal{C}^1$. □

[3] Indeed, it suffices to check that the proof of [22, Theorem 3.1] (and the proofs of the lemmas used in it) goes through in the case $p=0$ with the norm replaced by the metric $d(x^1,x^2)=E\min\{|x^1(\omega)-x^2(\omega)|,1\}$. Alternatively, one can apply [9, Theorem 2.1] to the indicator function of $S$.

Since $C^\infty$ is a cone, the linearity condition in Theorem 5.2 can be expressed as

$$\{x \in \mathcal{N} \,|\, (x,0) \in C^\infty \ a.s.\} = \{x \in \mathcal{N} \,|\, (x,0) \in C^0 \ a.s.\},$$

where $C^0(\omega) = C^\infty(\omega) \cap (-C^\infty(\omega))$ is the *lineality space* of $C(\omega)$; see [39]. In applications where the set $C$ has more structure, one can use the calculus rules for recession cones given e.g. in Sections 8 and 9 of [39] to write the linearity condition in Theorem 5.2 in a more concrete form; see the examples below.

Theorem 5.2 can be seen as a lower-semicontinuity result for the value function $\varphi$ is situations where it takes the form of an indicator function. Theorem 5.1 and Lemma 5.3 allow the verification of the lower-semicontinuity of $\varphi$ in more general situations as well. This will be the subject of a separate article. We end this paper by showing how Theorem 5.2 yields some fundamental results in financial mathematics. An early application of recession analysis to portfolio optimization can be found in Bertsekas [5]. Korf [31] derives a version of the fundamental theorem of asset pricing for the classical perfectly liquid market model over a single period. The main result in [31] can be seen as an instance of Theorem 5.2 applied to Example 3.1 in the case where $T = 1$, $\mathcal{F}_0$ is the trivial sigma-field and $n_1 = 0$ (so that $\mathcal{N}$ is a finite-dimensional space).

**6. Applications to mathematical finance** Consider the market model studied in Section 4 and let (see Example 4.2)

$$\mathcal{C} = \{u \in L^p \,|\, \exists x \in \mathcal{N} : \ \Delta x_t + u_t \in C_t, \ x_t \in D_t, \ x_T = 0\}.$$

This is the set of (not necessarily adapted) claim processes that can be superhedged at zero cost. Its closedness is an important issue in mathematical finance; see [16] and [25] for comprehensive treatment of linear and conical market models without constraints. For example, the closedness is essential in proving various versions of the *fundamental theorem of asset pricing* that characterize certain arbitrage properties of a given market model. An example of this approach, initiated by Schachermayer [50], is given in Corollary 6.1 below. In this section, we consider claims with physical delivery, i.e. portfolio-valued claims, but similar (and often simpler) arguments apply to more traditional models on claims with cash-delivery.

The closedness of $\mathcal{C}$ is essential also for dual characterizations of superhedging costs. In particular, a straightforward adaptation of the closedness result in [35, Theorem 10] gives the following result for Example 4.3.

Proposition 6.1 *Consider the superhedging cost*

$$\varphi(u) = \inf\{\alpha \,|\, u - \alpha p \in \mathcal{C}\}$$

*associated with a market model $(C, D)$ and a premium process $p \in \mathcal{N}^p$. If $\mathcal{C}$ is closed and $\varphi(0) > -\infty$, then $\varphi$ is a closed convex function on $\mathcal{N}^p$.*

The above result allows us to express superhedging costs in terms of consistent price systems as explained in Example 4.3. This generalizes classical pricing formulas concerning claims with single payout date, perfectly liquid markets and premiums paid only at time $t = 0$; see [35] for further discussion in the case of claims with cash delivery. The condition $\varphi(0) > -\infty$ in Proposition 6.1 just means that the premium process $p$ is not freely available in the market at unlimited amounts; see [35, Section 4].

Theorem 5.2 gives the following result for the market model of Section 4. Here $C_t^\infty$ and $D_t^\infty$ denote the random sets $C_t(\omega)^\infty$ and $C_t(\omega)^\infty$, respectively.

Theorem 6.1 *Consider the market model of Section 4 and assume that*

$$\{x \in \mathcal{N} \,|\, \Delta x_t \in C_t^\infty, \ x_t \in D_t^\infty, \ x_T = 0 \ a.s.\}$$

*is a linear space. Then the set $\mathcal{C}$ is $\sigma(L^p, L^q)$-closed.*

Proof. This fits Theorem 5.2 with

$$C(\omega) = \{(x,u) \in \mathbb{R}^n \times \mathbb{R}^m \,|\, \Delta x_t + u_t \in C_t(\omega), \ x_t \in D_t(\omega), \ x_T = 0\}.$$

By [39, Corollary 8.3.3],

$$C^\infty(\omega) = \{(x,u) \in \mathbb{R}^n \times \mathbb{R}^m \,|\, \Delta x_t + u_t \in C_t(\omega)^\infty,\ x_t \in D_t(\omega)^\infty,\ x_T = 0\},$$

so the condition means that $\{x \in \mathcal{N} \,|\, (x,0) \in C^\infty \ a.s.\}$ is linear. □

Theorem 6.1 generalizes many earlier closedness results for discrete-time market models. The linearity condition holds, in particular, when

$$C_t^\infty(\omega) \cap D_t^\infty(\omega) = \{0\} \quad \forall t,\ \omega. \tag{10}$$

Indeed, since $x_{-1} = 0$ by definition, this implies that $\{x \in \mathcal{N} \,|\, \Delta x_t \in C_t^\infty,\ x_t \in D_t^\infty,\ x_T = 0\ a.s.\} = \{0\}$. Condition (10) is quite natural in practice. It means that portfolios that are freely available in the market at unlimited quantities become infeasible when scaled up by a large enough positive constant. Condition (10) was used for claims with cash delivery in [35, Theorem 6.3].

When there are no portfolio constraints, Theorem 6.1 reduces to [36, Theorem 3.3], which in turn, extends the closedness result of Schachermayer [51, Section 2]. Indeed, when there are no portfolio constraints, the linearity condition in Theorem 6.1 is implied by the so called *robust no scalable arbitrage* condition which generalizes the *robust no-arbitrage* condition introduced in [51] for certain conical models; see [36, Lemma 6.1]. Models with portfolio constraints have been studied in Kreher [33]. We illustrate the situation below without going to full detail.

A market model satisfies the *no-arbitrage* condition if

$$\mathcal{C} \cap \mathcal{N}_+ = \{0\}, \tag{11}$$

where $\mathcal{N}_+$ is the set of componentwise nonnegative adapted processes. The no-arbitrage condition means that it is not possible to superhedge nontrivial nonnegative claims by costless transactions in the financial market. In models without portfolio constraints, (11) can be expressed in a more conventional form in terms of terminal positions; see [36, Lemma 4.1]. In markets with portfolio constraints, however, the above formulation in terms of claims with multiple payout dates is more meaningful; see [34, 35] for a discussion in the case of claims with cash delivery.

It was shown in [50, Section 2] that, in the classical perfectly liquid market model, the no-arbitrage condition (in terms of cash-delivery) implies the closedness of the set of claims with cash delivery that can be superhedged at zero cost. Theorem 5.2 yields a simple proof of this important result. The following example (which is a special case of [36, Lemma 6.1]) illustrates the idea with contingent claims with physical delivery.

Example 6.1 (The no-arbitrage condition) Assume that $C_t(\omega)$ is a half-space (as e.g. in the classical perfectly liquid market model; see Example 4.1) and that there are no portfolio constraints, i.e. $D_t(\omega) = \mathbb{R}^d$ for every $t$ and $\omega$. The condition in Theorem 6.1 can then be written as

$$\{x \in \mathcal{N} \,|\, \Delta x_t \in C_t,\ x_T = 0\ a.s.\} = \{x \in \mathcal{N} \,|\, \Delta x_t \in C_t^0,\ x_T = 0\ a.s.\},$$

where $C_t^0(\omega) := C_t(\omega) \cap (-C_t(\omega))$. If this condition fails, there is an $\bar{x} \in \mathcal{N}$ such that $\Delta \bar{x}_t \in C_t$ and $\bar{x}_T = 0$ almost surely but $\Delta \bar{x}_t \notin C_t^0$ for some $t$ on a set $A \in \mathcal{F}_t$ of positive probability. Since $C_t(\omega)$ is a half-space, we have $C_t(\omega) \setminus C_t^0(\omega) = \operatorname{int} C_t(\omega)$. Given a nonzero vector $e \in \mathbb{R}^d_+$, the $\mathcal{F}_t$-measurable nonnegative random variable

$$\varepsilon(\omega) := \max\{\alpha \,|\, \Delta \bar{x}_t(\omega) + \alpha e \in C_t(\omega)\}$$

is thus strictly positive on $A$. We then have that $\bar{x}$ superhedges the nontrivial claim process defined by $u_t(\omega) = \varepsilon(\omega) e$ and $u_s = 0$ for $s \neq t$, so the no-arbitrage condition (11) cannot hold. The no arbitrage condition thus implies the closedness condition of Theorem 6.1.

The above argument extends to market models with transaction costs. In particular, the generalized no-arbitrage condition in [36] implies, by [36, Lemma 6.1], the linearity condition of Theorem 6.1 for models without portfolio constraints. Kreher [33, Lemma 30] gives conditions under which the linearity condition of Theorem 6.1 is satisfied in models with portfolio constraints.

We will next derive a version of the "fundamental theorem of asset pricing" using the above closedness result and the Kreps-Yan theorem. In our setting, the Kreps-Yan theorem can be stated as follows; see [23, 49] for more general formulations.

Theorem 6.2 (Kreps-Yan) *Let $K \subset \mathcal{N}^p$ be a closed convex cone such that $\mathcal{N}^p_- \subset K$ and $K \cap \mathcal{N}^p_+ = \{0\}$. Then there exists a $y \in \mathcal{N}^q$ such that $\langle u, y\rangle \le 0$ for every $u \in K$ and $\langle u, y\rangle > 0$ for every $u \in \mathcal{N}^p_+ \setminus \{0\}$.*

Schachermayer [50] used the Kreps-Yan theorem to prove the famous result of Dalang, Morton and Willinger [11], which gives the fundamental theorem of asset pricing for claims with cash delivery. The following applies to claims with physical delivery.

Corollary 6.1 (Fundamental theorem of asset pricing) *Assume that $C_t(\omega)$ are half-spaces containing the negative orthant $\mathbb{R}^d_-$ and that $D_t(\omega) \equiv \mathbb{R}^d$. The model has the no-arbitrage property if and only if there is a componentwise strictly positive martingale $y \in \mathcal{N}^q$ such that $y_t \in C^*_t$ almost surely for every $t$.*

Proof. It was shown in Example 6.1 that the no-arbitrage condition implies that $\mathcal{C}$ is closed. The Kreps-Yan theorem then gives the existence of a strictly positive $y \in \mathcal{N}^q$ such that $y \in \mathcal{C}^*$, where $\mathcal{C}^*$ is the polar cone of $\mathcal{C}$. It was shown in Example 4.2 that in conical market models,

$$\mathcal{C}^* = \{y \in \mathcal{N}^q \,|\, E_t\Delta y_{t+1} \in D^*_t,\ y_t \in C^*_t\}.$$

Since $D_t(\omega) = \mathbb{R}^d$, by assumption, we have $D^*_t(\omega) = \{0\}$ so $y$ is a martingale. On the other hand, if there exists a strictly positive $y \in \mathcal{C}^*$, every non-zero $u \in \mathcal{C} \cap \mathcal{N}_+$ would satisfy the contradicting inequalities $E(u \cdot y) \le 0$ and $E(u \cdot y) > 0$. □

The closedness result in Theorem 6.1 together with the Kreps-Yan theorem allow for extending Corollary 6.1 to market models with proportional transaction costs; see [36, Section 5] and the references therein.

Remark 6.1 (Equivalent martingale measures) *When $C_t(\omega)$ is a half-space, $C^*_t(\omega)$ is a half-line. In particular, in the classical perfectly liquid market model described in Example 4.1 where $C_t(\omega) = \{(z^0, z) \,|\, z^0 + s_t(\omega) \cdot z \le 0\}$, we have*

$$C^*_t(\omega) = \{(y^0, y) \,|\, y^0 \ge 0,\ y = y^0 s_t(\omega)\}.$$

*If $y^0 = (y^0_t)^T_{t=0}$ is the first component of the strictly positive process $y$ in Corollary 6.1, then $dQ/dP = y^0_T/Ey^0_T$ defines a measure $Q$ which is equivalent to $P$ and under which the price process $s = (s_t)^T_{t=0}$ is a martingale. Note however, that Corollary 6.1 requires that $s$ be componentwise nonnegative which is not needed in the classical result of [11]. On the other hand, the conclusion of Corollary 6.1 is stronger than that of the classical result which concerns claims with cash delivery. Similarly, while condition* (11) *is stronger than the classical no-arbitrage condition in terms of cash delivery, it yields the closedness of $\mathcal{C}$ (Example 11) which is more than the classical closedness result [50, Lemma 2.1] for claims with cash delivery.*

The above arguments are easily adapted to claims with cash-delivery. In particular, Theorem 8 can be used to generalize the closedness result in [35, Section 6], which is concerned with cash-delivery in illiquid markets.

## References

[1] H. Attouch, G. Buttazzo, and G. Michaille. *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*, volume 6 of *MPS/SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.

[2] K. Back and S. R. Pliska. The shadow price of information in continuous time decision problems. *Stochastics*, 22(2):151–186, 1987.

[3] E. J. Balder. Infinite-dimensional extension of a theorem of Komlós. *Probab. Theory Related Fields*, 81(2):185–188, 1989.

[4] E. M. L. Beale. On minimizing a convex function subject to linear inequalities. *J. Roy. Statist. Soc. Ser. B.*, 17:173–184; discussion, 194–203, 1955. (Symposium on linear programming.).

[5] D. P. Bertsekas. Necessary and sufficient conditions for existence of an optimal portfolio. *Journal of Economic Theory*, 8(2):235–247, 1974.

[6] S. Biagini. Expected utility maximization: the dual approach. In R. Cont, editor, *Encyclopedia of Quantitative Finance*. Wiley, to appear.

[7] 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.

[8] C. Castaing and M. Valadier. *Convex analysis and measurable multifunctions*. Springer-Verlag, Berlin, 1977. Lecture Notes in Mathematics, Vol. 580.

[9] C. Choirat, C. Hess, and R. Seri. A functional version of the birkhoff ergodic theorem for a normal integrand: A variational approach. *Annals of probability*, 31(1):63–92, 2003.

[10] J. Cvitanić and I. Karatzas. Convex duality in constrained portfolio optimization. *Ann. Appl. Probab.*, 2(4):767–818, 1992.

[11] R. C. Dalang, A. Morton, and W. Willinger. Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics Stochastics Rep.*, 29(2):185–201, 1990.

[12] G. B. Dantzig. Linear programming under uncertainty. *Management Sci.*, 1:197–206, 1955.

[13] M. H. A. Davis. Dynamic optimization: a grand unification. In *Proceedings of the 31st IEEE Conference on Decision and Control*, volume 2, pages 2035 – 2036, 1992.

[14] M. H. A. Davis and G. Burstein. A deterministic approach to stochastic optimal control with application to anticipative control. *Stochastics and Stochastics Reports*, 40(3&4):203–256, 1992.

[15] M. H. A. Davis and I. Karatzas. A deterministic approach to optimal stopping. In *Probability, statistics and optimisation*, Wiley Ser. Probab. Math. Statist. Probab. Math. Statist., pages 455–466. Wiley, Chichester, 1994.

[16] F. Delbaen and W. Schachermayer. *The Mathematics of Arbitrage*. Springer Finance. Springer-Verlag, Berlin Heidelberg, 2006.

[17] M. A. H. Dempster, I. V. Evstigneev, and M.I. Taksar. Asset pricing and hedging in financial markets with transaction costs: An approach based on the von neumanngale model. *Annals of Finance*, 2(4):327–355, 2006.

[18] I. Ekeland and R. Temam. *Convex analysis and variational problems*. North-Holland Publishing Co., Amsterdam, 1976. Translated from the French, Studies in Mathematics and its Applications, Vol. 1.

[19] H. Föllmer and A. Schied. *Stochastic finance*, volume 27 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2004. An introduction in discrete time.

[20] A. Grothendieck. *Topological vector spaces*. Gordon and Breach Science Publishers, New York, 1973. Translated from the French by Orlando Chaljub, Notes on Mathematics and its Applications.

[21] M. B. Haugh and L. Kogan. Pricing american options: a duality approach. *Oper. Res.*, 52(2):258–270, 2004.

[22] F. Hiai and H. Umegaki. Integrals, conditional expectations, and martingales of multivalued functions. *J. Multivariate Anal.*, 7(1):149–182, 1977.

[23] E. Jouini, C. Napp, and W. Schachermayer. Arbitrage and state price deflators in a general intertemporal framework. *J. Math. Econom.*, 41(6):722–734, 2005.

[24] Y. M. Kabanov. Hedging and liquidation under transaction costs in currency markets. *Finance and Stochastics*, 3(2):237–248, 1999.

[25] Y. M. Kabanov and M. Safarian. *Markets with transaction costs*. Springer Finance. Springer-Verlag, Berlin, 2009. Mathematical theory.

[26] Y. M. Kabanov and Ch. Stricker. A teachers' note on no-arbitrage criteria. In *Séminaire de Probabilités, XXXV*, volume 1755 of *Lecture Notes in Math.*, pages 149–152. Springer, Berlin, 2001.

[27] I. Karatzas and Žitković G. Optimal consumption from investment and random endowment in incomplete semimartingale markets. *The Annals of Probability*, 31(4):18211858, 2003.

[28] A. J. King. Duality and martingales: a stochastic programming perspective on contingent claims. *Math. Program.*, 91(3, Ser. B):543–562, 2002. ISMP 2000, Part 1 (Atlanta, GA).

[29] I. Klein and L. C. G. Rogers. Duality in optimal investment and consumption problems with market frictions. *Math. Finance*, 17(2):225–247, 2007.

[30] J. Komlós. A generalization of a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.*, 18:217–229, 1967.

[31] L. A. Korf. Stochastic programming duality: $\mathcal{L}^\infty$ multipliers for unbounded constraints with an application to mathematical finance. *Math. Program.*, 99(2, Ser. A):241–259, 2004.

[32] D. Kramkov and W. Schachermayer. The condition on the asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Annals of Applied Probability*, 9(3):904–950, 1999.

[33] D. Kreher. Hedging of portfolio-valued claims under convex transaction costs and portfolio constraints. Master's thesis, Humboldt-Universität zu Berlin, 2009.

[34] T. Pennanen. Arbitrage and deflators in illiquid markets. *Finance and Stochastics*, 15(1):57–83, 2011.

[35] T. Pennanen. Superhedging in illiquid markets. *Mathematical Finance*, to appear.

[36] T. Pennanen and I. Penner. Hedging of claims with physical delivery under convex transaction costs. *SIAM Journal on Financial Mathematics*, 1:158–178, 2010.

[37] S. R. Pliska. *Introduction to Mathematical Finance: Discrete Time Models*. Blackwell, Malden, MA, 1997.

[38] R. T. Rockafellar. Integrals which are convex functionals. *Pacific J. Math.*, 24:525–539, 1968.

[39] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[40] R. T. Rockafellar. Integrals which are convex functionals. II. *Pacific J. Math.*, 39:439–469, 1971.

[41] R. T. Rockafellar. *Conjugate duality and optimization*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974.

[42] 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.

[43] R. T. Rockafellar and R. J.-B. Wets. Nonanticipativity and $L^1$-martingales in stochastic optimization problems. *Math. Programming Stud.*, (6):170–187, 1976. Stochastic systems: modeling, identification and optimization, II (Proc. Sympos., Univ Kentucky, Lexington, Ky., 1975).

[44] R. T. Rockafellar and R. J.-B. Wets. Measures as Lagrange multipliers in multistage stochastic programming. *J. Math. Anal. Appl.*, 60(2):301–313, 1977.

[45] 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.

[46] 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.

[47] 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.

[48] L. C. G. Rogers. Monte Carlo valuation of American options. *Math. Finance*, 12(3):271–286, 2002.

[49] D. B. Rokhlin. The Kreps-Yan theorem for $L^\infty$. *Int. J. Math. Math. Sci.*, (17):2749–2756, 2005.

[50] W. Schachermayer. A Hilbert space proof of the fundamental theorem of asset pricing in finite discrete time. *Insurance Math. Econom.*, 11(4):249–257, 1992.

[51] W. Schachermayer. The fundamental theorem of asset pricing under proportional transaction costs in finite discrete time. *Math. Finance*, 14(1):19–48, 2004.

[52] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory*, volume 9 of *MPS/SIAM Series on Optimization*. SIAM, Philadelphia, 2009.

[53] A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1996. Translated from the first (1980) Russian edition by R. P. Boas.

[54] S. Tian and R. J.-B. Wets. Pricing contingent claims: a computational compatible approach. *Manuscript*, 2006.

[55] R. J-B Wets. On the relation between stochastic and deterministic optimization. In A. Bensoussan and J.L. Lions, editors, *Control Theory, Numerical Methods and Computer Systems Modelling*, volume 107 of *Lecture Notes in Economics and Mathematical Systems*, pages 350–361. Springer, 1975.