

Shadow price of information in discrete time stochastic optimization

Teemu Pennanen · Ari-Pekka Perkkiö

Received: date / Accepted: date

Dedicated to R. T. Rockafellar on his 80th Birthday

Abstract The shadow price of information has played a central role in stochastic optimization ever since its introduction by Rockafellar and Wets in the mid-seventies. This article studies the concept in an extended formulation of the problem and gives relaxed sufficient conditions for its existence. We allow for general adapted decision strategies, which enables one to establish the existence of solutions and the absence of a duality gap e.g. in various problems of financial mathematics where the usual boundedness assumptions fail. As applications, we calculate conjugates and subdifferentials of integral functionals and conditional expectations of normal integrands. We also give a dual form of the general dynamic programming recursion that characterizes shadow prices of information.

Keywords Convex stochastic optimization · Shadow price of information

1 Introduction

Let (Ω, \mathcal{F}, P) be a probability space with a filtration $(\mathcal{F}_t)_{t=0}^T$ and consider the multistage stochastic optimization problem

$$\text{minimize } Eh(x) \quad \text{over } x \in \mathcal{N}, \quad (\text{SP})$$

The second author is grateful to the Einstein Foundation for the financial support.

T. Pennanen
Department of Mathematics, King's College London
E-mail: teemu.pennanen@kcl.ac.uk

A.-P. Perkkiö
Department of Mathematics, Technische Universität Berlin

where $\mathcal{N} = \{(x_t)_{t=0}^T \mid x_t \in L^0(\Omega, \mathcal{F}_t, P; \mathbb{R}^{n_t})\}$ denotes the space of decision strategies adapted to the filtration, h is a convex normal integrand on $\mathbb{R}^n \times \Omega$ and

$$Eh(x) := \int_{\Omega} h(x(\omega), \omega) dP(\omega)$$

is the associated integral functional on $L^0 := L^0(\Omega, \mathcal{F}, P; \mathbb{R}^n)$. Here and in what follows, $n = n_0 + \dots + n_T$ and $L^0(\Omega, \mathcal{F}, P; \mathbb{R}^n)$ denotes the linear space of equivalence classes of \mathbb{R}^n -valued \mathcal{F} -measurable functions. As usual, two functions are equivalent if they are equal P -almost surely. Throughout, we define the expectation of a measurable function as $+\infty$ unless its positive part is integrable.

Problems of the form (SP) have been extensively studied since their introduction in the mid 70's; see [18, 19, 21]. Despite its simple appearance, problem (SP) is a very general format of stochastic optimization. Indeed, various pointwise (almost sure) constraints can be incorporated in the objective by assigning f the value $+\infty$ when the constraints are violated. Several examples can be found in the above references. Applications to financial mathematics are given in [9, 11, 10].

Our formulation of problem (SP) extends the formulation of [21], where Eh was minimized over the space

$$\mathcal{N}^\infty := \mathcal{N} \cap L^\infty$$

of essentially bounded adapted strategies. Here and in what follows, $L^\infty := L^\infty(\Omega, \mathcal{F}, P; \mathbb{R}^n)$. Allowing for general decision strategies $x \in \mathcal{N}$, we can relax many of the assumptions of [21] while still obtaining the existence of optimal strategies and their scenario-wise characterization as in [21].

Our approach is to analyze the value function $\phi : L^\infty \rightarrow \overline{\mathbb{R}}$,

$$\phi(z) := \inf_{x \in \mathcal{N}} Eh(x + z)$$

in the conjugate duality framework of Rockafellar [18]. Being the infimal projection of a convex function, ϕ is a convex function on L^∞ . Clearly $\phi(0)$ is the optimum value of (SP) while in general, $\phi(z)$ gives the optimum value that can be achieved in combination with an essentially bounded possibly nonadapted strategy z . *We assume throughout that $\phi(0)$ is finite and that Eh is proper on L^∞ .*

The space L^∞ is in separating duality with $L^1 := L^1(\Omega, \mathcal{F}, P; \mathbb{R}^n)$ under the bilinear form

$$\langle z, v \rangle := E(z \cdot v).$$

A $v \in L^1$ is said to be a *shadow price of information* for problem (SP) if it is a subgradient of ϕ at the origin, i.e., if

$$\phi(z) \geq \phi(0) + \langle z, v \rangle \quad \forall z \in L^\infty,$$

or, equivalently, if it solves the *dual problem*

$$\text{minimize } \phi^*(v) \quad \text{over } v \in L^1,$$

where

$$\phi^*(v) = \sup_{z \in L^\infty} \{\langle z, v \rangle - \phi(z)\}$$

is the *conjugate* of ϕ . Clearly, $\phi(0) + \inf_{v \in L^1} \phi^*(v) \geq 0$. If the inequality is strict, a *duality gap* is said to exist.

The following result, the proof of which is given in the appendix, shows that the shadow price of information has the same fundamental properties here as in [21] where strategies were restricted to \mathcal{N}^∞ . In particular, it shows that the dual problem can be written as

$$\text{minimize } Eh^*(v) \quad \text{over } v \in \mathcal{N}^\perp, \quad (\text{DSP})$$

where h^* is the normal integrand conjugate to h and

$$\mathcal{N}^\perp := \{v \in L^1 \mid \langle z, v \rangle = 0 \ \forall z \in \mathcal{N}^\infty\}.$$

Recall that the *recession function* of a closed proper convex function g is given by

$$g^\infty(x) = \sup_{\alpha > 0} \frac{g(\bar{x} + \alpha x) - g(\bar{x})}{\alpha},$$

where $\bar{x} \in \text{dom } g$; see [14, Corollary 3C]. We define the function h^∞ scenario-wise by $h^\infty(\cdot, \omega) = h(\cdot, \omega)^\infty$. By [25, Exercise 14.54], h^∞ is a normal integrand.

Theorem 1 *We have $\phi^* = Eh^* + \delta_{\mathcal{N}^\perp}$. In particular, the following are equivalent for a $v \in L^1$*

- (a) v is a shadow price of information,
- (b) v solves the dual problem and there is no duality gap,
- (c) $v \in \mathcal{N}^\perp$ and the optimum value of (SP) equals $\inf_{x \in L^\infty} E[h(x) - x \cdot v]$,
- (d) $v \in \mathcal{N}^\perp$ and the optimum value of (SP) equals $\inf_{x \in L^0} E[h(x) - x \cdot v]$.

In this case, an $x \in \mathcal{N}$ is optimal if and only if $Eh(x) < \infty$ and it minimizes the function $x \mapsto h(x, \omega) - v(\omega) \cdot x$ almost surely. There is no duality gap, in particular, if

$$\{x \in \mathcal{N} \mid h^\infty(x) \leq 0 \text{ } P\text{-a.s.}\}$$

is a linear space and there exists $v \in \mathcal{N}^\perp$ such that $Eh^(\lambda v) < \infty$ for two different values of $\lambda \in \mathbb{R}$. Moreover, in this case, the primal optimum is attained and*

$$\phi^\infty(z) = \inf_{x \in \mathcal{N}} Eh^\infty(x + z).$$

The linearity condition in terms of h^∞ holds in particular if $\text{dom } h(\cdot, \omega)$ is bounded for P -almost every ω . Indeed, we then have $h^\infty(x, \omega) = \infty$ unless $x = 0$, so $\{x \in \mathcal{N} \mid h^\infty(x) \leq 0 \text{ } P\text{-a.s.}\} = \{0\}$. The condition involving λ , holds in particular if the normal integrand h is bounded from below by some integrable function m since then, $h^*(0, \omega) \leq -m(\omega)$, so the condition is satisfied with $v = 0$. These conditions are clearly implied by Assumption C of [21] where the sets $\text{dom } h(\cdot, \omega)$ are uniformly bounded and there exists an integrable function μ such that $|h(x, \omega)| \leq \mu(\omega)$ for every $x \in \text{dom } h(\cdot, \omega)$.

The notion of a shadow price of information first appeared in a general single period model in Rockafellar [18, Example 6 in Section 10] and Rockafellar and Wets [20, Section 4]. Extension to finite discrete time was given in [21]. Continuous-time extensions have been studied in Wets [29], Back and Pliska [2], Davis [5] and Davis and Burstein [6] under various structural assumptions.

The shadow price of information has been found useful e.g. in duality theory and in deriving optimality conditions in general parametric stochastic optimization problems; see e.g. [22, 2, 3]. It is the basis for the “progressive hedging algorithm” introduced in [26]. The shadow price of information is useful also in subdifferential calculus involving conditional expectations; see [23] and Section 4.2 below. As a further application, we give a dual formulation of the general dynamic programming recursion from [21] and [7]; see Section 5.

The main result of this paper, Theorem 2 below, gives new generalized sufficient conditions for the existence of a shadow price of information for problem (SP). Its proof is obtained by extending the original argument of [21] and by relaxing some of the technical assumptions made there. We will use the notation $x^t := (x_0, \dots, x_t)$ and we denote the conditional expectation with respect to \mathcal{F}_t by E_t .

Assumption 1 *For every $z \in \text{dom } Eh \cap L^\infty$ and every $t = 0, \dots, T$, there exists $\hat{z} \in \text{dom } Eh \cap L^\infty$ with $E_t z^t = \hat{z}^t$.*

Assumption 1 relaxes the assumptions of [21]. Indeed, Assumptions C and D of [21] require the existence of a $\mu \in L^1$ such that $|h(x, \omega)| \leq \mu(\omega)$ for all $x \in \text{dom } h(\cdot, \omega)$ and that the sets $\text{dom } h(\cdot, \omega)$ are closed, uniformly bounded and that the projection mappings $\omega \mapsto \{x^t \mid x \in \text{dom } h(\cdot, \omega)\}$ are \mathcal{F}_t -measurable for all t . The following example gives more general conditions in the spirit of the “bounded recourse condition” given in [24].

Example 1 (Bounded recourse condition) Let \mathbb{B}_r be the Euclidean unit ball of radius r . Assumption 1 holds if for every $r > 0$ large enough, there exists $\beta \in L^1$ such that the projection mappings $D_r^t(\omega) := \{x^t \mid x \in \text{dom } h(\cdot, \omega) \cap \mathbb{B}_r\}$ are closed-valued and \mathcal{F}_t -measurable, and that

$$h(x, \omega) \leq \beta(\omega) \quad \forall x \in \text{dom } h(\cdot, \omega) \cap \mathbb{B}_r. \quad (1)$$

Indeed, if $z \in \text{dom } Eh \cap L^\infty$, then there exists $r > 0$ and $\beta \in L^1$ satisfying the above conditions together with $z \in \text{dom } h \cap \mathbb{B}_r$ almost surely. By Jensen’s inequality (see e.g. [9, Corollary 2.1] or Remark 2 in Section 4.2 below), the \mathcal{F}_t -measurability and closed-valuedness of $D_r^t(\omega)$ implies that $E_t z^t \in D_r^t$ almost surely as well. By the measurable selection theorem (see [25, Corollary 14.6]), there exists a $\hat{z} \in L^0$ such that $\hat{z}^t = E_t z^t$ and $\hat{z} \in \text{dom } h \cap \mathbb{B}_r$ almost surely. The upper bound β now gives $Eh(\hat{z}) < \infty$.

We will also use the following assumption where $\|\cdot\|$ denotes the essential supremum norm of L^∞ and \mathcal{L} is the linear subspace of L^∞ generated by the set $\text{dom } Eh \cap L^\infty - \bar{x}$, where $\bar{x} \in \text{dom } Eh \cap L^\infty$. Note that $\mathcal{L} = \text{aff dom } Eh \cap L^\infty - \bar{x}$,

where “aff” stands for the *affine hull* of a set. Throughout, the *strong topology* refers to the norm topology of L^∞ .

Assumption 2 *The function Eh is strongly continuous at a point of \mathcal{N}^∞ relative to $\text{aff dom } Eh \cap L^\infty$. There exists $\rho \in \mathbb{R}$ such that, for every $z \in \mathcal{N}^\infty + \mathcal{L}$, there exist $x \in \mathcal{N}^\infty$ and $w \in \mathcal{L}$ with $z = x + w$ and $\|w\| \leq \rho\|z\|$.*

Assumption 2 is implied by (1) and the *strict feasibility* condition assumed in [21, Theorem 2]. Indeed, these conditions imply that Eh is strongly continuous at a point of \mathcal{N}^∞ and, in particular, that $\text{dom } Eh$ contains an open ball of L^∞ so that $\mathcal{L} = L^\infty$. One can then simply take $x = 0$, so Assumption 2 holds with $\rho = 1$.

Recall that a convex function is continuous at a point if and only if it is bounded from above on a neighbourhood of the point; see e.g. [1, Theorem 5.43]. A sufficient condition for the relative continuity of Eh is given in Theorem 4 below. The second condition of Assumption 2 holds if \mathcal{L} and $\mathcal{N}^\infty + \mathcal{L}$ are both strongly closed, since then $\mathcal{N}^\infty + \mathcal{L}$ is a Banach space, so the condition holds by [27, Theorem 5.20]. In particular, the second condition holds automatically for finite Ω since affine sets in a Euclidean space are closed. A sufficient condition for the closedness of \mathcal{L} in the general case is given in Theorem 4. In the single-period case where $T = 0$, the second property of Assumption 2 is implied by Assumption 1. Indeed, if $z = x + w$ for $x \in \mathcal{N}^\infty$ and $w \in \mathcal{L}$, then $z = (x + E_0w) + (w - E_0w)$, where $\|w - E_0w\| = \|z - E_0z\| \leq 2\|z\|$ and, by Assumption 1, $E_0w \in \mathcal{L}$.

Combining Lemmas 1 and 2 and Theorem 3 below gives the following extension of [21, Theorem 2].

Theorem 2 *Under Assumptions 1 and 2, shadow price of information exists.*

The rest of this paper is organized as follows. Section 2 proves Theorem 3. In order to clarify the structure and its logic, we have split the proof in three statements of independent interest. Section 3 gives a sufficient conditions for relative continuity of integral functionals and for Assumption 2. Section 4 applies the main results to calculate conjugates and subdifferentials of integral functionals. Section 5 develops a dynamic programming recursion for the dual problem, by applying the results of Section 4.

2 Proof of Theorem 2

Given a function g on L^∞ , we denote its *closure* by $\text{cl } g := g^{**}$ and its *subdifferential* (i.e. the set of subgradients) at $x \in L^\infty$ by $\partial g(x)$, both defined with respect to the pairing of L^∞ with L^1 . Accordingly, all topological properties on L^∞ refer to the weak topology generated by L^1 , unless otherwise specified.

The proof of Theorem 2 is largely based on analyzing the auxiliary value function $\tilde{\phi} : L^\infty \rightarrow \overline{\mathbb{R}}$ defined by

$$\tilde{\phi}(z) = \inf_{x \in \mathcal{N}^\infty} Eh(x + z).$$

Here decision strategies are restricted to be essentially bounded like in [21]. Clearly, $\tilde{\phi}$ is convex and $\tilde{\phi} \geq \phi$.

Lemma 1 *We have $\text{cl } \tilde{\phi} = \text{cl } \phi$. If $\partial\tilde{\phi}(0)$ is nonempty, then $\partial\tilde{\phi}(0) = \partial\phi(0)$.*

Proof As shown in the proof of Theorem 1, $\tilde{\phi}^* = \phi^*$ which is equivalent to $\text{cl } \tilde{\phi} = \text{cl } \phi$. If $\partial\tilde{\phi}(0) \neq \emptyset$, then $\tilde{\phi}(0) = \text{cl } \tilde{\phi}(0)$ so $\tilde{\phi}(0) = \phi(0)$, since we always have $\tilde{\phi} \geq \phi \geq \text{cl } \phi$. Thus, $v \in \partial\tilde{\phi}(0)$ if and only if $v \in \partial\phi(0)$. \square

To prove Theorem 2, it suffices, by Lemma 1, to show that $\tilde{\phi}$ is subdifferentiable at the origin. We will do this as in [21], by first establishing the existence of a subgradient \bar{v} of $\tilde{\phi}$ with respect to the pairing of L^∞ with its Banach dual $(L^\infty)^*$, and then using \bar{v} to construct another subgradient of $\tilde{\phi}$ which belongs to L^1 . The first step is established by the following (purely functional analytic) lemma, the proof of which is given in the appendix.

Lemma 2 *Under Assumption 2, $\tilde{\phi}$ is strongly subdifferentiable at the origin.*

By [30], any $v \in (L^\infty)^*$ can be expressed as $v = v^a + v^s$ where $v^a \in L^1$ and $v^s \in (L^\infty)^*$ is such that there is a decreasing sequence of sets $A^\nu \in \mathcal{F}$ such that $P(A^\nu) \searrow 0$ and

$$\langle z, v^s \rangle = 0$$

for any $z \in L^\infty$ that vanishes on A^ν .

Theorem 3 *Under Assumption 1, $\tilde{\phi}$ subdifferentiable (resp. closed) at the origin if and only if it is strongly subdifferentiable (resp. strongly closed) at the origin.*

Proof Clearly, subdifferentiability (resp. closedness) implies strong subdifferentiability (resp. closedness). Strong closedness of ϕ at the origin means that $\tilde{\phi}(0) = \tilde{\phi}^{**}(0)$, i.e. that for every $\epsilon > 0$ there is a $v \in (L^\infty)^*$ such that $\tilde{\phi}(0) \leq -\tilde{\phi}^*(v) + \epsilon$, or equivalently,

$$\begin{aligned} \tilde{\phi}(w) &\geq \tilde{\phi}(0) + \langle w, v \rangle - \epsilon && \forall w \in L^\infty \\ \iff Eh(x+w) &\geq \tilde{\phi}(0) + \langle w, v \rangle - \epsilon && \forall w \in L^\infty, x \in \mathcal{N}^\infty \\ \iff Eh(z) &\geq \tilde{\phi}(0) + \langle z-x, v \rangle - \epsilon && \forall z \in L^\infty, x \in \mathcal{N}^\infty, \end{aligned}$$

which means that $v \perp \mathcal{N}^\infty$ and

$$Eh(z) \geq \tilde{\phi}(0) + \langle z, v \rangle - \epsilon \quad \forall z \in L^\infty. \quad (2)$$

Similarly, $\tilde{\phi}$ is strongly subdifferentiable at the origin if and only if $v \perp \mathcal{N}^\infty$ and (2) holds with $\epsilon = 0$.

We will prove the existence of a $v \perp \mathcal{N}^\infty$ which has $v^s = 0$ and satisfies (2) with ϵ multiplied with 2^{T+1} . Similarly to the above, this means that ϕ is closed (if (2) holds with all $\epsilon > 0$) or subdifferentiable (if $\epsilon = 0$) at the origin with respect to the weak topology. The existence will be proved recursively by showing that if $v \perp \mathcal{N}^\infty$ satisfies (2) and $v_{t'}^s = 0$ for $t' > t$ (this does hold for

$t = T$ as noted above), then there exists a $\tilde{v} \perp \mathcal{N}^\infty$ which satisfies (2) with ϵ multiplied by 2 and $\tilde{v}_{t'}^s = 0$ for $t' \geq t$.

Thus, assume that $v_{t'}^s = 0$ for $t' > t$ and let $\bar{\epsilon} > 0$ and $\bar{x} \in \mathcal{N}^\infty$ be such that $\tilde{\phi}(0) \geq Eh(\bar{x}) - \bar{\epsilon}$. Combining this inequality with (2) and noting that $\langle \bar{x}, v \rangle = 0$, we get

$$Eh(z) \geq Eh(\bar{x}) + \langle z - \bar{x}, v \rangle - \epsilon - \bar{\epsilon} \quad \forall z \in L^\infty.$$

Let $z \in \text{dom } Eh \cap L^\infty$ and let \hat{z} be as in Assumption 1. By Theorem 10 in the appendix,

$$Eh(z) \geq Eh(\bar{x}) + \langle z - \bar{x}, v^a \rangle - \epsilon - \bar{\epsilon}, \quad (3)$$

and (since $\hat{z} \in \text{dom } Eh \cap L^\infty$)

$$0 \geq \langle \hat{z} - \bar{x}, v^s \rangle - \epsilon - \bar{\epsilon}. \quad (4)$$

Since $\hat{z}^t = E_t z^t$ and $v_{t'}^s = 0$ for $t' > t$, by assumption, (4) means that

$$0 \geq \sum_{t'=0}^t \langle E_t z_{t'} - \bar{x}_{t'}, v_{t'}^s \rangle - \epsilon - \bar{\epsilon}.$$

Each term in the sum can be written as $\langle z_{t'} - \bar{x}_{t'}, A_{t'}^* v_{t'}^s \rangle$, where $A_{t'}^*$ denotes the adjoint of the linear mapping $A_{t'} : L^\infty(\Omega, \mathcal{F}, P; \mathbb{R}^{n_{t'}}) \rightarrow L^\infty(\Omega, \mathcal{F}, P; \mathbb{R}^{n_{t'}})$ defined by $A_{t'} x_{t'} := E_t x_{t'}$. Moreover, since $v \perp \mathcal{N}^\infty$, we have $A_t^* v_t = 0$ so, in the last term, $A_t^* v_t^s = -A_t^* v_t^a$. By the tower property of the conditional expectation, $A_t^* v_t^a = E_t v_t^a$. Thus, combining (4) and (3) gives

$$Eh(z) \geq Eh(\bar{x}) + \langle z - \bar{x}, \tilde{v} \rangle - 2\epsilon - 2\bar{\epsilon},$$

where

$$\tilde{v}_{t'} = \begin{cases} v_{t'}^a + A_{t'}^* v_{t'}^s & \text{for } t' < t, \\ v_{t'}^a - E_t v_{t'}^a & \text{for } t' = t, \\ v_{t'} & \text{for } t' > t. \end{cases}$$

We still have $\tilde{v} \in \mathcal{N}^\perp$ but now $\tilde{v}_{t'}^s = 0$ for every $t' \geq t$ as desired. Since $\bar{\epsilon} > 0$ was arbitrary and $\langle \bar{x}, \tilde{v} \rangle = 0$, we see that \tilde{v} satisfies (2) with ϵ multiplied by 2. This completes the proof since $z \in \text{dom } Eh \cap L^\infty$ was arbitrary. \square

In summary, Assumption 2 implies, by Lemma 2, the existence of a strong subgradient of $\tilde{\phi}$ at the origin. By Theorem 3, Assumption 1 then implies $\partial\tilde{\phi}(0) \neq \emptyset$ so by Lemma 1, $\partial\phi(0) \neq \emptyset$, which completes the proof of Theorem 2.

It is clear that, in the above proof of Theorem 2, Assumption 2 could be replaced by the more abstract requirement that $\tilde{\phi}$ be strongly subdifferentiable at the origin. Assumption 2 is merely a sufficient condition for this.

3 Relative continuity of integral functionals

If Eh is closed proper and convex with $\text{aff}(\text{dom } Eh \cap L^\infty)$ closed, then Eh is continuous on $\text{rint}_s(\text{dom } Eh \cap L^\infty)$, the *relative strong interior* of $\text{dom } Eh \cap L^\infty$ (recall that the relative interior of a set is defined as its interior with respect to its affine hull). Indeed, the closedness of $\text{aff } \text{dom } Eh \cap L^\infty$ implies that it is a translation of a Banach space, and then Eh is strongly continuous relative to $\text{rint}_s(\text{dom } Eh \cap L^\infty)$; see e.g. [18, Corollary 8B].

The following result gives sufficient conditions for $\text{aff } \text{dom } Eh$ to be closed (not just strongly but weakly) and $\text{rint}_s(\text{dom } Eh \cap L^\infty)$ to be nonempty. Its proof is obtained by modifying the proof of [17, Theorem 2] which required, in particular, that $\text{aff } \text{dom } h = \mathbb{R}^n$ almost surely. Recall that the set-valued mappings $\omega \mapsto \text{dom } h(\cdot, \omega)$ and $\omega \mapsto \text{aff } \text{dom } h(\cdot, \omega)$ are measurable; see [25, Proposition 14.8 and Exercise 14.12]. Given a measurable mapping $D : \Omega \rightrightarrows \mathbb{R}^n$, we will use the notation $L^p(D) := \{x \in L^p \mid x \in D \text{ } P\text{-a.s.}\}$.

Theorem 4 *Assume that the set*

$$\mathcal{D} = \{x \in L^\infty(\text{dom } h) \mid \exists r > 0 : \mathbb{B}_r(x) \cap \text{aff } \text{dom } h \subseteq \text{dom } h \text{ } P\text{-a.e.}\}$$

is nonempty and contained in $\text{dom } Eh$. Then $Eh : L^\infty \rightarrow \overline{\mathbb{R}}$ is closed proper and convex, $\text{aff}(\text{dom } Eh \cap L^\infty)$ is closed and $\text{rint}_s(\text{dom } Eh \cap L^\infty) = \mathcal{D}$. In particular, Eh is strongly continuous throughout \mathcal{D} relative to $\text{aff}(\text{dom } Eh \cap L^\infty)$.

Proof Translating, if necessary, we may assume $0 \in \mathcal{D}$ so that $L^\infty(\text{aff } \text{dom } h) \subseteq \cup_{\lambda > 0} \lambda \mathcal{D} \subseteq \text{aff } \mathcal{D}$. By assumption,

$$\mathcal{D} \subseteq \text{dom } Eh \cap L^\infty \subseteq L^\infty(\text{dom } h) \subseteq L^\infty(\text{aff } \text{dom } h),$$

so $\text{aff } \mathcal{D} = \text{aff}(\text{dom } Eh \cap L^\infty) = \text{aff } L^\infty(\text{dom } h) = L^\infty(\text{aff } \text{dom } h)$ which is a closed set. Thus, the above also implies

$$\text{rint}_s \mathcal{D} \subseteq \text{rint}_s(\text{dom } Eh \cap L^\infty) \subseteq \text{rint}_s L^\infty(\text{dom } h).$$

Clearly $\text{rint}_s L^\infty(\text{dom } h) \subseteq \mathcal{D}$ while $\text{rint}_s \mathcal{D} = \mathcal{D}$. It remains to prove that Eh is closed and proper.

Let $\bar{r} > 0$ be such that $\mathbb{B}_{\bar{r}}(0) \cap \text{aff } \text{dom } h \subseteq \text{rint } \text{dom } h$ almost surely (here, the relative interior is taken scenario-wise with respect to the usual Euclidean topology) and let $\pi(\omega)$ be the orthogonal projection of \mathbb{R}^n to $\text{aff } \text{dom } h(\cdot, \omega)$. Let $x^i \in \mathbb{B}_{\bar{r}}(0)$, $i = 0, \dots, n$ and $r > 0$ be such that $\mathbb{B}_r(0)$ is contained in the convex hull of $\{x^i \mid i = 0, \dots, n\}$. By [25, Exercise 14.17], πx is measurable for every measurable x , so each πx^i belongs to \mathcal{D} and thus,

$$\alpha := \max_{i=0, \dots, n} h(\pi x^i)$$

is integrable. Since $0 \in \text{rint } \text{dom } h$ almost surely, the closed convex-valued mapping $\omega \mapsto \partial h(0, \omega)$ is nonempty-valued ([16, Theorem 23.4]) and measurable ([25, Theorem 14.56]). Thus, by [25, Corollary 14.6], it admits a measurable

selection $w \in L^0(\partial h(0))$. We also have $y := \pi w \in L^0(\partial h(0))$, so $h^*(y) \leq -h(0)$ and thus, $Eh^*(y) < \infty$. Moreover, by Fenchel's inequality and convexity of h ,

$$\begin{aligned} r|y(\omega)| &= \sup_{x \in \mathbb{B}_r(0)} \{y(\omega) \cdot x\} \\ &= \sup_{x \in \mathbb{B}_r(0)} \{w(\omega) \cdot \pi(\omega)x\} \\ &\leq \sup_{x \in \mathbb{B}_r(0)} h(\pi(\omega)x, \omega) - h(0, \omega) \\ &\leq \sup_{i=0, \dots, n} h(\pi(\omega)x^i, \omega) - h(0, \omega) \\ &\leq \alpha(\omega) - h(0, \omega). \end{aligned}$$

Thus, $y \in L^1$ so Eh is closed and proper, by [15, Theorem 2]. \square

Remark 1 Under the assumptions of Theorem 4, Eh is subdifferentiable throughout \mathcal{D} with respect to the pairing of L^∞ with L^1 . Indeed, the proof of Theorem 4 constructs a $y \in L^1$ with $y \in \partial h(x)$ almost surely, which implies $y \in \partial Eh(x)$.

Example 2 The extension of the strict feasibility condition of [17, Theorem 2] in Theorem 4 is needed, for example, in problems of the form

$$\begin{aligned} &\text{minimize} && Eh_0(x) \quad \text{over } x \in \mathcal{N}^\infty \\ &\text{subject to} && Ax = b \quad P\text{-a.s.}, \end{aligned}$$

where h_0 is a convex normal integrand such that $h_0(x, \cdot) \in L^1$ for every $x \in \mathbb{R}^n$, A is a measurable matrix and b is a measurable vector of appropriate dimensions such that the problem is feasible. This fits the general format of (SP) with

$$h(x, \omega) = \begin{cases} h_0(x, \omega) & \text{if } A(\omega)x = b(\omega), \\ +\infty & \text{otherwise.} \end{cases}$$

The integrability of h_0 implies that $\text{aff dom } h = \text{dom } h$ almost surely and $\mathcal{D} = \{x \in L^\infty \mid Ax = b \text{ } P\text{-a.s.}\} = \text{dom } Eh$. The conditions of Theorem 4 are then satisfied but the strict feasibility assumption of [17, Theorem 2] fails unless $\mathcal{D} = L^\infty$.

Corollary 1 *Let \mathcal{D} be as in Theorem 4 and assume that $\mathcal{N}^\infty \cap \mathcal{D} \neq \emptyset$. Then Assumption 2 holds if and only if $\mathcal{N}^\infty + \mathcal{L}$ is strongly closed.*

Proof By Theorem 4, the assumptions imply the first property in Assumption 2 and that \mathcal{L} is strongly closed. By [27, Theorem 5.20], the closedness of $\mathcal{N}^\infty + \mathcal{L}$ implies the existence of a $\rho > 0$ in Assumption 2. The converse has been proved in [31]. We reproduce here the simple argument. Let (z^ν) be a Cauchy sequence in $\mathcal{N}^\infty + \mathcal{L}$. Passing to a subsequence, we may assume that $\|z^\nu - z^{\nu-1}\| \leq C/2^\nu$ for some $C > 0$. By Assumption 2, there exist $\rho > 0$, $x^\nu \in \mathcal{N}^\infty$ and $w^\nu \in \mathcal{L}$ with $z^\nu - z^{\nu-1} = x^\nu + w^\nu$ and $\|w^\nu\| \leq \rho\|z^\nu - z^{\nu-1}\|$. It follows that $\bar{w}^\mu := \sum_{\nu=1}^\mu w^\nu$ and $\bar{x}^\mu := \sum_{\nu=1}^\mu x^\nu$ converge strongly to some \bar{w} and \bar{x} , respectively, whose sum equals the limit of (z^ν) . The closedness of \mathcal{L} and \mathcal{N}^∞ then implies the closedness of $\mathcal{N}^\infty + \mathcal{L}$. \square

4 Calculating conjugates and subgradients

This section applies the results of the previous sections to calculate subdifferentials and conjugates of certain integral functionals and conditional expectations of normal integrands.

4.1 Integral functionals on \mathcal{N}^∞

Let f be a convex normal integrand such that Ef is proper on \mathcal{N}^∞ . The space \mathcal{N}^∞ is in separating duality with $\mathcal{N}^1 := \mathcal{N} \cap L^1$ under the bilinear form

$$\langle x, v \rangle := E(x \cdot v).$$

We will use the results of the previous section to calculate the conjugate and the subdifferential of Ef with respect to this pairing.

If $x \in \mathcal{N}^\infty$ and $v \in L^1(\partial f(x))$, then $Ef(x') \geq Ef(x) + \langle x' - x, v \rangle$ for all $x' \in \mathcal{N}^\infty$, so

$${}^aL^1(\partial f(x)) \subseteq \partial Ef(x), \quad (5)$$

where ${}^aL^1(\partial f(x)) := \{{}^av \mid v \in L^1(\partial f(x)) \text{ } P\text{-a.s.}\}$. Here and in what follows av denotes the *adapted projection* of an $x \in L^1$, that is, $({}^av)_t = E_tv_t$.

The following theorem gives sufficient conditions for (5) to hold as an equality.

Theorem 5 *Assume that $x^* \in \mathcal{N}^1$ is such that the function $\tilde{\phi}_{x^*} : L^\infty \rightarrow \overline{\mathbb{R}}$,*

$$\tilde{\phi}_{x^*}(z) := \inf_{x \in \mathcal{N}^\infty} E[f(x+z) - (x+z) \cdot x^*]$$

is closed at the origin. Then

$$(Ef)^*(x^*) = \inf_{v \in \mathcal{N}^\perp} Ef^*(x^* + v).$$

If $\tilde{\phi}_{x^}$ is subdifferentiable at the origin, then the infimum over \mathcal{N}^\perp is attained. If this holds for every $x^* \in \partial Ef(x)$ with $x \in \mathcal{N}^\infty$, then*

$$\partial Ef(x) = {}^aL^1(\partial f(x)).$$

Proof When $\tilde{\phi}_{x^*}$ is closed at the origin, $(Ef)^*(x^*) = -\tilde{\phi}_{x^*}(0) = -\text{cl } \tilde{\phi}_{x^*}(0) = \inf_{y \in L^1} \tilde{\phi}_{x^*}^*(y)$. For any $v \in \mathcal{N}^\perp$,

$$\begin{aligned} (Ef)^*(x^*) &= \sup_{x \in \mathcal{N}^\infty} \{ \langle x, x^* \rangle - Ef(x) \} \\ &= \sup_{x \in \mathcal{N}^\infty} E\{x \cdot (x^* + v) - f(x)\} \\ &\leq E \sup_{x \in \mathbb{R}^n} \{x \cdot (x^* + v) - f(x)\} \\ &= Ef^*(x^* + v) \end{aligned}$$

so the conjugate formula holds trivially unless $\tilde{\phi}_{x^*}$ is proper. When $\tilde{\phi}_{x^*}$ is proper then Ef is proper on L^∞ as well and $\tilde{\phi}_{x^*}^*(y) = Ef^*(x^* + y) + \delta_{\mathcal{N}^\perp}(y)$ exactly like in the proof of Theorem 1. The second claim follows from the identity $\partial\tilde{\phi}_{x^*}^*(0) = \operatorname{argmin}_{y \in L^1} \tilde{\phi}_{x^*}^*(y)$; see e.g. [18, Corollary 12B].

Assume now that $\tilde{\phi}_{x^*}$ is subdifferentiable at the origin for $x^* \in \partial Ef(x)$. Then the infimum in the expression for $(Ef)^*(x^*)$ is attained and $Ef(x) + (Ef)^*(x^*) = \langle x, x^* \rangle$, so there is a $v \in \mathcal{N}^\perp$ such that $E[f(x) + f^*(x^* + v)] = E[x \cdot (x^* + v)]$, and thus $x^* + v \in \partial f(x)$ almost surely. Clearly, $x^* = {}^a(x^* + v)$. Thus, $\partial Ef(x) \supseteq {}^aL^1(\partial f(x))$ while the reverse inclusion is given in (5). \square

The results of Section 2 provide global conditions that imply the local conditions in Theorem 5.

Corollary 2 *If f satisfies Assumptions 1 and 2, then*

$$(Ef)^*(x^*) = \inf_{v \in \mathcal{N}^\perp} Ef^*(x^* + v) \quad \forall x^* \in \mathcal{N}^1$$

where the infimum is attained, and

$$\partial Ef(x) = {}^aL^1(\partial f(x))$$

for every $x \in \mathcal{N}^\infty$.

Proof Let $x^* \in \mathcal{N}^1$. Since $\operatorname{dom} Ef \cap \mathcal{N}^\infty \neq \emptyset$, we have $\tilde{\phi}_{x^*}(0) < \infty$. If $\tilde{\phi}_{x^*}(0) = -\infty$, then $\tilde{\phi}_{x^*}$ is trivially closed at the origin. Assume now that $\tilde{\phi}_{x^*}(0) > -\infty$. The assumed properties of f imply that Assumptions 1 and 2 are satisfied by $h(x, \omega) := f(x, \omega) - x \cdot x^*(\omega)$ and that Eh is continuous at a point of \mathcal{N}^∞ relative to $\operatorname{aff} \operatorname{dom} h \cap L^\infty$. By Lemma 2 and Theorem 3, $\tilde{\phi}_{x^*}$ is subdifferentiable at the origin. If $x^* \in \partial(Ef)(x)$, Fenchel's inequality $\tilde{\phi}_{x^*}(0) \geq E[f(x) - x \cdot x^*] \geq -(Ef)^*(x^*)$ implies $\tilde{\phi}_{x^*}(0) > -\infty$. The assumptions of Theorem 5 are thus satisfied. \square

Without the assumptions of Corollary 2, inclusion (5) may be strict. The following simple example is from page 176 of [23]: Let $T = 0$, $n = 1$, $\mathcal{F}_0 = \{\Omega, \emptyset\}$ (so that \mathcal{N}^∞ may be identified with \mathbb{R}) and $f(\cdot, \omega) = \delta_{[-\infty, \xi(\omega)]}$, where ξ is a random variable uniformly distributed on $[0, 1]$. One then has $Ef = \delta_{\mathbb{R}_-}$ so $\partial Ef(0) = \mathbb{R}_+$ but $\partial f(0) = \{0\}$ almost surely so ${}^aL^1(\partial f(x)) = \{0\}$. Here $\operatorname{dom} Ef = \{x \in L^\infty \mid x \leq 0 \text{ P-a.s.}\}$, so Assumption 2 is satisfied but Assumption 1 fails because $\xi \in \operatorname{dom} Ef$ but $E_0\xi \notin \operatorname{dom} Ef$.

4.2 Conditional expectation of a normal integrand

Results of the previous section allow for a simple proof of the interchange rule for subdifferentiation and conditional expectation of a normal integrand. Commutation of the two operations has been extensively studied ever since the introduction of the notion of a conditional expectation of a normal integrand in Bismut [4]; see Rockafellar and Wets [23], Truffert [28] and the references

there in. The results of the previous section allow us to relax some of the continuity assumption made in earlier works.

Given a sub-sigma-algebra $\mathcal{G} \subseteq \mathcal{F}$, the \mathcal{G} -conditional expectation of a normal integrand f is a \mathcal{G} -measurable normal integrand $E^{\mathcal{G}}f$ such that

$$(E^{\mathcal{G}}f)(x(\omega), \omega) = E^{\mathcal{G}}[f(x(\cdot), \cdot)](\omega) \quad P\text{-a.s.}$$

for every $x \in L^0(\Omega, \mathcal{G}, P; \mathbb{R}^n)$ such that either the positive part $f(x)^+$ or the negative part $f(x)^-$ of $f(x)$ is integrable. If $\text{dom } Ef^* \cap L^1 \neq \emptyset$, then the conditional expectation exists and is unique in the sense that if \tilde{f} is another function with the above property, then $\tilde{f}(\cdot, \omega) = (E^{\mathcal{G}}f)(\cdot, \omega)$ almost surely; see e.g. [28, Corollary 2.1.2].

The \mathcal{G} -conditional expectation of an \mathcal{F} -measurable set-valued mapping $S : \Omega \rightrightarrows \mathbb{R}^n$ is a \mathcal{G} -measurable closed-valued mapping $E^{\mathcal{G}}S$ such that

$$L^1(\mathcal{G}; E^{\mathcal{G}}S) = \text{cl}\{E^{\mathcal{G}}v \mid v \in L^1(S)\}.$$

The conditional expectation is well-defined and unique as soon as S admits at least one integrable selection; see e.g. [28, Section 2.1.1]. In this case, the support function of $E^{\mathcal{G}}S$ is the \mathcal{G} -conditional expectation of the support function of S . This is a special case of Theorem 7 below.

The general form of ‘‘Jensen’s inequality’’ in the following lemma is from [28, Corollary 2.1.2]. We give a direct proof for completeness.

Lemma 3 *If f is a convex normal integrand such that $\text{dom } Ef \cap L^\infty(\mathcal{G}) \neq \emptyset$ and $\text{dom } Ef^* \cap L^1 \neq \emptyset$, then*

$$(E^{\mathcal{G}}f)^*(E^{\mathcal{G}}v) \leq E^{\mathcal{G}}f^*(v)$$

almost surely for all $v \in L^1$ and

$$\partial(E^{\mathcal{G}}f)(x) \supseteq E^{\mathcal{G}}\partial f(x)$$

for every $x \in \text{dom } Ef \cap L^0(\mathcal{G})$.

Proof Fenchel’s inequality $f^*(v) \geq x \cdot v - f(x)$ and the assumption that $\text{dom } Ef \cap L^\infty(\mathcal{G}) \neq \emptyset$ imply that $E^{\mathcal{G}}f^*(v)$ is well-defined for all $v \in L^1$. To prove the first claim, assume, for contradiction, that there is a $v \in L^1$ and a set $A \in \mathcal{G}$ with $P(A) > 0$ on which the inequality is violated. Passing to a subset of A if necessary, we may assume that $E[\mathbb{1}_A E^{\mathcal{G}}f^*(v)] < \infty$ and thus,

$$E[\mathbb{1}_A (E^{\mathcal{G}}f)^*(E^{\mathcal{G}}v)] > E[\mathbb{1}_A E^{\mathcal{G}}f^*(v)] = E[\mathbb{1}_A f^*(v)].$$

This cannot happen since, by Fenchel’s inequality

$$E[\mathbb{1}_A f^*(v)] \geq \sup_{x \in L^\infty(\mathcal{G})} E\mathbb{1}_A[x \cdot E^{\mathcal{G}}v - (E^{\mathcal{G}}f)(x)] = E[\mathbb{1}_A (E^{\mathcal{G}}f)^*(E^{\mathcal{G}}v)],$$

where the equality follows by applying the interchange rule in $L^\infty(A, \mathcal{G}, P; \mathbb{R}^n)$.

Given $v \in L^1(\partial f(x))$, we have

$$f(x) + f^*(v) = x \cdot v$$

almost surely. Let $A^\nu = \{\|x\| \leq \nu\}$ so that $\mathbb{1}_{A^\nu}x$ is bounded. Since $\text{dom } Ef^* \cap L^1 \neq \emptyset$, Fenchel's inequality implies that $\mathbb{1}_{A^\nu}f(x)$ integrable. Taking conditional expectations,

$$\mathbb{1}_{A^\nu}E^\mathcal{G}f(x) + \mathbb{1}_{A^\nu}E^\mathcal{G}f^*(v) = \mathbb{1}_{A^\nu}x \cdot E^\mathcal{G}v,$$

so by the first part,

$$\mathbb{1}_{A^\nu}(E^\mathcal{G}f)(x) + \mathbb{1}_{A^\nu}(E^\mathcal{G}f)^*(E^\mathcal{G}v) \leq \mathbb{1}_{A^\nu}x \cdot E^\mathcal{G}v,$$

which means that $E^\mathcal{G}v \in \partial(E^\mathcal{G}f)(x)$ almost surely on A^ν . This finishes the proof since ν was arbitrary. \square

Remark 2 If in Lemma 3, f is normal \mathcal{G} -integrand, then the inequality can be written in the more familiar form $f^*(E^\mathcal{G}v) \leq E^\mathcal{G}f^*(v)$. It is clear from its proof that Lemma 3 remains valid if we replace $L^\infty(\mathcal{G})$ by $L^1(\mathcal{G})$ and L^1 by L^∞ throughout. More generally, one could replace $L^\infty(\mathcal{G})$ by $\mathcal{U} \cap L^0(\mathcal{G})$ and L^1 by \mathcal{Y} , where \mathcal{U} and \mathcal{Y} are decomposable spaces such that $x \cdot y \in L^1$ for all $x \in \mathcal{U}$ and $y \in \mathcal{Y}$.

The following gives sufficient conditions for the inequalities in Lemma 3 to hold as equalities.

Theorem 6 *Let f be a convex normal integrand such that $\text{dom } Ef \cap L^\infty(\mathcal{G}) \neq \emptyset$ and $\text{dom } Ef^* \cap L^1 \neq \emptyset$. If $x^* \in L^1(\mathcal{G})$ is such that the function $\phi : L^\infty \rightarrow \overline{\mathbb{R}}$,*

$$\tilde{\phi}(z) := \inf_{x \in L^\infty(\mathcal{G})} E[f(x+z) - (x+z) \cdot x^*]$$

is subdifferentiable at the origin, then there is a $v \in L^1$ such that $E^\mathcal{G}v = 0$ and

$$(E^\mathcal{G}f)^*(x^*) = E^\mathcal{G}f^*(x^* + v).$$

If $x \in \text{dom } Ef \cap L^0(\mathcal{G})$ and the above holds for every $x^ \in L^1(\mathcal{G}; \partial E^\mathcal{G}f(x))$, then*

$$\partial(E^\mathcal{G}f)(x) = E^\mathcal{G}\partial f(x).$$

Proof Applying Theorem 5 with $T = 0$ and $\mathcal{F}_0 = \mathcal{G}$ gives the existence of a $v \in L^1$ such that $E^\mathcal{G}v = 0$ and

$$(Ef)^*(x^*) = Ef^*(x^* + v).$$

On the other hand, $Ef = E(E^\mathcal{G}f)$ by definition, so $(Ef)^*(x^*) = E(E^\mathcal{G}f)^*(x^*)$, by [15, Theorem 2]. The first claim now follows from the fact that $E^\mathcal{G}f^*(x^* + v) \geq (E^\mathcal{G}f)^*(x^*)$ almost surely, by Lemma 3.

If $x^* \in L^1(\mathcal{G}; \partial E^\mathcal{G}f(x))$, we have

$$(E^\mathcal{G}f)(x) + (E^\mathcal{G}f)^*(x^*) = x \cdot x^* \quad P\text{-a.s.}$$

By the first part, there is a $v \in L^1$ such that $E^\mathcal{G}v = 0$ and

$$(E^\mathcal{G}f)(x) + E^\mathcal{G}f^*(x^* + v) = x \cdot x^* \quad P\text{-a.s.}$$

It follows that

$$E[f(x) + f^*(x^* + v) - x \cdot (x^* + v)] = 0,$$

which by Fenchel's inequality, implies $x^* + v \in \partial f(x)$ almost surely so $\partial(E^{\mathcal{G}}f)(x) \subseteq E^{\mathcal{G}}\partial f(x)$. Combining this with Lemma 3 completes the proof. \square

Sufficient conditions for the subdifferentiability condition are again obtained from Lemma 2 and Theorem 4.

Corollary 3 *Let f be a convex normal integrand such that $\text{dom } Ef^* \cap L^1 \neq \emptyset$, $E^{\mathcal{G}}x \in \text{dom } Ef$ for all $x \in \text{dom } Ef \cap L^\infty$ and Ef is strongly continuous at a point of $L^\infty(\mathcal{G})$ relative to $\text{aff } \text{dom } Ef \cap L^\infty$. Then for every $x^* \in L^1(\mathcal{G})$ there is a $v \in L^1$ such that $E^{\mathcal{G}}v = 0$ and*

$$(E^{\mathcal{G}}f)^*(x^*) = E^{\mathcal{G}}f^*(x^* + v).$$

Moreover,

$$\partial(E^{\mathcal{G}}f)(x) = E^{\mathcal{G}}\partial f(x)$$

for every $x \in \text{dom } Ef \cap L^0(\mathcal{G})$.

Proof Analogously to Corollary 2, the additional conditions guarantee the subdifferentiability condition in Theorem 6. Indeed, the condition $E^{\mathcal{G}}x \in \text{dom } Ef$ for all $x \in \text{dom } Ef \cap L^\infty$ implies both Assumption 1 and the second condition of Assumption 2; see the remarks after Assumption 2. \square

The above subdifferential formula was obtained in [23] while the expression for the conjugate was given in [28, Section 2.2, Corollary 3]. Both assumed the stronger condition that Ef be continuous at a point $x \in L^\infty(\mathcal{G})$ relative to all of L^∞ . A more abstract condition (not requiring the relative continuity assumed here) for the subdifferential formula is given in the corollary in Section 2.2.2 of [28].

Let g be a convex normal integrand. As soon as $\text{epi } g$ has an integrable selection (which happens exactly when $\text{dom } Eg \cap L^1 \neq \emptyset$), the \mathcal{G} -conditional expectation of the epigraphical mapping $\text{epi } g$ is also an epigraphical mapping of some normal integrand; see [28, page 140]. We denote this normal integrand by ${}^{\mathcal{G}}g$. Combining [28, Theorem 2.1.2 and Corollary 2.1.1] gives the following.

Theorem 7 (Truffert [28]) *Let g be a convex normal integrand such that $\text{dom } Eg \cap L^1 \neq \emptyset$ and $\text{dom } Eg^* \cap L^0(\mathcal{G}) \neq \emptyset$. Then ${}^{\mathcal{G}}g$ and $E^{\mathcal{G}}g^*$ are well defined and conjugates of each other.*

Combined with Theorem 7, the results of this section on conditional expectations yield expressions for ${}^{\mathcal{G}}(f^*)$ as well.

5 Dual dynamic programming

Consider again problem (SP) and define extended real-valued functions $h_t, \tilde{h}_t : \mathbb{R}^{n_1+\dots+n_t} \times \Omega \rightarrow \bar{\mathbb{R}}$ by the recursion

$$\begin{aligned} \tilde{h}_T &= h, \\ h_t &= E^{\mathcal{F}_t} \tilde{h}_t, \\ \tilde{h}_{t-1}(x^{t-1}, \omega) &= \inf_{x_t \in \mathbb{R}^{n_t}} h_t(x^{t-1}, x_t, \omega). \end{aligned} \quad (6)$$

This far reaching generalization of the classical dynamic programming recursion for control systems was introduced in [21] and [7]. The following result from [11] relaxes the compactness assumptions made in [21] and [7]. In the context of financial mathematics, this allows for various extensions of certain fundamental results in financial mathematics; see [11] for details. An extension to nonconvex stochastic optimization can be found in [12]. Recall that the optimum value of (SP) equals $\phi(0)$.

Theorem 8 (Pennanen and Perkkiö [11]) *Assume that $h \geq m$ for an $m \in L^1$ and that*

$$\{x \in \mathcal{N} \mid h^\infty(x) \leq 0 \text{ P-a.s.}\}$$

is a linear space. The functions h_t and \tilde{h}_t are then well-defined normal integrands and we have for every $x \in \mathcal{N}$ that

$$Eh_t(x^t) \geq \phi(0) \quad t = 0, \dots, T. \quad (7)$$

Optimal solutions $x \in \mathcal{N}$ exist and they are characterized by the condition

$$x_t(\omega) \in \underset{x_t}{\operatorname{argmin}} h_t(x^{t-1}(\omega), x_t, \omega) \quad \text{P-a.s.} \quad t = 0, \dots, T,$$

which is equivalent to having equalities in (7).

Consider now the dual problem (DSP). We know that the optimum dual value is at least $-\phi(0)$ and that if the values are equal, the shadow prices of information are exactly the dual solutions. Theorem 7 gives sufficient conditions for $(h_t)^* = \mathcal{F}_t(\tilde{h}_t^*)$ to hold. This suggests that the conjugates of h_t and \tilde{h}_t solve the *dual dynamic programming* equations

$$\begin{aligned} \tilde{g}_T &= h^*, \\ g_t &= \mathcal{F}_t \tilde{g}_t, \\ \tilde{g}_{t-1}(v^{t-1}, \omega) &= g_t(v^{t-1}, 0, \omega). \end{aligned} \quad (8)$$

Much like Theorem 8 characterizes optimal primal solutions in terms of the dynamic programming equations (13), the following result characterizes optimal dual solutions in terms of the dual recursion (8). It also gives conditions for (8) to be well-defined.

Theorem 9 *Assume that $\text{dom } Eh \cap \mathcal{N}^\infty \neq \emptyset$ and $\text{dom } Eh^* \cap \mathcal{N}^\perp \neq \emptyset$. Then the dual dynamic programming equations are well-defined and we have for every $v \in \mathcal{N}^\perp$ that*

$$Eg_t(E_tv^t) \geq -\phi(0) \quad t = 0, \dots, T. \quad (9)$$

In the absence of a duality gap, optimal dual solutions are characterized by having equalities in (9) while $x \in \mathcal{N}$ and $v \in \mathcal{N}^\perp$ are primal and dual optimal, respectively, if and only if $Eh(x) < \infty$, $Eh^(v) < \infty$ and*

$$Eg_t^*(x^t) + Eg_t(E_tv^t) = 0 \quad t = 0, \dots, T,$$

which is equivalent to having

$$E_tv^t \in \partial g_t^*(x^t) \quad P\text{-a.s.} \quad t = 0, \dots, T.$$

If the assumptions of Theorem 8 are satisfied, then there is no duality gap, $g_t = (h_t)^$ and $\tilde{g}_t = (\tilde{h}_t)^*$.*

Proof Let $x \in \text{dom } Eh \cap \mathcal{N}$, $v \in \text{dom } Eh^* \cap \mathcal{N}^\perp$ and $\bar{x} \in \text{dom } Eh \cap \mathcal{N}^\infty$. We start by showing inductively that $g_t = \mathcal{F}_t \tilde{g}_t$ is well-defined, $x^t, \bar{x}^t \in \text{dom } E\tilde{g}_t^*$ and $E_{t+1}v^t \in \text{dom } E\tilde{g}_t$ (where E_{T+1} is understood as the identity mapping on L^1). These conditions imply that

$$\tilde{g}_{t-1}(E_tv^{t-1}) = g_t(E_tv^t) \leq E_t \tilde{g}_t(E_{t+1}v^t), \quad (10)$$

and

$$\tilde{g}_{t-1}^*(x^{t-1}) \leq g_t^*(x^t) = E^{\mathcal{F}_t} \tilde{g}_t^*(x^t). \quad (11)$$

Indeed, the inequality in (10) follows from Lemma 3 and Theorem 7 while the equality comes from the definition of \tilde{g}_{t-1} . The inequality (11) holds since $\tilde{g}_{t-1}^*(x^{t-1}, \omega) = \text{cl} \inf_{x_t} g_t^*(x^{t-1}, x_t, \omega)$, by the definition of \tilde{g}_{t-1} . The equality in (11) holds by Theorem 7 and the definition of the conditional expectation of a normal integrand. By the assumptions, the induction hypothesis holds for $t = T$. The induction argument is then completed by (10), (11) and Theorem 7.

Combining (10) and (11) with Fenchel's inequality $g_0(0) \geq -g_0^*(x_0)$ gives

$$Eh^*(v) \geq Eg_t(E_tv^t) \geq Eg_0(0) \geq -Eg_0^*(x_0) \geq -Eg_t^*(x^t) \geq -Eh(x) \quad (12)$$

for all t . Since $x \in \text{dom } Eh \cap \mathcal{N}$ was arbitrary, we get (9). In the absence of duality gap, (12) also implies that optimal dual solutions are characterized by having equalities in (9). Likewise, we get from (12) that x and v are primal and dual optimal, respectively, if and only if

$$Eg_t^*(x^t) + Eg_t(E_tv^t) = 0 \quad t = 0, \dots, T.$$

By Fenchel's inequality, $g_t^*(x^t) + g_t(E_tv^t) \geq x^t \cdot (E_tv^t)$, so, by [13, Lemma 2], $E[x^t \cdot (E_tv^t)] = 0$ whenever the left side is integrable. Thus $Eg_t^*(x^t) + Eg_t(E_tv^t) = 0$ is equivalent to having $g_t^*(x^t) + g_t(E_tv^t) = x^t \cdot (E_tv^t)$ almost surely, which means that

$$E_tv^t \in \partial g_t^*(x^t)$$

almost surely.

Under the assumptions of Theorem 8, the absence of a duality gap follows from Theorem 1. We have already verified that $g_t^* = E^{\mathcal{F}_t} \tilde{g}_t^*$. Conjugating each line of (8) gives

$$\begin{aligned} \tilde{g}_T^* &= h, \\ g_t^* &= E^{\mathcal{F}_t} \tilde{g}_t^*, \\ \tilde{g}_{t-1}^*(x^{t-1}, \omega) &= \inf_{x_t \in \mathbb{R}^{n_t}} g_t^*(x^{t-1}, x_t, \omega), \end{aligned} \tag{13}$$

as soon as the last expression defines a normal integrand. By Theorem 8, this is indeed the case so $g_t^* = h_t$ and $(\tilde{g}_t)^* = \tilde{h}_t$ by uniqueness of h_t and \tilde{h}_t . \square

6 Appendix

This appendix contains the proofs of Theorem 1 and Lemma 2, and Theorem 10 below which was used in the proof of Theorem 3.

Proof of Theorem 1. The first two claims are simple consequences of Theorems 2 and 8 of [3] but for convenience of the reader, we reproduce the proofs in the present notation. Let $\tilde{\phi}(z) := \inf_{x \in \mathcal{N}^\infty} Eh(x+z)$. For any $v \in L^1$,

$$\begin{aligned} \tilde{\phi}^*(v) &= \sup_{x \in \mathcal{N}^\infty, z \in L^\infty} \{E(z \cdot v) - Eh(x+z)\} \\ &= \sup_{x \in \mathcal{N}^\infty, z' \in L^\infty} \{E(z' \cdot v) - E(x \cdot v) - Eh(z')\} \\ &= Eh^*(v) + \delta_{\mathcal{N}^\perp}(v), \end{aligned}$$

where the last line follows from the interchange rule [25, Theorem 14.60]. Since $\phi \leq \tilde{\phi}$, we have $\phi^* \geq \tilde{\phi}^*$. On the other hand, by Fenchel's inequality,

$$h(x+z) + h^*(v) \geq (x+z) \cdot v, \tag{14}$$

so if $x+z \in \text{dom } Eh$ and $v \in \text{dom } Eh^* \cap \mathcal{N}^\perp$, we have

$$Eh(x+z) + Eh^*(v) \geq E(z \cdot v),$$

by [13, Lemma 2]. Thus,

$$\phi(z) + Eh^*(v) \geq E(z \cdot v)$$

for all $z \in L^\infty$ and $v \in \mathcal{N}^\perp$ so $\phi^* \leq Eh^* + \delta_{\mathcal{N}^\perp} = \tilde{\phi}^*$. This proves the first claim. The equivalence of (a) and (b) follows by noting that $v \in \partial\phi(0)$ if and only if $-\phi^*(v) = \phi(0)$. By the interchange rule [25, Theorem 14.60], both (c) and (d) mean that $v \in \mathcal{N}^\perp$ and that the optimum value of (SP) equals $E[-h^*(v)]$, which is (b).

The optimality condition for x follows by observing that $x \in \mathcal{N}$ and $v \in \mathcal{N}^\perp$ are primal and dual optimal, respectively, with $Eh(x) + Eh^*(v) = 0$, if and

only if $Eh(x) < \infty$, $Eh^*(v) < \infty$ and (14) holds with $z = 0$ as an equality, or equivalently, $v \in \partial h(x)$ almost surely. The last two claims involving the recession functions are direct applications of [13, Theorem 5 and Lemma 6] with $f(x, u) = h(x + u)$, $\mathcal{U} = L^\infty$ and $\mathcal{Y} = L^1$. \square

Proof of Lemma 2. The continuity assumption means that there exist $\bar{x} \in \mathcal{N}^\infty$, $M \in \mathbb{R}$ and $\varepsilon > 0$ such that $Eh(\bar{x} + w) \leq M$ when $\bar{x} + w \in \text{aff dom } Eh$ and $\|w\| \leq \varepsilon$. Since $\text{dom } \tilde{\phi} = \mathcal{N}^\infty + \text{dom } Eh \cap L^\infty$, we have $\text{aff dom } \tilde{\phi} = \mathcal{N}^\infty + \text{aff}(\text{dom } Eh \cap L^\infty)$. Thus, if $z \in \text{aff dom } \tilde{\phi}$ is such that $\|z\| \leq \varepsilon/\rho$, Assumption 2 gives the existence of $x \in \mathcal{N}^\infty$ and $w \in L^\infty$ such that $\bar{x} + w \in \text{aff dom } Eh$, $z = x + w$ and $\|w\| \leq \rho\|z\| \leq \varepsilon$. Thus,

$$\tilde{\phi}(z) = \inf_{x' \in \mathcal{N}^\infty} Eh(x' + x + w) \leq Eh(\bar{x} + w) \leq M.$$

Since $\tilde{\phi}(0)$ is finite by assumption, this implies that $\tilde{\phi}$ is strongly continuous and thus strongly subdifferentiable with respect to $\text{aff dom } \tilde{\phi}$; see [18, Theorem 11]. By the Hahn–Banach theorem, relative subgradients on $\text{aff dom } \tilde{\phi}$ can be extended to subgradients on L^∞ . \square

The following is a simple refinement of [17, Corollary 1B].

Theorem 10 *Let h be a convex normal integrand and $\bar{z} \in L^\infty$ such that $Eh(\bar{z})$ is finite. If $v \in (L^\infty)^*$ and $\varepsilon \geq 0$ are such that*

$$Eh(z) \geq Eh(\bar{z}) + \langle z - \bar{z}, v \rangle - \varepsilon \quad \forall z \in L^\infty, \quad (15)$$

then

$$Eh(z) \geq Eh(\bar{z}) + \langle z - \bar{z}, v^a \rangle - \varepsilon \quad \forall z \in L^\infty$$

and

$$0 \geq \langle z - \bar{z}, v^s \rangle - \varepsilon \quad \forall z \in \text{dom } Eh \cap L^\infty.$$

Proof Let $z \in \text{dom } Eh \cap L^\infty$ and define $z^\nu := \mathbb{1}_{A^\nu} \bar{z} + \mathbb{1}_{\Omega \setminus A^\nu} z$ where A^ν are the sets in the characterization of the singular component v^s . For almost every $\omega \in \Omega$, we have $z^\nu(\omega) = z(\omega)$ for ν large enough, so $h(z^\nu) \rightarrow h(z)$ almost surely and $z^\nu \rightarrow z$ both weakly and almost surely. Thus, since $h(z^\nu) \leq \max\{h(\bar{z}), h(z)\}$, Fatou's lemma and (15) give,

$$\begin{aligned} Eh(z) &\geq \limsup Eh(z^\nu) \geq Eh(\bar{z}) + \limsup \langle z^\nu - \bar{z}, v \rangle - \varepsilon \\ &= Eh(\bar{z}) + \langle z - \bar{z}, v^a \rangle - \varepsilon, \end{aligned}$$

where the equality holds since $z^\nu - \bar{z} = \mathbb{1}_{\Omega \setminus A^\nu}(z - \bar{z})$, so that

$$\langle z^\nu - \bar{z}, v \rangle = \langle z^\nu - \bar{z}, v^a \rangle \rightarrow \langle z - \bar{z}, v^a \rangle.$$

Now let $z^\nu := \mathbb{1}_{A^\nu} z + \mathbb{1}_{\Omega \setminus A^\nu} \bar{z}$. We now have that $h(z^\nu) \rightarrow h(\bar{z})$ almost surely and $z^\nu \rightarrow \bar{z}$ both weakly and almost surely. Since $h(z^\nu) \leq \max\{h(z), h(\bar{z})\}$, Fatou's lemma and (15) give,

$$\begin{aligned} Eh(\bar{z}) &\geq \limsup Eh(z^\nu) \geq Eh(\bar{z}) + \limsup \langle z^\nu - \bar{z}, v \rangle - \varepsilon \\ &= Eh(\bar{z}) + \langle z - \bar{z}, v^s \rangle - \varepsilon, \end{aligned}$$

where the equality holds since $z^\nu - \bar{z} = \mathbb{1}_{A^\nu}(z - \bar{z})$ so that

$$\langle z^\nu - \bar{z}, v \rangle = \langle z^\nu - \bar{z}, v^a \rangle + \langle z^\nu - \bar{z}, v^s \rangle \rightarrow \langle z - \bar{z}, v^s \rangle$$

which completes the proof. \square

References

1. Charalambos D. Aliprantis and Kim C. Border. *Infinite-dimensional analysis*. Springer-Verlag, Berlin, second edition, 1999. A hitchhiker's guide.
2. K. Back and S. R. Pliska. The shadow price of information in continuous time decision problems. *Stochastics*, 22(2):151–186, 1987.
3. S. Biagini, T. Pennanen, and A.-P. Perkkiö. Duality and optimality conditions in stochastic optimization and mathematical finance. *Journal of Convex Analysis*, to appear.
4. J.-M. Bismut. Intégrales convexes et probabilités. *J. Math. Anal. Appl.*, 42:639–673, 1973.
5. 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.
6. 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.
7. I. V. Evstigneev. Measurable selection and dynamic programming. *Math. Oper. Res.*, 1(3):267–272, 1976.
8. F. Hiai and H. Umegaki. Integrals, conditional expectations, and martingales of multi-valued functions. *J. Multivariate Anal.*, 7(1):149–182, 1977.
9. T. Pennanen. Convex duality in stochastic optimization and mathematical finance. *Mathematics of Operations Research*, 36(2):340–362, 2011.
10. T. Pennanen. Optimal investment and contingent claim valuation in illiquid markets. *Finance Stoch.*, 18(4):733–754, 2014.
11. T. Pennanen and A.-P. Perkkiö. Stochastic programs without duality gaps. *Mathematical Programming*, 136(1):91–110, 2012.
12. T. Pennanen, A.-P. Perkkiö, and M. Rásonyi. Existence of solutions in non-convex dynamic programming and optimal investment. *Mathematics and Financial Economics*, pages 1–16, 2016.
13. A.-P. Perkkiö. Stochastic programs without duality gaps for objectives without a lower bound. *manuscript*, 2016.
14. R. T. Rockafellar. Level sets and continuity of conjugate convex functions. *Trans. Amer. Math. Soc.*, 123:46–63, 1966.
15. R. T. Rockafellar. Integrals which are convex functionals. *Pacific J. Math.*, 24:525–539, 1968.
16. R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
17. R. T. Rockafellar. Integrals which are convex functionals. II. *Pacific J. Math.*, 39:439–469, 1971.
18. R. T. Rockafellar. *Conjugate duality and optimization*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974.
19. R. T. Rockafellar and R. J.-B. Wets. Continuous versus measurable recourse in N -stage stochastic programming. *J. Math. Anal. Appl.*, 48:836–859, 1974.
20. R. T. Rockafellar and R. J.-B. Wets. Stochastic convex programming: Kuhn-Tucker conditions. *J. Math. Econom.*, 2(3):349–370, 1975.
21. 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).

22. 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.
23. R. T. Rockafellar and R. J.-B. Wets. On the interchange of subdifferentiation and conditional expectations for convex functionals. *Stochastics*, 7(3):173–182, 1982.
24. 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.
25. 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.
26. R. T. Rockafellar and Roger J.-B. Wets. Scenarios and policy aggregation in optimization under uncertainty. *Math. Oper. Res.*, 16(1):119–147, 1991.
27. Walter Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.
28. A. Truffert. Conditional expectation of integrands and random sets. *Ann. Oper. Res.*, 30(1-4):117–156, 1991. Stochastic programming, Part I (Ann Arbor, MI, 1989).
29. 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.
30. Kôzaku Yosida and Edwin Hewitt. Finitely additive measures. *Trans. Amer. Math. Soc.*, 72:46–66, 1952.
31. Zhe-Ming Zheng and Hui-Sheng Ding. A note on closedness of the sum of two closed subspaces in a Banach space. *Commun. Math. Anal.*, 19(2):62–67, 2016.