

OPTIMAL STOPPING UNDER MODEL UNCERTAINTY: RANDOMIZED STOPPING TIMES APPROACH

BY DENIS BELOMESTNY^{*,†} AND VOLKER KRÄTSCHMER^{*}

Duisburg-Essen University and Higher School of Economics[†]*

In this work, we consider optimal stopping problems with conditional convex risk measures of the form

$$\rho_t^\Phi(X) = \sup_{Q \in \mathcal{Q}_t} \left(\mathbb{E}_Q[-X | \mathcal{F}_t] - \mathbb{E} \left[\Phi \left(\frac{dQ}{dP} \right) | \mathcal{F}_t \right] \right),$$

where $\Phi : [0, \infty[\rightarrow [0, \infty]$ is a lower semicontinuous convex mapping and \mathcal{Q}_t stands for the set of all probability measures Q which are absolutely continuous w.r.t. a given measure P and $Q = P$ on \mathcal{F}_t . Here, the *model uncertainty risk* depends on a (random) divergence $\mathbb{E}[\Phi(\frac{dQ}{dP}) | \mathcal{F}_t]$ measuring the distance between a hypothetical probability measure we are uncertain about and a reference one at time t . Let $(Y_t)_{t \in [0, T]}$ be an adapted nonnegative, right-continuous stochastic process fulfilling some proper integrability condition and let \mathcal{T} be the set of stopping times on $[0, T]$; then without assuming any kind of time-consistency for the family (ρ_t^Φ) , we derive a novel representation

$$\sup_{\tau \in \mathcal{T}} \rho_0^\Phi(-Y_\tau) = \inf_{x \in \mathbb{R}} \left\{ \sup_{\tau \in \mathcal{T}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] \right\},$$

which makes the application of the standard dynamic programming based approaches possible. In particular, we generalize the additive dual representation of Rogers [*Math. Finance* **12** (2002) 271–286] to the case of optimal stopping under uncertainty. Finally, we develop several Monte Carlo algorithms and illustrate their power for optimal stopping under Average Value at Risk.

1. Introduction. In this paper, we study the optimal stopping problems in an uncertain environment. The classical solution to the optimal stopping problems based on the dynamic programming principle assumes that there is a unique subjective prior distribution driving the reward process. However, for example, in incomplete financial markets, we have to deal with multiple equivalent martingale measures not being sure which one underlies the market. In fact, under the presence of the multiple possible distributions, a solution of the optimal stopping problem by maximization with respect to some subjective prior cannot be reliable. Instead, it is reasonable to view the multitude of possible distributions as a kind

Received March 2014; revised April 2015.

MSC2010 subject classifications. Primary 60G40; secondary 91G80.

Key words and phrases. Optimized certainty equivalents, optimal stopping, primal representation, additive dual representation, randomized stopping times, thin sets.

of *model uncertainty risk* which should be taken into account while formulating an optimal stopping problem. Here, one may draw on concepts from the theory of risk measures. As the established generic notion for static risk assessment at present time 0, convex risk measures are specific functionals ρ_0 on vector spaces of random variables viewed as financial risks (see [26] and [27]). They typically have the following type of robust representation:

$$(1.1) \quad \rho_0(X) = \sup_{Q \in \mathcal{Q}(\mathbb{P})} \{ \mathbb{E}_Q[-X] - \gamma_0(Q) \},$$

where $\mathcal{Q}(\mathbb{P})$ denotes the set of probability measures which are absolutely continuous w.r.t. a given reference probability measure \mathbb{P} , and γ_0 is some penalty function (see, e.g., [14] and [25]). In this way, model uncertainty is incorporated, as no specific probability measure is assumed. Moreover, the penalty function scales the plausibility of models.

Turning over from static to dynamic risk assessment, convex risk measures have been extended to the concept of conditional convex risk measures ρ_t at a future time t , which are specific functions on the space of financial risks with random outcomes (see [11, 18] and [15]). Under some regularity conditions, they have a robust representation of the form (see, e.g., [17, 24] or [25], Chapter 11)

$$(1.2) \quad \rho_t(X) = \sup_{Q \in \mathcal{Q}_t} \{ \mathbb{E}_Q[-X | \mathcal{F}_t] - \gamma_t(Q) \},$$

where γ_t is a (random) penalty function and \mathcal{Q}_t consists of all $Q \in \mathcal{Q}(\mathbb{P})$ with $Q = \mathbb{P}$ on \mathcal{F}_t . As in (1.1), the robust representation (1.2) mirrors the model uncertainty, but now at a future time t .

In recent years, the optimal stopping with families $(\rho_t)_{t \in [0, T]}$ of conditional convex risk measures was subject of several studies. For example, the works [33] and [30] are settled within a time-discrete framework, where in addition the latter one provides some dual representations extending the well-known ones from the classical optimal stopping. Optimal stopping in continuous time was considered in [5–7, 13]. All these contributions restrict their analysis to the families $(\rho_t)_{t \in [0, T]}$ satisfying the property of time consistency, sometimes also called recursiveness, defined to mean

$$\rho_s(-\rho_t) = \rho_s, \quad 0 \leq s < t \leq T.$$

Hence, the results of the above papers cannot be, for example, used to solve optimal stopping problems under such very popular convex risk measure as Average Value at Risk. The only paper which tackled the case of nontime-consistent families of conditional convex risk measures so far is [37], where the authors considered the so-called distorted mean payoff functionals. However, the analysis of [37] excludes the case of Average Value at Risk as well. Moreover, the class of processes to be stopped is limited to the functions of a one-dimensional geometric Brownian motion. The main probabilistic tool used in [37] is the Skorokhod embedding.

In this paper, we consider a rather general class of conditional convex risk measures having representation (1.2) with $\gamma_t(Q) = \mathbb{E}[\Phi(dQ/dP)|\mathcal{F}_t]$ for some lower semicontinuous convex mapping $\Phi : [0, \infty[\rightarrow [0, \infty]$. The related class of risk measures ρ_0 known as the class of *divergence risk measures* or *optimized certainty equivalents* was first introduced in [9, 10]. Any divergence risk measure has the representation

$$\rho_0(X) = \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x - X) - x]$$

with

$$\Phi^* : \mathbb{R} \rightarrow [0, \infty], \quad y \mapsto \sup_{x \geq 0} (xy - \Phi(x))$$

(cf. [9, 10, 16], or Appendix A). Here, we study the problem of optimally stopping the reward process $\rho_0(-Y_t)$, where $(Y_t)_{t \in [0, T]}$ is an adapted nonnegative, right-continuous stochastic process with $\sup_{t \in [0, T]} Y_t$ satisfying some suitable integrability condition. We do not assume any time-consistency for the family ρ_t and basically impose no further restrictions on (Y_t) . Our main result is the representation

$$(1.3) \quad \sup_{\tau \in \mathcal{T}} \rho_0(-Y_\tau) = \inf_{x \in \mathbb{R}} \left\{ \sup_{\tau \in \mathcal{T}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] \right\},$$

which allows one to apply the well-known methods from the theory of ordinary optimal stopping problems. In particular, we derive the so-called additive dual representation of the form:

$$(1.4) \quad \inf_{x \in \mathbb{R}} \inf_{M \in \mathcal{M}_0} \mathbb{E} \left[\sup_{t \in [0, T]} (\Phi^*(x + Y_t) - x - M_t) \right],$$

where \mathcal{M}_0 is the class of adapted martingales vanishing at time 0. This dual representation generalizes the well-known dual representation of Rogers, [35]. The representation (1.4) together with (1.3) can be used to efficiently construct lower and upper bounds for the optimal value (1.3) by Monte Carlo, see, for example, [8] and references therein.

The paper is organized as follows. In Section 2, we introduce notation and set up the optimal stopping problem. The main results are presented in Section 3 where in particular a criterion ensuring the existence of a saddle-point in (1.3) is formulated. Section 4 contains some discussion on the main results and on their relation to the previous literature. A Monte Carlo algorithm for computing lower and upper bounds for the value function is formulated in Section 5, where also an example of optimal stopping under Average Value at Risk is numerically analyzed.

The crucial idea to derive representation (1.3) is to consider the optimal stopping problem

$$\text{maximize} \quad \rho_0(-Y_{\tau^r}) \quad \text{over} \quad \tau^r \in \mathcal{T}^r,$$

where \mathcal{T}^r denotes the set of all randomized stopping times on $[0, T]$. It will be studied in Section 6, where in particular it will turn out that this optimal stopping problem has the same optimal value as the original one. Finally, the proofs are collected in Section 7.

2. The set-up. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and denote by $L^0 := L^0(\Omega, \mathcal{F}, \mathbb{P})$ the class of all finitely-valued random variables (modulo the P-a.s. equivalence). Let Ψ be a Young function, that is, a left-continuous, nondecreasing convex function $\Psi : \mathbb{R}_+ \rightarrow [0, \infty]$ such that $0 = \Psi(0) = \lim_{x \rightarrow 0} \Psi(x)$ and $\lim_{x \rightarrow \infty} \Psi(x) = \infty$. The Orlicz space associated with Ψ is defined as

$$L^\Psi := L^\Psi(\Omega, \mathcal{F}, \mathbb{P}) = \{X \in L^0 : \mathbb{E}[\Psi(c|X|)] < \infty \text{ for some } c > 0\}.$$

It is a Banach space when endowed with the Luxemburg norm

$$\|X\|_\Psi := \inf\{\lambda > 0 : \mathbb{E}[\Psi(|X|/\lambda)] \leq 1\}.$$

The Orlicz heart is

$$H^\Psi := H^\Psi(\Omega, \mathcal{F}, \mathbb{P}) = \{X \in L^0 : \mathbb{E}[\Psi(c|X|)] < \infty \text{ for all } c > 0\}.$$

For example, if $\Psi(x) = x^p/p$ for some $p \in [1, \infty[$, then $H^\Psi = L^\Psi = L^p := L^p(\Omega, \mathcal{F}, \mathbb{P})$ is the usual L^p -space. In this case $\|Y\|_\Psi = p^{-1/p} \|Y\|_p$, where $\|\cdot\|_p$ stands for L^p -norm. If Ψ takes the value $+\infty$, then $H^\Psi = \{0\}$ and $L^\Psi = L^\infty := L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ is defined to consist of all P-essentially bounded random variables. By Jensen inequality, we always have $H^\Psi \subseteq L^1$. In the case of finite Ψ , we see that L^∞ is a linear subspace of H^Ψ , which is dense w.r.t. $\|\cdot\|_\Psi$ (see Theorem 2.1.14 in [20]).

Let $0 < T < \infty$ and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ be a filtered probability space, where $(\mathcal{F}_t)_{t \in [0, T]}$ is a right-continuous filtration with \mathcal{F}_0 containing only the sets with probability 0 or 1 as well as all the null sets of \mathcal{F} . Furthermore, consider a lower semicontinuous convex mapping $\Phi : [0, \infty[\rightarrow [0, \infty]$ satisfying $\Phi(x_0) < \infty$ for some $x_0 > 0$, $\inf_{x \geq 0} \Phi(x) = 0$, and $\lim_{x \rightarrow \infty} \frac{\Phi(x)}{x} = \infty$. Its Fenchel–Legendre transform

$$\Phi^* : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}, \quad y \mapsto \sup_{x \geq 0} (xy - \Phi(x))$$

is a finite nondecreasing convex function whose restriction $\Phi^*|_{[0, \infty[}$ to $[0, \infty[$ is a finite Young function (cf. Lemma A.1 in Appendix A). We shall use H^{Φ^*} to denote the Orlicz heart w.r.t. $\Phi^*|_{[0, \infty[}$. Then we can define a conditional convex risk measure $(\rho_t^\Phi)_{t \in [0, T]}$ via

$$\rho_t^\Phi(X) = \operatorname{ess\,sup}_{Q \in \mathcal{Q}_{\Phi, t}} \left(\mathbb{E}_Q[-X | \mathcal{F}_t] - \mathbb{E} \left[\Phi \left(\frac{dQ}{dP} \right) | \mathcal{F}_t \right] \right)$$

for all $X \in H^{\Phi^*}$, where $\mathcal{Q}_{\Phi, t}$ denotes the set of all probability measures Q which are absolutely continuous w.r.t. \mathbb{P} such that $\Phi \left(\frac{dQ}{dP} \right)$ is \mathbb{P} -integrable and $Q = \mathbb{P}$ on \mathcal{F}_t . Note that $\frac{dQ}{dP} X$ is \mathbb{P} -integrable for every $Q \in \mathcal{Q}_{\Phi, 0}$ and any $X \in H^{\Phi^*}$ due to the Young's inequality. Consider now a right-continuous nonnegative stochastic process (Y_t) adapted to (\mathcal{F}_t) . Furthermore, let \mathcal{T} contain all finite stopping times

$\tau \leq T$ w.r.t. (\mathcal{F}_t) . The main object of our study is the following optimal stopping problem:

$$(2.1) \quad \sup_{\tau \in \mathcal{T}} \rho_0^\Phi(-Y_\tau).$$

If we set $\Phi(x) = 0$ for $x \leq 1$, and $\Phi(x) = \infty$ otherwise, we end up with the classical stopping problem

$$(2.2) \quad \sup_{\tau \in \mathcal{T}} \mathbb{E}[Y_\tau].$$

It is well known that the optimal value of the problem (2.2) may be viewed as a risk neutral price of the American option with the discounted payoff $(Y_t)_{t \in [0, T]}$ at time $t = 0$. However, in face of incompleteness, it seems to be not appropriate to assume the uniqueness of the risk neutral measure. Instead, the uncertainty about the stochastic process driving the payoff Y_t should be taken into account. Considering the optimal value of the problem (2.1) as an alternative pricing rule, model uncertainty risk is incorporated by taking the supremum over $\mathcal{Q}_{\Phi, t}$, where the penalty function is used to assess the plausibility of possible models. The more plausible is the model, the lower is the value of the penalty function.

EXAMPLE 2.1. Let us illustrate our setup in the case of the so-called Average Value at Risk risk measure. The Average Value at Risk risk measure at level $\alpha \in]0, 1[$ is defined as the following functional:

$$AV@R_\alpha : X \mapsto -\frac{1}{\alpha} \int_0^\alpha F_X^{\leftarrow}(\beta) d\beta,$$

where $X \in L^1$ and F_X^{\leftarrow} denotes the left-continuous quantile function of the distribution function F_X of X defined by $F_X^{\leftarrow}(\alpha) = \inf\{x \in \mathbb{R} | F_X(x) \geq \alpha\}$ for $\alpha \in]0, 1[$. Note that $AV@R_1(X) = \mathbb{E}[-X]$ for any $X \in L^1$. Moreover, it is well known that

$$AV@R_\alpha(X) = \sup_{Q \in \mathcal{Q}_{\Phi_\alpha, 0}} \mathbb{E}_Q[-X] \quad \text{for } X \in L^1,$$

where Φ_α is the Young function defined by $\Phi_\alpha(x) = 0$ for $x \leq 1/\alpha$, and $\Phi_\alpha(x) = \infty$ otherwise (cf. [25], Theorem 4.52, and [28]). Observe that the set $\mathcal{Q}_{\Phi_\alpha, 0}$ consists of all probability measures on \mathcal{F} with $\frac{dQ}{dP} \leq 1/\alpha$ P-a.s. Hence, the optimal stopping problem (2.1) reads as follows:

$$(2.3) \quad \sup_{\tau \in \mathcal{T}} AV@R_\alpha(-Y_\tau) = \sup_{\tau \in \mathcal{T}} \left\{ \frac{1}{\alpha} \int_{1-\alpha}^1 F_{Y_\tau}^{\leftarrow}(\beta) d\beta \right\}.$$

The family $(\rho_t^{\Phi_\alpha})_{t \in [0, T]}$ of conditional convex risk measure associated with Φ_α is also known as the conditional AV@R $(AV@R_\alpha(\cdot | \mathcal{F}_t))_{t \in [0, T]}$ at level α (cf. [25], Definition 11.8).

EXAMPLE 2.2. Let us consider, for any $\gamma > 0$, the continuous convex mapping $\Phi_{[\gamma]} : [0, \infty[\rightarrow \mathbb{R}$ defined by $\Phi_{[\gamma]}(x) = (x \ln(x) - x + 1)/\gamma$ for $x > 0$ and $\Phi_{[\gamma]}(0) = 1/\gamma$. The Fenchel–Legendre transform of $\Phi_{[\gamma]}$ is given by $\Phi_{[\gamma]}^*(y) = (\exp(\gamma y) - 1)/\gamma$ for $y \in \mathbb{R}$. In view of Lemma A.1 (cf. Appendix A), the corresponding risk measure $\rho_0^{\Phi_{[\gamma]}}$ has the representation

$$(2.4) \quad \rho_0^{\Phi_{[\gamma]}}(X) = \inf_{x \in \mathbb{R}} \mathbb{E} \left[\frac{\exp(\gamma x - \gamma X) - 1}{\gamma} - x \right] = \frac{\ln(\mathbb{E}[\exp(-\gamma X)])}{\gamma}$$

for $X \in H^{\Phi_{[\gamma]}^*}$. This is the well-known entropic risk measure. Optimal stopping with the entropic risk measures is easy to handle, since they can be reduced to the standard optimal stopping problems via

$$(2.5) \quad \sup_{\tau \in \mathcal{T}} \rho_0^{\Phi_{[\gamma]}}(-Y_\tau) = \frac{1}{\gamma} \cdot \ln \left(\sup_{\tau \in \mathcal{T}} \mathbb{E}[\exp(\gamma Y_\tau)] \right).$$

EXAMPLE 2.3. Set $\Phi^{[p]} = x^p/p$ for any $p \in]1, \infty[$, then the set $\mathcal{Q}_{\Phi^{[p]},0}$ contains all probability measures Q on \mathcal{F} with $\frac{dQ}{dP} \in L^p$, and

$$\rho^{\Phi^{[p]}}(X) = \sup_{Q \in \mathcal{Q}_{\Phi^{[p]},0}} \left(\mathbb{E}_Q[-X] - \frac{1}{p} \mathbb{E} \left[\left(\frac{dQ}{dP} \right)^p \right] \right) \quad \text{for } X \in L^{p/(p-1)}.$$

3. Main results. Let $\text{int}(\text{dom}(\Phi))$ denote the topological interior of the effective domain of the mapping $\Phi : [0, \infty[\rightarrow [0, \infty]$. We shall assume Φ to be a lower semicontinuous convex function satisfying

$$(3.1) \quad 1 \in \text{int}(\text{dom}(\Phi)), \quad \inf_{x \geq 0} \Phi(x) = 0 \quad \text{and} \quad \lim_{x \rightarrow \infty} \frac{\Phi(x)}{x} = \infty.$$

3.1. *Primal representation.* The following theorem is our main result.

THEOREM 3.1. Let $(\Omega, \mathcal{F}_t, P|_{\mathcal{F}_t})$ be atomless with countably generated \mathcal{F}_t for every $t > 0$. Furthermore, let (3.1) be fulfilled, and let $\sup_{t \in [0, T]} Y_t \in H^{\Phi^*}$, then

$$\begin{aligned} \sup_{\tau \in \mathcal{T}} \rho_0^\Phi(-Y_\tau) &= \sup_{\tau \in \mathcal{T}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] \\ &= \inf_{x \in \mathbb{R}} \sup_{\tau \in \mathcal{T}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] < \infty. \end{aligned}$$

REMARK 3.2. The functional

$$\rho_{\Phi^*} : H^{\Phi^*} \rightarrow \mathbb{R}, \quad X \mapsto \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + X) - x]$$

is known as the optimized certainty equivalent w.r.t. Φ^* (cf. [9, 10]). Thus, the relationship

$$(3.2) \quad \sup_{\tau \in \mathcal{T}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] = \inf_{x \in \mathbb{R}} \sup_{\tau \in \mathcal{T}} \mathbb{E}[\Phi^*(x + Y_\tau) - x]$$

may also be viewed as a representation result for optimal stopping with optimized certainty equivalents.

Let us illustrate Theorem 3.1 for the case $\Phi = \Phi_\alpha$ with some $\alpha \in]0, 1]$. The Young function Φ_α satisfies the conditions of Theorem 3.1 if and only if $\alpha < 1$. The Fenchel–Legendre transform Φ_α^* of Φ is given by $\Phi_\alpha^*(x) = x^+/\alpha$ and it fulfills the inequality $\Phi_\alpha^*(x + y) - x \geq \Phi_\alpha^*(y)$ for $x, y \geq 0$. Then, as an immediate consequence of Theorem 3.1, we obtain the following primal representation for the optimal stopping problem (2.3).

COROLLARY 3.3. *Let $(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ be atomless with countably generated \mathcal{F}_t for every $t > 0$. If $\sup_{t \in [0, T]} Y_t \in L^1$, then it holds for $\alpha \in]0, 1[$*

$$\begin{aligned} \sup_{\tau \in \mathcal{T}} AV @ R_\alpha(-Y_\tau) &= \inf_{x \in \mathbb{R}} \sup_{\tau \in \mathcal{T}} \mathbb{E} \left[\frac{1}{\alpha} (x + Y_\tau)^+ - x \right] \\ &= \inf_{x \leq 0} \sup_{\tau \in \mathcal{T}} \mathbb{E} \left[\frac{1}{\alpha} (x + Y_\tau)^+ - x \right] < \infty. \end{aligned}$$

Let us now consider the case $\Phi = \Phi^{[p]}$ for some $p \in]1, \infty[$. This mapping meets all requirements of Theorem 3.1, and $\Phi^{[p]^*}(x) = \Phi^{[p/(p-1)]}(x^+)$. Then by Theorem 3.1, we have the following primal representation of the corresponding optimal stopping problem.

COROLLARY 3.4. *Let $(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ be atomless with countably generated \mathcal{F}_t for every $t > 0$. If $\sup_{t \in [0, T]} Y_t \in L^{p/(p-1)}$ for some $p \in]1, \infty[$, then*

$$\sup_{\tau \in \mathcal{T}} \rho^{\Phi^{[p]} }(-Y_\tau) = \inf_{x \in \mathbb{R}} \sup_{\tau \in \mathcal{T}} \mathbb{E} \left[\frac{(p-1)((x + Y_\tau)^+)^{p/(p-1)}}{p} - x \right] < \infty.$$

3.2. The existence of solutions. A natural question is whether we can find a real number x^* and a (\mathcal{F}_t) -stopping time τ^* which solve (3.2). We may give a fairly general answer within the context of discrete time optimal stopping problems. In order to be more precise, let $\mathcal{T}_{\mathbb{T}}$ denote all stopping times from \mathcal{T} with values in \mathbb{T} , where \mathbb{T} is any finite subset of $[0, T]$ containing $\{0, T\}$. Consider now the stopping problem

$$(3.3) \quad \text{maximize} \quad \rho_0^\Phi(-Y_\tau) \quad \text{over} \quad \tau \in \mathcal{T}_{\mathbb{T}}.$$

Turning over to the filtration $(\mathcal{F}^\mathbb{T})_{t \in [0, T]}$ defined by $\mathcal{F}_t^\mathbb{T} := \mathcal{F}_{[t]}$ with $[t] := \max\{s \in \mathbb{T} | s \leq t\}$, we see that $(Y^\mathbb{T})_{t \in [0, T]}$ with $Y_t^\mathbb{T} := Y_{[t]}$ describes some

$(\mathcal{F}_t^{\mathbb{T}})$ -adapted process. Hence, we can apply Theorem 3.1 to get

$$(3.4) \quad \begin{aligned} \sup_{\tau \in \mathcal{T}_{\mathbb{T}}} \rho_0^{\Phi}(-Y_{\tau}) &= \sup_{\tau \in \mathcal{T}_{\mathbb{T}}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau}) - x] \\ &= \inf_{x \in \mathbb{R}} \sup_{\tau \in \mathcal{T}_{\mathbb{T}}} \mathbb{E}[\Phi^*(x + Y_{\tau}) - x]. \end{aligned}$$

In this section, we want to find conditions which guarantee the existence of a saddle point for the optimization problems

$$(3.5) \quad \text{maximize} \quad \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau}) - x] \quad \text{over} \quad \tau \in \mathcal{T}_{\mathbb{T}}$$

and

$$(3.6) \quad \text{minimize} \quad \sup_{\tau \in \mathcal{T}_{\mathbb{T}}} \mathbb{E}[\Phi^*(x + Y_{\tau}) - x] \quad \text{over} \quad x \in \mathbb{R}.$$

To this end, we shall borrow some arguments from the theory of Lyapunoff’s theorem for infinite-dimensional vector measures. A central concept in this context is the notion of *thin subsets* of integrable mappings. So let us first recall it. For a fixed probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$, a subset $M \subseteq L^1(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ is called thin if for any $A \in \bar{\mathcal{F}}$ with $\bar{\mathbb{P}}(A) > 0$, there is some nonzero $g \in L^{\infty}(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ vanishing outside A and satisfying $\mathbb{E}[g \cdot Z] = 0$ for every $Z \in M$ (cf. [29], or [1]). Best known examples are finite subsets of $L^1(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ or finite-dimensional linear subspaces of $L^1(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ if $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ is atomless (cf. [29], or [1]).

PROPOSITION 3.5. *Let the assumptions of Theorem 3.1 be fulfilled, and let $\mathbb{T} := \{t_0, \dots, t_{r+1}\}$ with $t_0 = 0 < t_1 < \dots < t_{r+1} = T$. Moreover, let $\{\mathbb{E}[\mathbb{1}_A \cdot \Phi^*(x + Y_s) | \mathcal{F}_t] | x \in \mathbb{R}\}$ be a thin subset of $L^1(\Omega, \mathcal{F}_t, \mathbb{P} |_{\mathcal{F}_t})$ for $s, t \in \mathbb{T}$ with $t \leq s$ and $A \in \mathcal{F}_T$. Then there are $\tau^* \in \mathcal{T}_{\mathbb{T}}$ and $x^* \in \mathbb{R}$ satisfying*

$$\begin{aligned} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau^*}) - x] &= \sup_{\tau \in \mathcal{T}_{\mathbb{T}}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau}) - x] \\ &= \inf_{x \in \mathbb{R}} \sup_{\tau \in \mathcal{T}_{\mathbb{T}}} \mathbb{E}[\Phi^*(x + Y_{\tau}) - x] \\ &= \sup_{\tau \in \mathcal{T}_{\mathbb{T}}} \mathbb{E}[\Phi^*(x^* + Y_{\tau}) - x^*]. \end{aligned}$$

In particular, it holds

$$\mathbb{E}[\Phi^*(x^* + Y_{\tau}) - x^*] \leq \mathbb{E}[\Phi^*(x^* + Y_{\tau^*}) - x^*] \leq \mathbb{E}[\Phi^*(x + Y_{\tau^*}) - x]$$

for any $x \in \mathbb{R}$ and $\tau \in \mathcal{T}_{\mathbb{T}}$.

The proof of Proposition 3.5 can be found in Section 7.5.

EXAMPLE 3.6. Let the mapping $\Phi_e^* : \mathbb{R} \rightarrow \mathbb{R}$ be defined by $\Phi_e^*(y) := \sum_{k=1}^n \alpha_k (\exp(\beta_k y) - 1)$ for some $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n > 0$. Obviously, Φ_e^* is convex, nondecreasing and satisfies $\lim_{y \rightarrow \infty} (\Phi_e^*(y) - y) = \infty$ as well as $\Phi_e^*(0) = 0$. Hence, $\Phi_e(x) := \sup_{y \in \mathbb{R}} (xy - \Phi_e^*(y))$ defines a lower semicontinuous convex

function which satisfies (3.1), and whose Fenchel–Legendre transform coincides with Φ_e^* , since Φ_e^* is continuous. Moreover, for any $s, t \in \mathbb{T}$ such that $t \leq s$, and $A \in \mathcal{F}_T$, the set $\{\mathbb{E}[\mathbb{1}_A \cdot \Phi_e^*(x + Y_s) | \mathcal{F}_t] | x \in \mathbb{R}\}$ is contained in the finite-dimensional linear subspace of $L^1(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ spanned by the sequence of r.v.

$$\{\mathbb{E}[\mathbb{1}_A \cdot \exp(\beta_k Y_s) | \mathcal{F}_t] | k = 0, \dots, n\},$$

where by definition $\beta_0 := 0$. As a result, $\{\mathbb{E}[\mathbb{1}_A \cdot \Phi_e^*(x + Y_s) | \mathcal{F}_t] | x \in \mathbb{R}\}$ is a thin subset of $L^1(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ in the case of atomless $(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ (cf. e.g., [1], Proposition 2.6).

3.3. Additive dual representation. In this section, we generalize the celebrated additive dual representation for optimal stopping problems (see [35]) to the case of optimal stopping under uncertainty. The result in [35] is formulated in terms of martingales M with $M_0 = 0$ satisfying $\sup_{t \in [0, T]} |M_t| \in L^1$. The set of all such adapted martingales will be denoted by \mathcal{M}_0 .

THEOREM 3.7. *Let $V_t := \text{ess sup}_{\tau \in \mathcal{T}, \tau \geq t} \mathbb{E}[Z_\tau | \mathcal{F}_t]$ be the Snell envelope of an integrable right-continuous stochastic process $(Z_t)_{t \in [0, T]}$ adapted to $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$. If $\sup_{t \in [0, T]} |Z_t| \in L^p$ for some $p > 1$, then*

$$V_0 = \sup_{\tau \in \mathcal{T}} \mathbb{E}[Z_\tau] = \inf_{M \in \mathcal{M}_0} \mathbb{E}\left[\sup_{t \in [0, T]} (Z_t - M_t) \right],$$

where the infimum is attained for $M = M^*$ with M^* being the martingale part of the Doob–Meyer decomposition of $(V_t)_{t \in [0, T]}$. Even more it holds

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}[Z_\tau] = \sup_{t \in [0, T]} (Z_t - M_t^*) \quad \text{P-a.s.}$$

REMARK 3.8. By inspection of the proof of Theorem 2.1 in [35], one can see that the assumption $\sup_{t \in [0, T]} \mathbb{E}[Z_t] \in L^p$ for some $p > 1$ is only used to guarantee the existence of the Doob–Meyer decomposition of the Snell envelope $(V_t)_{t \in [0, T]}$. Therefore, this assumption may be relaxed, if we consider discrete time optimal stopping problems on the set \mathbb{T} for some finite $\mathbb{T} \subseteq [0, T]$ containing $\{0, T\}$. In this case, the Doob–Meyer decomposition always exists if $(Z_t)_{t \in \mathbb{T}}$ is integrable, and Theorem 3.7 holds with \mathcal{T} replaced by $\mathcal{T}_{\mathbb{T}}$ and $[0, T]$ replaced by \mathbb{T} (see also [30], Theorem 5.5).

Theorem 3.1 allows us to extend the additive dual representation to the case of stopping problems (2.1). We shall use the following notation. For a fixed Φ and $x \in \mathbb{R}$, we shall denote by $V^{\Phi, x} = (V_t^{\Phi, x})_{t \in [0, T]}$ the Snell-envelope w.r.t. to $(\Phi^*(x + Y_t) - x)_{t \in [0, T]}$ defined via

$$V_t^{\Phi, x} := \text{ess sup}_{\tau \in \mathcal{T}, \tau \geq t} \mathbb{E}[(\Phi^*(x + Y_\tau) - x) | \mathcal{F}_t].$$

The application of Theorem 3.1 together with Theorem 3.7 provides us with the following additive dual representation of the stopping problem (2.1).

THEOREM 3.9. *Under assumptions on Φ and (\mathcal{F}_t) of Theorem 3.1 and under the condition $\sup_{t \in [0, T]} |\Phi^*(x + Y_t)| \in L^p$ for some $p > 1$ and any $x \in \mathbb{R}$, the following dual representation holds*

$$\begin{aligned} \sup_{\tau \in \mathcal{T}} \rho_0^\Phi(-Y_\tau) &= \inf_{x \in \mathbb{R}} \inf_{M \in \mathcal{M}_0} \mathbb{E} \left[\sup_{t \in [0, T]} (\Phi^*(x + Y_t) - x - M_t) \right] \\ &= \inf_{x \in \mathbb{R}} \mathbb{E} \left[\sup_{t \in [0, T]} (\Phi^*(x + Y_t) - x - M_t^{*, \Phi, x}) \right] \\ &= \text{ess inf}_{x \in \mathbb{R}} \sup_{t \in [0, T]} (\Phi^*(x + Y_t) - x - M_t^{*, \Phi, x}) \quad \text{P-a.s.} \end{aligned}$$

Here, $M^{*, \Phi, x}$ stands for the martingale part of the Doob–Meyer decomposition of the Snell-envelope $V^{\Phi, x}$.

REMARK 3.10. Under the assumptions of Theorem 3.1, we have that $\sup_{t \in [0, T]} Y_t \in H^{\Phi^*}$. Furthermore, Φ^* is convex and nondecreasing with $\Phi^*(0) = 0$ (see Lemma A.1 in Appendix A) so that for any $y < 0$

$$|\Phi^*(y)| = \int_y^0 \Phi^{*'}(z) dz \leq \Phi^{*'}(0)|y| \leq \int_0^{|y|} \Phi^{*'}(z) dz = \Phi^*(|y|),$$

where $\Phi^{*'}$ denotes the right-sided derivative of Φ^* . Using the monotonicity of Φ^* again, we conclude that

$$|\Phi^*(x + Y_t)| \leq \Phi^*(|x| + Y_t) \leq \Phi^* \left(|x| + \sup_{t \in [0, T]} Y_t \right) \in L^1$$

for all $x \in \mathbb{R}$ and $t \in [0, T]$. Hence, the application of Theorem 3.9 to (3.3) is already possible under the assumptions of Theorem 3.1.

The dual representation for the optimal stopping problem under Average Value at Risk reads as follows.

COROLLARY 3.11. *Let the assumptions on Φ and (\mathcal{F}_t) be as in Theorem 3.1. If $\sup_{t \in [0, T]} Y_t \in L^p$ for some $p > 1$, then it holds P-a.s.*

$$\begin{aligned} &\sup_{\tau \in \mathcal{T}} AV@R_\alpha(-Y_\tau) \\ (3.7) \quad &= \inf_{x \in \mathbb{R}} \inf_{M \in \mathcal{M}_0} \mathbb{E} \left[\sup_{t \in [0, T]} \left(\frac{1}{\alpha} (x + Y_t)^+ - x - M_t \right) \right] \\ &= \inf_{x \leq 0} \mathbb{E} \left[\sup_{t \in [0, T]} \left(\frac{1}{\alpha} (x + Y_t)^+ - x - M_t^{*, \alpha, x} \right) \right] \\ &= \text{ess inf}_{x \leq 0} \sup_{t \in [0, T]} \left(\frac{1}{\alpha} (x + Y_t)^+ - x - M_t^{*, \alpha, x} \right) \quad \text{P-a.s.} \end{aligned}$$

Here, $M^{*,\alpha,x}$ denotes the martingale part of the Doob–Meyer decomposition of the Snell-envelope $V^{\Phi_\alpha,x}$.

REMARK 3.12. Let us consider a discrete time optimal stopping problem $\sup_{\tau \in \mathbb{T}} AV@R_\alpha(-Y_\tau)$ for some finite $\mathbb{T} \subseteq [0, T]$ with $\{0, T\} \in \mathbb{T}$. In view of Remark 3.10, the assumptions of Theorem 3.1 are already sufficient to obtain the dual representation (3.7) with \mathcal{T} replaced by $\mathcal{T}_\mathbb{T}$ and $[0, T]$ replaced by \mathbb{T} .

4. Discussion. In [30], the optimal stopping problems of the type

$$(4.1) \quad \sup_{\tau \in \mathcal{T}} \mathcal{U}_0(Y_\tau)$$

were studied, where for any $t \geq 0$, the functional \mathcal{U}_t maps a linear subspace \mathcal{X} of the space L^0 into $\mathcal{X} \cap L^0(\Omega, \mathcal{F}_t, P|_{\mathcal{F}_t})$ and satisfies $\mathcal{U}_t(X) \leq \mathcal{U}_t(Y)$ for $X \leq Y$ P-a.s. In fact, there is a one-to-one correspondence between conditional convex risk measures $(\rho_t)_{t \in [0, T]}$ and dynamic utility functionals $\mathcal{U} := (\mathcal{U}_t)_{t \in [0, T]}$ satisfying the following two properties:

- *conditional translation invariance:*
 $\mathcal{U}_t(X + Y) = \mathcal{U}_t(X) + Y$ for $Y \in \mathcal{X} \cap L^0(\Omega, \mathcal{F}_t, P|_{\mathcal{F}_t})$ and $X \in \mathcal{X}$,
- *conditional concavity:*
 $\mathcal{U}_t(\lambda X + (1 - \lambda)Y) \geq \lambda \mathcal{U}_t(X) + (1 - \lambda)\mathcal{U}_t(Y)$ for $X, Y \in \mathcal{X}$ and $\lambda \in \mathcal{X} \cap L^0(\Omega, \mathcal{F}_t, P|_{\mathcal{F}_t})$ with $0 \leq \lambda \leq 1$.

More precisely, any conditionally translation invariant and conditionally concave dynamic utility functional $(\mathcal{U}_t)_{t \in [0, T]}$ defines a family $(\rho_t^\mathcal{U})_{t \in [0, T]}$ of conditional convex risk measures via $\rho_t^\mathcal{U}(X) = -\mathcal{U}_t(X)$ and vice versa. The results of [30] essentially rely on the following additional assumptions:

- *regularity:*
 $\mathcal{U}_t(\mathbb{1}_A X) = \mathbb{1}_A \cdot \mathcal{U}_t(X)$ for $A \in \mathcal{F}_t$ and $X \in \mathcal{X}$,
- *recursiveness:*
 $\mathcal{U}_s \circ \mathcal{U}_t = \mathcal{U}_s$ for $s \leq t$.

Recursiveness is often also referred to as time consistency. Obviously, the dynamic utility functional $(\mathcal{U}_t^{\Phi_\alpha})_{t \in [0, T]}$, defined by $\mathcal{U}_t^{\Phi_\alpha}(X) := AV@R_\alpha(-X|\mathcal{F}_t)$, satisfies the regularity and the conditional translation invariance, but it fails to be recursive (cf. [25], Example 11.13). Even worse, according to Theorem 1.10 in [32] for any $\alpha < 1$, there is in general no regular conditionally translation invariant and recursive dynamic utility functional \mathcal{U} such that $\mathcal{U}_0 = \mathcal{U}_0^{\Phi_\alpha}$. This means that we cannot in general reduce the stopping problem (2.3) to the stopping problem (4.1) with a regular, conditionally translation invariant and recursive dynamic utility functional \mathcal{U} . Note that this conclusion can be drawn from Theorem 1.10 of [32], because $AV@R_\alpha$ is law-invariant, that is, $AV@R_\alpha(X) = AV@R_\alpha(Y)$ for identically distributed X and Y , and satisfies the properties $AV@R_\alpha(0) = 0$ as well as $AV@R_\alpha(-\varepsilon \mathbb{1}_A) > 0$ for any $\varepsilon > 0$ and $A \in \mathcal{F}$ with $P(A) > 0$.

The stopping problem (2.3) may also be viewed as a special case of the following stopping problem:

$$(4.2) \quad \sup_{\tau \in \mathcal{T}} \int_0^\infty w(\mathbb{P}(Y_\tau > x)) dx,$$

where $w : [0, 1] \mapsto [0, 1]$ is the so-called distortion function, that is, w is nondecreasing and satisfies $w(0) = 0$, $w(1) = 1$. Indeed, if for $\alpha \in]0, 1[$ the distortion function w_α is defined by $w_\alpha(u) := \frac{u}{\alpha} \wedge 1$, then the stopping problems (2.3) and (4.2) coincide. Recalling Theorem 1.10 of [32] again, we see that the stopping problem (4.2) is not in general representable in the form (4.1) with some regular, conditionally translation invariant and recursive dynamic utility functional. The stopping problem (4.2) was recently considered by [37]. However, the analysis in [37] relies on some additional assumptions. First of all, the authors allow for all finite stopping times w.r.t. to some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ instead of restricting to those which are bounded by a fixed number. Second, they assume a special structure for the process $(Y_t)_{t \geq 0}$, namely it is supposed that $Y_t = u(S_t)$ for an absolutely continuous nonnegative function u on $[0, \infty[$ and for a one-dimensional geometric Brownian motion $(S_t)_{t \geq 0}$. Third, the authors focus on strictly increasing absolutely continuous distortion functions w so that their analysis does not cover the case of Average Value at Risk. More precisely, in [37] the optimal stopping problems of the form

$$(4.3) \quad \sup_{\tau \in \mathcal{T}^\infty} D_w(u(S_\tau)) = \sup_{\tau \in \mathcal{T}^\infty} \int_0^\infty w(\mathbb{P}(u(S_\tau) > x)) dx,$$

are studied, where \mathcal{T}^∞ denotes the set of all finite stopping times. A crucial step in the authors' argumentation is the reformulation of the optimal stopping problem (4.3) as

$$\begin{aligned} \sup_{\tau \in \mathcal{T}^\infty} D_w(u(S_\tau)) &= \sup_{F \in \mathcal{D}} \int_0^\infty w(1 - F(x)) u'(x) dx \\ &= \sup_{F \in \mathcal{D}} \int_0^1 u(F^{\leftarrow}(u)) w'(1 - u) du, \end{aligned}$$

where u' and w' are derivatives of u and w , respectively, and \mathcal{D} denotes the set of all distribution functions F with a nonnegative support such that $\int_0^\infty (1 - F(x)) dx \leq S_0$. The main idea of the approach in [37] is that any such distribution function may be described as the distribution function of S_τ for some finite stopping time $\tau \in \mathcal{T}^\infty$ and this makes the application of the Skorokhod embedding technique possible. Hence, the results essentially rely on the special structure of the stochastic process $(Y_t)_{t \geq 0}$ and seem to be not extendable to stochastic processes of the form $Y_t = U(X_t)$, where $(X_t)_{t \geq 0}$ is a multivariate Markov process. Moreover, it remains unclear whether the analysis of [37] can be carried over to the case of bounded stopping times, as the Skorokhod embedding cannot be applied to the general sets of stopping times \mathcal{T} (see, e.g., [3]).

5. Numerical example. In this section, we illustrate how our results can be used to price Bermudan-type options in uncertain environment. Specifically, we consider the model with d identically distributed assets, where each underlying has dividend yield δ . The dynamic of assets is given by

$$(5.1) \quad \frac{dX_t^i}{X_t^i} = (r - \delta) dt + \sigma dW_t^i, \quad i = 1, \dots, d,$$

where $W_t^i, i = 1, \dots, d$, are independent one-dimensional Brownian motions and r, δ, σ are constants. At any time $t \in \{t_0, \dots, t_J\}$ the holder of the option may exercise it and receive the payoff

$$Y_t = G(X_t) = e^{-rt} (\max(X_t^1, \dots, X_t^d) - K)^+.$$

If we are uncertain about our modelling assumption and if the Average Value at Risk is used to measure the risk related to this uncertainty, then the risk-adjusted price of the option is given by

$$(5.2) \quad \begin{aligned} \sup_{\tau \in \mathcal{T}[t_0, \dots, t_J]} AV@R_\alpha(-Y_\tau) &= \sup_{\tau \in \mathcal{T}} \sup_{Q \in \mathcal{Q}_{\Phi_\alpha, 0}} \mathbb{E}_Q[-Y_\tau] \\ &= \inf_{x \leq 0} \sup_{\tau \in \mathcal{T}[t_0, \dots, t_J]} \mathbb{E} \left[\frac{1}{\alpha} (x + Y_\tau)^+ - x \right], \end{aligned}$$

where $\mathcal{Q}_{\Phi_\alpha, t}$ consists of all probability measures Q on \mathcal{F} with

$$(5.3) \quad \left. \frac{dQ}{dP} \right|_{\mathcal{F}_t} \leq 1/\alpha, \quad P|_{\mathcal{F}_t}\text{-a.s.}$$

If we restrict our attention to the class of generalized Black–Scholes models of the type

$$dX_t^i = X_t^i (\alpha_t^i dt + \sigma_t^i dW_t^i), \quad i = 1, \dots, d$$

with adapted processes $(\alpha_t^i), (\sigma_t^i)$ and independent Brownian motions W_t^1, \dots, W_t^d , then

$$\left. \frac{dQ}{dP} \right|_{\mathcal{F}_t} = \exp \left(- \sum_{i=1}^d \int_0^t \theta_s^i dW_s^i - \frac{1}{2} \sum_{i=1}^d \int_0^t (\theta_s^i)^2 ds \right)$$

with $\theta_t^i = (\alpha_t^i - r + \delta)/\sigma_t^i$ and the condition (5.3) transforms to

$$\exp \left(- \sum_{i=1}^d \int_0^t \theta_s^i dW_s^i - \frac{1}{2} \sum_{i=1}^d \int_0^t (\theta_s^i)^2 ds \right) \leq 1/\alpha, \quad P|_{\mathcal{F}_t}\text{-a.s.}$$

Due to Corollary 3.3, one can use the standard methods based on dynamic programming principle to solve (5.2). Indeed, for any fixed x , the optimal value of the stopping problem

$$V = \sup_{\tau \in \mathcal{T}[t_0, \dots, t_J]} \mathbb{E} \left[\frac{1}{\alpha} (x + Y_\tau)^+ - x \right]$$

can be, for example, numerically approximated via the well-known regression methods like Longstaff–Schwartz method. In this way, one can get a (suboptimal) stopping rule

$$\hat{\tau}_x := \inf\{0 \leq j \leq J : (x + Y_{t_j})^+ / \alpha - x \geq \hat{C}_j(X_{t_j}, x)\},$$

where $\hat{C}_1, \dots, \hat{C}_J$ are continuation values estimates. Then

$$(5.4) \quad V_N^l := \inf_{x \leq 0} \left\{ \frac{1}{N} \sum_{n=1}^N (x + Y_{t_{\frac{x}{\alpha}}^{(n)}})^+ / \alpha - x \right\}$$

is a low-biased estimate for V . Note that the infimum in (5.4) can be easily computed using a simple search algorithm. An upper-biased estimate can be constructed using the well-known Andersen–Broadie dual approach (see [2]). For any fixed $x \leq 0$, this approach would give us a discrete time martingale $(M_j^x)_{j=0, \dots, J}$ which in turn can be used to build an upper-biased estimate via the representation (3.7):

$$(5.5) \quad V_N^u := \inf_{x \leq 0} \left\{ \sum_{n=1}^N \left[\sup_{j=0, \dots, J} \left(\frac{1}{\alpha} (x + Y_{t_j}^{(n)})^+ - x - M_j^{x, (n)} \right) \right] \right\}.$$

Note that (5.5) remains upper biased even if we replace the infimum of the objective function in (5.5) by its value at a fixed point x . In Table 1, we present the bounds V_N^l and V_N^u together with their standard deviations for different values of α . As to implementation details, we used 12 basis functions for regression (see [2]) and 10^4 training paths to compute $\hat{C}_1, \dots, \hat{C}_J$. In the dual approach of Andersen and Broadie, 10^3 inner simulations were done to approximate M^x . In both cases, we simulated $N = 10^4$ testing paths to compute the final estimates.

For comparison, let us consider a problem of pricing the above Bermudan option under entropic risk measure (2.4). Due to (2.5), we need to solve the optimal stopping problem

$$V^\gamma = \sup_{\tau \in \mathcal{T}[t_0, \dots, t_J]} \mathbb{E}[\exp(\gamma Y_\tau)].$$

TABLE 1

Bounds (with standard deviations) for 2-dimensional Bermudan max-call with parameters $K = 100$, $r = 0.05$, $\sigma = 0.2$, $\delta = 0.1$ under AV@R at level α

| α | Lower bound V_N^l | Upper bound V_N^u |
|----------|---------------------|---------------------|
| 0.33 | 23.64 (0.026) | 23.92 (0.108) |
| 0.50 | 16.06 (0.019) | 16.12 (0.045) |
| 0.67 | 12.05 (0.014) | 12.09 (0.034) |
| 0.75 | 10.71 (0.013) | 10.75 (0.030) |

TABLE 2
Bounds (with standard deviations) for 2-dimensional Bermudan max-call with parameters $K = 100, r = 0.05, \sigma = 0.2, \delta = 0.1$ under entropic risk measure with parameter γ

| γ | Lower bound | Upper bound |
|----------|------------------|------------------|
| 0.0025 | 8.218979 (0.011) | 8.262082 (0.029) |
| 0.005 | 8.399141 (0.015) | 8.454748 (0.032) |
| 0.01 | 8.797425 (0.017) | 8.888961 (0.041) |
| 0.02 | 9.698094 (0.020) | 10.03958 (0.058) |
| 0.03 | 12.72327 (0.020) | 12.74784 (0.072) |
| 0.04 | 17.47090 (0.022) | 17.50481 (0.095) |

The latter problem can be solved via the standard dynamic programming combined with regression as described above. In Table 2, the upper and lower MC bounds for $\log(V)/\gamma$ are presented for different values of the parameter γ . Unfortunately, for larger values of γ , the corresponding MC estimates become unstable due to the presence of exponent in (2.5). In Figure 1, the lower bounds for AV@R and the entropic risk measure are shown graphically. As can be seen the quality of upper and lower bounds are quite similar. However, due to above mentioned instability, AV@R should be preferred under higher uncertainty.

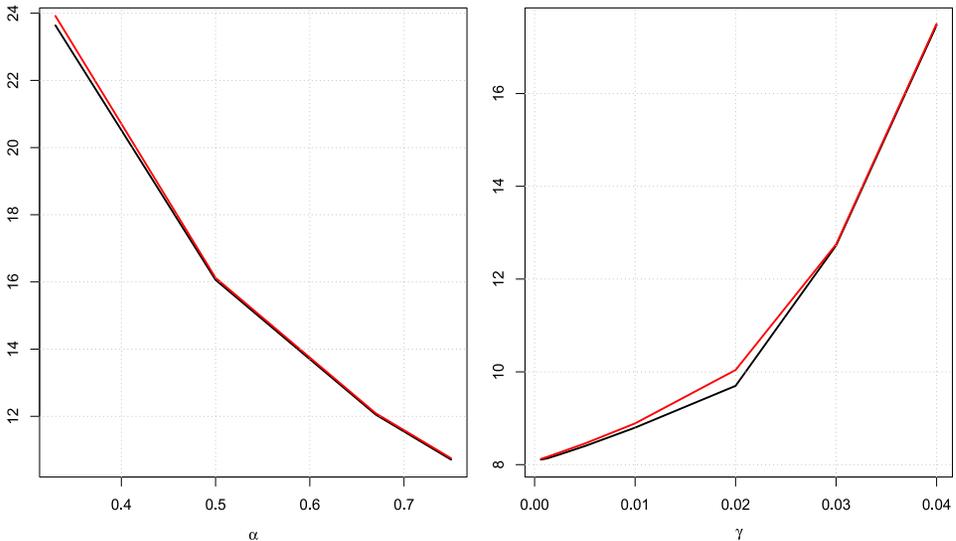


FIG. 1. Lower and upper bounds for Bermudan option prices under AV@R (left) and entropic risk (right) measures.

6. The optimal stopping problem with randomized stopping times. In order to prove Theorem 3.1, we shall proceed as follows. First, by Lemma A.1 (cf. Appendix A), we obtain immediately

$$(6.1) \quad \sup_{\tau \in \mathcal{T}} \sup_{Q \in \mathcal{Q}_{\Phi,0}} \left(\mathbb{E}_Q[Y_\tau] - \mathbb{E} \left[\Phi \left(\frac{dP}{dQ} \right) \right] \right) = \sup_{\tau \in \mathcal{T}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x].$$

The proof of Theorem 3.1 would be completed, if we can show that

$$(6.2) \quad \sup_{\tau \in \mathcal{T}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] = \inf_{x \in \mathbb{R}} \sup_{\tau \in \mathcal{T}} \mathbb{E}[\Phi^*(x + Y_\tau) - x].$$

Using Fubini's theorem, we obtain for any $\tau \in \mathcal{T}$ and every $x \in \mathbb{R}$

$$\mathbb{E}[\Phi^*((x + Y_\tau)^+) - x] = \int_{x^-}^{\infty} \Phi^{*'}(x + z)[1 - F_{Y_\tau}(z)] dz + \Phi^*(x^+) - x,$$

where F_{Y_τ} stands for the distribution function of Y_τ and $\Phi^{*'}$ denotes the right-sided derivative of the convex function Φ^* . In the same way, we may also find

$$\mathbb{E}[\Phi^*(-(x + Y_\tau)^-)] = - \int_0^{x^-} \Phi^{*'}(x + z) F_{Y_\tau}(z) dz.$$

Hence, the property $\Phi^*(x) = \Phi^*(x^+) + \Phi^*(-x^-)$ for $x \in \mathbb{R}$ yields

$$(6.3) \quad \mathbb{E}[\Phi^*(x + Y_\tau) - x] = \int_0^{\infty} \Phi^{*'}(x + z)[1 - F_{Y_\tau}(z)] dz + \Phi^*(x) - x$$

for $\tau \in \mathcal{T}$ and $x \in \mathbb{R}$. Since the set $\mathbb{F} := \{F_{Y_\tau} | \tau \in \mathcal{T}\}$ of distribution functions F_{Y_τ} of Y_τ is not, in general, a convex subset of the set of distribution functions on \mathbb{R} , we cannot apply the known minimax results. The idea is to first establish (6.2) for the larger class of randomized stopping times, and then to show that the optimal value coincides with the optimal value $\sup_{\tau \in \mathcal{T}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x]$.

Let us recall the notion of randomized stopping times. By definition (see, e.g., [19]), a randomized stopping time w.r.t. $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ is a mapping $\tau^r : \Omega \times [0, 1] \rightarrow [0, \infty]$ which is nondecreasing and left-continuous in the second component such that $\tau^r(\cdot, u)$ is a stopping time w.r.t. $(\mathcal{F}_t)_{t \in [0, T]}$ for any $u \in [0, 1]$. Notice that any randomized stopping time τ^r is also an ordinary stopping time w.r.t. the enlarged filtered probability space $(\Omega \times [0, 1], \mathcal{F} \otimes \mathcal{B}([0, 1]), (\mathcal{F}_t \otimes \mathcal{B}([0, 1]))_{t \in [0, T]}, P \otimes P^U)$. Here, P^U denotes the uniform distribution on $[0, 1]$, defined on $\mathcal{B}([0, 1])$, the usual Borel σ -algebra on $[0, 1]$. We shall call a randomized stopping time τ^r to be degenerated if $\tau^r(\omega, \cdot)$ is constant for every $\omega \in \Omega$. There is an obvious one-to-one correspondence between stopping times and degenerated randomized stopping times.

Consider the stochastic process $(Y_t^r)_{t \geq 0}$, defined by

$$Y_t^r : \Omega \times [0, 1] \rightarrow \mathbb{R}, \quad (\omega, u) \mapsto Y_t(\omega),$$

which is adapted w.r.t. the enlarged filtered probability space. Denoting by \mathcal{T}^r the set of all randomized stopping times $\tau^r \leq T$, we shall study the following new stopping problem:

$$(6.4) \quad \text{maximize} \quad \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau^r}^r) - x] \quad \text{over} \quad \tau^r \in \mathcal{T}^r.$$

Obviously, $\inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] = \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau^r}^r) - x]$ is valid for every stopping time $\tau \in \mathcal{T}$, where $\tau^r \in \mathcal{T}^r$ is the corresponding degenerated randomized stopping time such that $\tau^r(\omega, u) = \tau(\omega)$, $u \in [0, 1]$. Thus, in general the optimal value of the stopping problem (6.4) is at least as large as the one of the original stopping problem (2.1) due to (6.1). One reason to consider the new stopping problem (6.4) is that it has a solution under fairly general conditions.

PROPOSITION 6.1. *Let $(Y_t)_{t \in [0, T]}$ be quasi-left-continuous, defined to mean $Y_{\tau_n} \rightarrow Y_\tau$ P-a.s. whenever $(\tau_n)_{n \in \mathbb{N}}$ is a sequence in \mathcal{T} satisfying $\tau_n \nearrow \tau$ for some $\tau \in \mathcal{T}$. If $\sup_{t \in [0, T]} Y_t \in H^{\Phi^*}$ and if \mathcal{F}_T is countably generated, then there exists a randomized stopping time $\tau_*^r \in \mathcal{T}^r$ such that*

$$\inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau_*^r}^r) - x] = \sup_{\tau^r \in \mathcal{T}^r} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau^r}^r) - x].$$

Moreover, the following important minimax result for the stopping problem (6.4) holds.

PROPOSITION 6.2. *If (3.1) is fulfilled, and if $\sup_{t \in [0, T]} Y_t \in H^{\Phi^*}$, then*

$$\sup_{\tau^r \in \mathcal{T}^r} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau^r}^r) - x] = \inf_{x \in \mathbb{R}} \sup_{\tau^r \in \mathcal{T}^r} \mathbb{E}[\Phi^*(x + Y_{\tau^r}^r) - x].$$

Moreover, if $(Y_t)_{t \in [0, T]}$ is quasi-left-continuous and if \mathcal{F}_T is countably generated, then there exist $\tau^{r*} \in \mathcal{T}^r$ and $x^* \in \mathbb{R}$ such that

$$\mathbb{E}[\Phi^*(x^* + Y_{\tau^{r*}}^r) - x^*] \leq \mathbb{E}[\Phi^*(x^* + Y_{\tau^{r*}}^r) - x^*] \leq \mathbb{E}[\Phi^*(x + Y_{\tau^{r*}}^r) - x]$$

for $x \in \mathbb{R}$ and $\tau \in \mathcal{T}^r$.

The proof of Proposition 6.2 can be found in Section 7.2. In the next step, we shall provide conditions ensuring that the stopping problems (2.1) and (6.4) have the same optimal value.

PROPOSITION 6.3. *Let $(\Omega, \mathcal{F}_t, P|_{\mathcal{F}_t})$ be atomless with countably generated \mathcal{F}_t for every $t > 0$. If (3.1) is fulfilled, and if $\sup_{t \in [0, T]} Y_t$ belongs to H^{Φ^*} , then*

$$\begin{aligned} \sup_{\tau^r \in \mathcal{T}^r} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau^r}^r) - x] &= \sup_{\tau \in \mathcal{T}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] \\ &= \sup_{\tau \in \mathcal{T}} \sup_{Q \in \mathcal{Q}_{\Phi, 0}} \left(\mathbb{E}_Q[Y_\tau] - \mathbb{E} \left[\Phi \left(\frac{dP}{dQ} \right) \right] \right). \end{aligned}$$

The proof of Proposition 6.3 is delegated to Section 7.3.

7. Proofs. We shall start with some preparations which also will turn out to be useful later on. Let us recall (cf. [19]) that every $\tau^r \in \mathcal{T}^r$ induces a stochastic kernel $K_{\tau^r} : \Omega \times \mathcal{B}([0, T]) \rightarrow [0, 1]$ with $K_{\tau^r}(\omega, \cdot)$ being the distribution of $\tau^r(\omega, \cdot)$ under \mathbb{P}^U for any $\omega \in \Omega$. Here $\mathcal{B}([0, T])$ stands for the usual Borel σ -algebra on $[0, T]$. This stochastic kernel has the following properties:

$$\begin{aligned} K_{\tau^r}(\cdot, [0, t]) &\text{ is } \mathcal{F}_t\text{-measurable for every } t \geq 0, \\ K_{\tau^r}(\omega, [0, t]) &= \sup\{u \in [0, 1] \mid \tau^r(\omega, u) \leq t\}. \end{aligned}$$

The associated stochastic kernel K_{τ^r} is useful to characterize the distribution function $F_{Y_{\tau^r}^r}$ of $Y_{\tau^r}^r$.

LEMMA 7.1. *For any $\tau^r \in \mathcal{T}^r$ with associated stochastic kernel K_{τ^r} , the distribution function $F_{Y_{\tau^r}^r}$ of $Y_{\tau^r}^r$ may be represented in the following way:*

$$F_{Y_{\tau^r}^r}(x) = \mathbb{E}[K_{\tau^r}(\cdot, \{t \in [0, T] \mid Y_t \leq x\})] \quad \text{for } x \in \mathbb{R}.$$

PROOF. Let $\tau^r \in \mathcal{T}^r$, and let us fix $x \in \mathbb{R}$. Then

$$\begin{aligned} F_{Y_{\tau^r}^r}(x) &= \mathbb{E}[\mathbb{1}_{]-\infty, x]}(Y_{\tau^r}^r)] = \int_0^1 \mathbb{E}[\mathbb{1}_{]-\infty, x]}(Y_{\tau^r}^r(\cdot, u))] du \\ &= \mathbb{E}\left[\int_0^1 \mathbb{1}_{]-\infty, x]}(Y_{\tau^r}^r(\cdot, u)) du\right] \end{aligned}$$

holds (cf. [19], Theorem 4.5), where the last equation on the right-hand side is due to the Fubini–Tonelli theorem. Then by definition of K_{τ^r} , we obtain for every $\omega \in \Omega$

$$\begin{aligned} \int_0^1 \mathbb{1}_{]-\infty, x]}(Y_{\tau^r}^r(\omega, u)) du &= \mathbb{E}_{\mathbb{P}^U}[\mathbb{1}_{]-\infty, x]}(Y_{\tau^r}^r(\omega, \cdot)(\omega))] \\ &= \mathbb{P}^U(\{Y_{\tau^r}^r(\omega, \cdot)(\omega) \leq x\}) \\ &= K_{\tau^r}(\omega, \{t \in [0, T] \mid Y_t(\omega) \leq x\}). \end{aligned}$$

This completes the proof. \square

7.1. Proof of Proposition 6.1. Let us introduce the filtered probability space $(\Omega, \mathcal{F}, (\tilde{\mathcal{F}}_t)_{0 \leq t \leq \infty}, \mathbb{P})$ defined by

$$\tilde{\mathcal{F}}_t = \begin{cases} \mathcal{F}_t, & t \leq T, \\ \mathcal{F}_T, & t > T. \end{cases}$$

We shall denote by $\tilde{\mathcal{T}}^r$ the set of randomized stopping times according to $(\Omega, \mathcal{F}, (\tilde{\mathcal{F}}_t)_{0 \leq t \leq \infty}, \mathbb{P})$. Furthermore, we may extend the processes $(Y_t)_{t \in [0, T]}$ and $(Y_t^r)_{t \in [0, T]}$ to right-continuous processes $(\tilde{Y}_t)_{t \in [0, \infty]}$ and $(\tilde{Y}_t^r)_{t \in [0, T]}$ in the following way:

$$\tilde{Y}_t = \begin{cases} Y_t, & t \leq T, \\ Y_T, & t > T, \end{cases} \quad \text{and} \quad \tilde{Y}_t^r = \begin{cases} Y_t^r, & t \leq T, \\ Y_T^r, & t > T. \end{cases}$$

Recall that we may equip $\tilde{\mathcal{T}}^r$ with the so called Baxter–Chacon topology which is compact in general, and even metrizable within our setting because \mathcal{F}_T is assumed to be countably generated (cf. Theorem 1.5 in [4] and discussion afterwards).

Next, consider the mapping

$$\tilde{h} : \tilde{\mathcal{T}}^r \times \mathbb{R} \rightarrow \mathbb{R}, \quad (\tilde{\tau}^r, x) \mapsto \mathbb{E}[\Phi^*(x + \tilde{Y}_{\tilde{\tau}^r}^r) - x].$$

By assumption on $(Y_t)_{t \in [0, T]}$, the processes $(\tilde{Y}_t)_{t \in [0, \infty]}$ and $(\tilde{Y}_t^r)_{t \in [0, T]}$ are quasi-left-continuous. Moreover, Φ^* is continuous due to Lemma A.1, (i) in Appendix A, so that $(\Phi^*(x + \tilde{Y}_t^r) - x)_{t \in [0, T]}$ is a quasi-left-continuous and right-continuous adapted process. Hence, in view of [19], Theorem 4.7, the mapping $\tilde{h}(\cdot, x)$ is continuous w.r.t. the Baxter–Chacon topology for every $x \in \mathbb{R}$, and thus $\inf_{x \in \mathbb{R}} h(\cdot, x)$ is upper semicontinuous w.r.t. the Baxter–Chacon topology. Then by compactness of the Baxter–Chacon topology, we may find some randomized stopping time $\tilde{\tau}^r \in \tilde{\mathcal{T}}^r$ such that

$$\inf_{x \in \mathbb{R}} h(\tilde{\tau}^r, x) = \sup_{\tilde{\tau}^r \in \tilde{\mathcal{T}}^r} \inf_{x \in \mathbb{R}} h(\tilde{\tau}^r, x).$$

This completes the proof because $\tilde{Y}_{\tilde{\tau}^r}^r = Y_{\tilde{\tau}^r \wedge T}^r$ and $\tilde{\tau}^r \wedge T$ belongs to \mathcal{T}^r for every $\tilde{\tau}^r \in \tilde{\mathcal{T}}^r$. □

7.2. *Proof of Proposition 6.2.* Let us define the mapping $h : \mathcal{T}^r \times \mathbb{R} \rightarrow \mathbb{R}$ by

$$h(\tau^r, x) := \mathbb{E}[\Phi^*(x + Y_{\tau^r}^r) - x].$$

Since $\sup_{t \in [0, T]} Y_t$ is assumed to belong to H^{Φ^*} , the mapping $\sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, \cdot)$ is finite and convex, and thus continuous. Moreover, by Lemma A.1 (cf. Appendix A)

$$\begin{aligned} \lim_{x \rightarrow -\infty} \sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, x) &\geq \lim_{x \rightarrow -\infty} (\Phi^*(x) - x) = \infty = \lim_{x \rightarrow \infty} (\Phi^*(x) - x) \\ &\leq \lim_{x \rightarrow \infty} \sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, x). \end{aligned}$$

Hence, $\inf_{x \in \mathbb{R}} \sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, x) = \inf_{x \in [-\varepsilon, \varepsilon]} \sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, x)$ for some $\varepsilon > 0$. Thus, $\sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, \cdot)$ attains its minimum at some x^* due to continuity of $\sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, \cdot)$. Moreover, if $(Y_t)_{t \in [0, T]}$ is quasi-left-continuous and if \mathcal{F}_T is countably generated, then $\inf_{x \in \mathbb{R}} h(\tau^{r*}, x) = \sup_{\tau^r \in \mathcal{T}^r} \inf_{x \in \mathbb{R}} h(\tau^r, x)$ for some $\tau^{r*} \in \mathcal{T}^r$ due to Proposition 6.1. It remains to show that $\sup_{\tau^r \in \mathcal{T}^r} \inf_{x \in \mathbb{R}} h(\tau^r, x) = \inf_{x \in \mathbb{R}} \sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, x)$. Following the same line of reasoning as for the derivation of (6.3), we may rewrite h in the following way.

$$(7.1) \quad h(\tau^r, x) = \int_0^\infty \Phi^{*'}(x + z)[1 - F_{Y_{\tau^r}^r}(z)] dz + \Phi^*(x) - x,$$

where $F_{Y_{\tau^r}^r}$ stands for the distribution function of $Y_{\tau^r}^r$, and $\Phi^{*'}$ denotes the right-sided derivative of the convex function Φ^* . Obviously, we have

$$(7.2) \quad h(\tau^r, \cdot) \text{ is convex and, therefore, continuous for every } \tau^r \in \mathcal{T}^r.$$

Set $\beta := \inf_{x \in \mathbb{R}} \sup_{\tau^r \in \mathcal{T}^r} \mathbb{E}[\Phi^*(x + Y_{\tau^r}^r) - x] + 1 = \inf_{x \in \mathbb{R}} \sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, x) + 1$ which is a real number because $\sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, \cdot)$ has been already proved to be a finite function which attains its minimum on some compact interval of \mathbb{R} . Furthermore, we may conclude from $h(\tau^r, x) \geq \Phi^*(x) - x$ for $x \in \mathbb{R}$ that

$$(7.3) \quad I_\beta := \{x \in \mathbb{R} \mid \Phi^*(x) - x \leq \beta\} \text{ is a compact interval,}$$

and

$$(7.4) \quad h(\tau^r, x) > \beta \quad \text{for } \tau^r \in \mathcal{T}^r \text{ and } x \in \mathbb{R} \setminus I_\beta.$$

By (7.4), we verify

$$\sup_{\tau^r \in \mathcal{T}^r} \inf_{x \in \mathbb{R}} h(\tau^r, x) = \sup_{\tau^r \in \mathcal{T}^r} \inf_{x \in I_\beta} h(\tau^r, x)$$

and

$$\inf_{x \in \mathbb{R}} \sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, x) = \inf_{x \in I_\beta} \sup_{\tau^r \in \mathcal{T}^r} h(\tau^r, x).$$

We want to apply Fan’s minimax theorem (cf. [22], Theorem 2 or [12]) to $h|_{\mathcal{T}^r \times I_\beta}$. In view of (7.2) and (7.3), it remains to show that for every $\tau_1^r, \tau_2^r \in \mathcal{T}^r$, and any $\lambda \in]0, 1[$ there exists some $\tau^r \in \mathcal{T}^r$ such that

$$(7.5) \quad \lambda h(\tau_1^r, x) + (1 - \lambda)h(\tau_2^r, x) \leq h(\tau^r, x) \quad \text{for all } x \in \mathbb{R}.$$

To this end, let $\tau_1^r, \tau_2^r \in \mathcal{T}^r$ with associated stochastic kernels $K_{\tau_1^r}, K_{\tau_2^r}$, and $\lambda \in]0, 1[$. First, $K := \lambda K_{\tau_1^r} + (1 - \lambda)K_{\tau_2^r} : \Omega \times \mathcal{B}([0, T]) \rightarrow [0, 1]$ defines a stochastic kernel satisfying

$$\begin{aligned} K(\cdot, [0, t]) &\text{ is } \mathcal{F}_t\text{-measurable for every } t \in [0, T], \\ K(\omega, [0, T]) &= 1. \end{aligned}$$

Then

$$\tau^r(\omega, u) := \inf\{t \in [0, T] \mid K(\omega, [0, t]) \geq u\}$$

defines some $\tau^r \in \mathcal{T}^r$ with $K_{\tau^r} = K$. Furthermore, we obtain

$$F_{Y_{\tau^r}^r} = \lambda F_{Y_{\tau_1^r}^r} + (1 - \lambda)F_{Y_{\tau_2^r}^r}$$

due to Lemma 7.1. In view of (7.1) this implies (7.5) and the proof of Proposition 6.2 is completed. □

7.3. *Proof of Proposition 6.3.* The starting idea for proving Proposition 6.3 is to reduce the stopping problem (6.4) to suitably discretized random stopping times. The choice of the discretized randomized stopping times is suggested by the following lemma.

LEMMA 7.2. *For $\tau^r \in \mathcal{T}^r$, the construction*

$$\tau^r[j](\omega, u) := \min\{k/2^j \mid k \in \mathbb{N}, \tau^r(\omega, u) \leq k/2^j\} \wedge T$$

defines a sequence $(\tau^r[j])_{j \in \mathbb{N}}$ in \mathcal{T}^r satisfying the following properties.

(i) $\tau^r[j] \searrow \tau^r$ *pointwise, in particular it follows:*

$$\lim_{j \rightarrow \infty} Y_{\tau^r[j](\omega, u)}^r(\omega, u) = Y_{\tau^r(\omega, u)}^r(\omega, u)$$

for any $\omega \in \Omega$ and every $u \in [0, 1]$.

(ii) $\lim_{j \rightarrow \infty} F_{Y_{\tau^r[j]}^r}(x) = F_{Y_{\tau^r}^r}(x)$ *holds for any continuity point x of $F_{Y_{\tau^r}^r}$.*

(iii) *For any $x \in \mathbb{R}$ and every $j \in \mathbb{N}$, we have*

$$F_{Y_{\tau^r[j]}^r}(x) = \mathbb{E}[\widehat{Y}_{t_{1j}}^x K_{\tau^r}(\cdot, [0, t_{1j}])] + \sum_{k=2}^{\infty} \mathbb{E}[\widehat{Y}_{t_{kj}}^x K_{\tau^r}(\cdot,]t_{(k-1)j}, t_{kj}])],$$

where $t_{kj} := (k/2^j) \wedge T$ for $k \in \mathbb{N}$, and $\widehat{Y}_t^x := \mathbb{1}_{]-\infty, x]} \circ Y_t$ for $t \in [0, T]$.

PROOF. Statements (i) and (ii) are obvious, so it remains to show (iii). To this end, recall from Lemma 7.1

$$(7.6) \quad F_{Y_{\tau^r[j]}^r}(x) = \mathbb{E}[K_{\tau^r[j]}(\cdot, \{t \in [0, T] \mid Y_t \leq x\})] \quad \text{for } x \in \mathbb{R}.$$

Since $K_{\tau^r[j]}(\omega, \cdot)$ is a probability measure, we also have

$$\begin{aligned} & K_{\tau^r[j]}(\omega, \{t \in [0, T] \mid Y_t(\omega) \leq x\}) \\ &= K_{\tau^r[j]}(\omega, \{t \in [0, t_{1j}] \mid Y_t(\omega) \leq x\}) \\ (7.7) \quad &+ \sum_{k=2}^{\infty} K_{\tau^r[j]}(\omega, \{t \in]t_{(k-1)j}, t_{kj}] \mid Y_t(\omega) \leq x\}) \\ &= K_{\tau^r[j]}(\omega, \{t \in [0, t_{1j}] \mid \widehat{Y}_t^x(\omega) = 1\}) \\ &+ \sum_{k=2}^{\infty} K_{\tau^r[j]}(\omega, \{t \in]t_{(k-1)j}, t_{kj}] \mid \widehat{Y}_t^x(\omega) = 1\}) \end{aligned}$$

for every $\omega \in \Omega$. Then by definitions of $K_{\tau^r[j]}$ and K_{τ^r} ,

$$(7.8) \quad \begin{aligned} & K_{\tau^r[j]}(\omega, \{t \in]t_{(k-1)j}, t_{kj}] \mid \widehat{Y}_t^x(\omega) = 1\}) \\ &= P^U(\{\tau^r[j](\omega, \cdot) \in]t_{(k-1)j}, t_{kj}[, \widehat{Y}_{\tau^r[j](\omega, \cdot)}^x(\omega) = 1\}) \end{aligned}$$

$$\begin{aligned}
 &= \mathbb{P}^U(\{\tau^r[j](\omega, \cdot) = t_{kj}, \widehat{Y}_{t_{kj}}^x(\omega) = 1\}) \\
 &= \widehat{Y}_{t_{kj}}^x(\omega) \mathbb{P}^U(\{\tau^r[j](\omega, \cdot) = t_{kj}\}) \\
 &= \widehat{Y}_{t_{kj}}^x(\omega) \mathbb{P}^U(\{\tau^r(\omega, \cdot) \in]t_{(k-1)j}, t_{kj}]\}) \\
 &= \widehat{Y}_{t_{kj}}^x(\omega) K_{\tau^r}(\omega,]t_{(k-1)j}, t_{kj}])
 \end{aligned}$$

for $\omega \in \Omega$ and $k \in \mathbb{N}$ with $k \geq 2$. Analogously, we also obtain

$$(7.9) \quad K_{\tau^r[j]}(\omega, \{t \in [0, t_{1j}] | \widehat{Y}_t^x(\omega) = 1\}) = \widehat{Y}_{t_{1j}}^x(\omega) K_{\tau^r}(\omega, [0, t_{1j}]).$$

Then statement (iii) follows from (7.6) combining (7.7) with (7.8) and (7.9). The proof is completed. \square

We shall use the discretized randomized stopping times, as defined in Lemma 7.2, to show that we can restrict ourselves to discrete randomized stopping times in the stopping problem (6.4).

COROLLARY 7.3. *If (3.1) is fulfilled, then for any $\tau^r \in \mathcal{T}^r$, we have:*

- (i) $\lim_{j \rightarrow \infty} \mathbb{E}[\Phi^*(x_j + Y_{\tau^r[j]}^r) - x_j] = \mathbb{E}[\Phi^*(x + Y_{\tau^r}^r) - x]$ for any sequence $(x_j)_{j \in \mathbb{N}}$ in $\mathbb{R}^{\mathbb{N}}$ converging to some $x \in \mathbb{R}$;
- (ii) $\lim_{j \rightarrow \infty} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau^r[j]}^r) - x] = \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau^r}^r) - x]$.

PROOF. Let the mapping $h : \mathcal{T}^r \times \mathbb{R}$ be defined by $h(\tau^r, x) = \mathbb{E}[\Phi^*(x + Y_{\tau^r}^r) - x]$. For every $\tau^r \in \mathcal{T}^r$, the mapping $h(\tau^r, \cdot)$ is convex, and thus continuous. Recalling that $\sup_{t \geq 0} Y_t \in H^{\Phi^*}(\Omega, \mathcal{F}, \mathbb{P})$, a direct application of Lemma 7.2, (i), along with the dominated convergence theorem yields part (i). Using terminology from [34] (see also [36]), statement (i) implies that the sequence $(h(\tau^r[j], \cdot))_{j \in \mathbb{N}}$ of continuous mappings $h(\tau^r[j], \cdot)$ epi-converges to the continuous mapping $h(\tau^r, \cdot)$. Moreover, in view of (7.3) and (7.4), we may conclude

$$\lim_{j \rightarrow \infty} \inf_{x \in \mathbb{R}} h(\tau^r[j], x) = \inf_{x \in \mathbb{R}} h(\tau^r, x),$$

drawing on Theorem 7.31 in [34] (see also Satz B 2.18 in [36]). \square

The following result provides the remaining missing link to prove Proposition 6.3.

LEMMA 7.4. *Let (3.1) be fulfilled. Furthermore, let $\tau^r \in \mathcal{T}^r$, and let us for any $j \in \mathbb{N}$ denote by $\mathcal{T}[j]$ the set containing all nonrandomized stopping times from \mathcal{T} taking values in $\{(k/2^j) \wedge T | k \in \mathbb{N}\}$ with probability 1. If $(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ is atomless with countably generated \mathcal{F}_t for every $t > 0$, and if $Y_t \in H^{\Phi^*}$ for $t > 0$, then*

$$(7.10) \quad \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau^r[j]}^r) - x] \leq \sup_{\tau \in \mathcal{T}[j]} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau}) - x].$$

PROOF. Let $m_j := \min\{k \in \mathbb{N} | k/2^j \geq T\}$. If $m_j = 1$, then the statement of Lemma 7.4 is obvious. So let us assume $m_j \geq 2$. Set $t_{kj} := (k/2^j) \wedge T$ and let the mapping $h : \mathcal{T}^r \times \mathbb{R} \rightarrow \mathbb{R}$ be defined via $h(\tau^r, x) := \mathbb{E}[\Phi^*(x + Y_{\tau^r}) - x]$. We already know from Lemma 7.2 that

$$(7.11) \quad F_{Y_{\tau^r[j]}}(x) = \mathbb{E}[\widehat{Y}_{t_{1j}}^x K_{\tau^r}(\cdot, [0, t_{1j}])] + \sum_{k=2}^{m_j} \mathbb{E}[\widehat{Y}_{t_{kj}}^x K_{\tau^r}(\cdot,]t_{(k-1)j}, t_{kj}])]$$

holds for any $x \in \mathbb{R}$. Here, $\widehat{Y}_t^x := \mathbb{1}_{]-\infty, x]} \circ Y_t$ for $t \in [0, T]$. Next,

$$Z_k := \begin{cases} K_{\tau^r}(\cdot, [0, t_{1j}]), & k = 1, \\ K_{\tau^r}(\cdot,]t_{(k-1)j}, t_{kj}), & k \in \{2, \dots, m_j\} \end{cases}$$

defines a random variable on $(\Omega, \mathcal{F}_{t_{kj}}, P_{|\mathcal{F}_{t_{kj}}})$ which satisfies $0 \leq Z_k \leq 1$ P-a.s.

In addition, we may observe that $\sum_{k=1}^{m_j} Z_k = 1$ holds P-a.s. Since the probability spaces $(\Omega, \mathcal{F}_{t_k}, P_{|\mathcal{F}_{t_k}})$ ($k = 1, \dots, m_j$) are assumed to be atomless and countably generated, we may draw on Corollary C.4 (cf. Appendix C) along with Lemma C.1 (cf. Appendix C) and Proposition B.1 (cf. Appendix B) to find a sequence $((B_{1n}, \dots, B_{m_j n}))_{n \in \mathbb{N}}$ in $\times_{k=1}^{m_j} \mathcal{F}_{t_{kj}}$ such that $B_{1n}, \dots, B_{m_j n}$ is a partition of Ω for $n \in \mathbb{N}$, and

$$\lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{1}_{B_{kn}} \cdot g] = \mathbb{E}[Z_k \cdot g]$$

holds for $g \in L^1(\Omega, \mathcal{F}_{t_{kj}}, P_{|\mathcal{F}_{t_{kj}}})$ and $k \in \{1, \dots, m_j\}$. In particular, we have by (7.11)

$$F_{Y_{\tau^r[j]}}(x) = \lim_{n \rightarrow \infty} \sum_{k=1}^{m_j} \mathbb{E}[\widehat{Y}_{t_{kj}}^x \mathbb{1}_{B_{kn}}] \quad \text{for } x \in \mathbb{R}.$$

So by Fatou's lemma along with (7.1),

$$(7.12) \quad \begin{aligned} & h(\tau^r[j], x) \\ & \leq \liminf_{n \rightarrow \infty} \int_0^\infty \Phi^{*'}(x + z) \left(1 - \sum_{k=1}^{m_j} \mathbb{E}[\widehat{Y}_{t_{kj}}^z \mathbb{1}_{B_{kn}}] \right) dz + \Phi^*(x) - x \end{aligned}$$

for $x \in \mathbb{R}$. Here, $\Phi^{*'}$ denotes the right-sided derivative of Φ^* . Next, we can define a sequence $(\tau_n)_{n \in \mathbb{N}}$ of nonrandomized stopping times from $\mathcal{T}[j]$ via

$$\tau_n := \sum_{k=1}^{m_j} t_{kj} \mathbb{1}_{B_{kn}}.$$

The distribution function $F_{Y_{\tau_n}}$ of Y_{τ_n} satisfies

$$F_{Y_{\tau_n}}(x) = \sum_{k=1}^{m_j} \mathbb{E}[\widehat{Y}_{t_{kj}}^x \mathbb{1}_{B_{kn}}] \quad \text{for } x \in \mathbb{R}$$

so that by (7.1)

$$(7.13) \quad h(\tau_n, x) = \int_0^\infty \Phi^{*'}(x+z) \left(1 - \sum_{k=1}^{m_j} \mathbb{E}[\widehat{Y}_{t_{kj}}^z \mathbb{1}_{B_{kn}}] \right) dz + \Phi^*(x) - x$$

for $x \in \mathbb{R}$. The crucial point now is to show that:

$$(\star) \quad \mathbb{H} := \{h(\tau, \cdot)|_{I_\beta} | \tau \in \mathcal{T}[j]\} \quad \text{is equicontinuous,}$$

where I_β is the interval defined in (7.3). Note that $(h(\tau_n, \cdot)|_{I_\beta})_{n \in \mathbb{N}}$ is a sequence in \mathbb{H} , and that $\{h(\tau, x) | \tau \in \mathcal{T}[j]\}$ is bounded for every $x \in \mathbb{R}$. Thus, in view of (7.3) the statement (\star) together with Arzela–Ascoli theorem implies that we can find a subsequence $(h(\tau_{i(n)}, \cdot)|_{I_\beta})_{n \in \mathbb{N}}$ such that

$$\lim_{n \rightarrow \infty} \sup_{x \in I_\beta} |h(\tau_{i(n)}, x) - g(x)| = 0$$

for some continuous mapping $g : I_\beta \rightarrow \mathbb{R}$. Hence, we may conclude from (7.13) and (7.12)

$$(7.14) \quad g(x) = \liminf_{n \rightarrow \infty} h(\tau_{i(n)}, x) \geq h(\tau^r[j], x) \quad \text{for } x \in I_\beta.$$

For any $\varepsilon > 0$, we may find some $n_0 \in \mathbb{N}$ such that $\sup_{x \in I_\beta} |h(\tau_{i(n_0)}, x) - g(x)| < \varepsilon$, which implies by (7.14) together with (7.4):

$$\begin{aligned} \inf_{x \in \mathbb{R}} h(\tau_{i(n_0)}, x) &\stackrel{(7.4)}{=} \inf_{x \in I_\beta} h(\tau_{i(n_0)}, x) \stackrel{(7.14)}{\geq} \inf_{x \in I_\beta} h(\tau^r[j], x) - \varepsilon \\ &\geq \inf_{x \in \mathbb{R}} h(\tau^r[j], x) - \varepsilon \end{aligned}$$

and (7.10) is proved. Therefore, it remains to show the statement (\star) .

Proof of (\star) . First, observe that for $\tau \in \mathcal{T}[j]$ and real numbers $x < y$, the inequality $h(\tau, x) + x \leq h(\tau, y) + y$ holds. Hence,

$$\begin{aligned} &|h(\tau, x) - h(\tau, y)| \\ &\leq \mathbb{E}[\Phi^*(y + Y_\tau)] - \mathbb{E}[\Phi^*(x + Y_\tau)] + |x - y| \\ (7.15) \quad &= \sum_{k=1}^{m_j} \mathbb{E}[\mathbb{1}_{\{t_{kj}\}} \circ \tau \underbrace{(\Phi^*(y + Y_{t_{kj}}) - \Phi^*(x + Y_{t_{kj}}))}_{\geq 0}] + |x - y| \\ &\leq \sum_{k=1}^{m_j} \mathbb{E}[\Phi^*(y + Y_{t_{kj}}) - \Phi^*(x + Y_{t_{kj}})] + |x - y| \\ &\leq \sum_{k=1}^{m_j} |h(t_{kj}, x) - h(t_{kj}, y)| + (m_j + 1)|x - y|. \end{aligned}$$

By convexity, the mappings $h(t_{kj}, \cdot), k = 1, \dots, m_j$, are also locally Lipschitz continuous. Thus, in view of (7.15), it is easy to verify that \mathbb{H} is equicontinuous at every $x \in I_\beta$. This proves (\star) . \square

Now, we are ready to prove Proposition 6.3. By (6.1) we have

$$\begin{aligned} \sup_{\tau' \in \mathcal{T}'} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau'}^r) - x] &\geq \sup_{\tau \in \mathcal{T}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] \\ &= \sup_{\tau \in \mathcal{T}} \sup_{Q \in \mathcal{Q}_{\Phi,0}} \left(\mathbb{E}_Q[Y_\tau] - \mathbb{E} \left[\Phi \left(\frac{dP}{dQ} \right) \right] \right). \end{aligned}$$

Moreover, due to (ii) of Corollary 7.3 and Lemma 7.4 we conclude that for any $\tau' \in \mathcal{T}'$

$$\begin{aligned} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau'}^r) - x] &= \lim_{j \rightarrow \infty} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau'}^r[j]) - x] \\ &\leq \sup_{\tau \in \mathcal{T}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x]. \end{aligned}$$

Thus, Proposition 6.3 is proved. \square

7.4. *Proof of Theorem 3.1.* First, we get from Propositions 6.2 and 6.3

$$\begin{aligned} \inf_{x \in \mathbb{R}} \sup_{\tau' \in \mathcal{T}'} \mathbb{E}[\Phi^*(x + Y_{\tau'}^r) - x] &= \sup_{\tau' \in \mathcal{T}'} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_{\tau'}^r) - x] \\ &= \sup_{\tau \in \mathcal{T}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x]. \end{aligned}$$

Furthermore,

$$\begin{aligned} \inf_{x \in \mathbb{R}} \sup_{\tau' \in \mathcal{T}'} \mathbb{E}[\Phi^*(x + Y_{\tau'}^r) - x] &\geq \inf_{x \in \mathbb{R}} \sup_{\tau \in \mathcal{T}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] \\ &\geq \sup_{\tau \in \mathcal{T}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x]. \end{aligned}$$

Thus,

$$\sup_{\tau \in \mathcal{T}'} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] = \inf_{x \in \mathbb{R}} \sup_{\tau \in \mathcal{T}} \mathbb{E}[\Phi^*(x + Y_\tau) - x]$$

which completes the proof of Theorem 3.1. \square

7.5. *Proof of Proposition 3.5.* Just simplifying notation, we assume that $\mathbb{T} = \{0, 1, \dots, T\}$ with T being a positive integer. By (3.4) we have

$$\sup_{\tau \in \mathbb{T}} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] = \inf_{x \in \mathbb{R}} \sup_{\tau \in \mathbb{T}} \mathbb{E}[\Phi^*(x + Y_\tau) - x].$$

So it is left to show that there exists a solution τ^* of the maximization problem (3.5) and a solution x^* of the minimization problem (3.6). Indeed such a pair (τ^*, x^*) would be as required.

In view of (7.4), we may find some compact interval I of \mathbb{R} such that

$$(7.16) \quad \sup_{\tau \in \mathcal{T}_T} \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + Y_\tau) - x] = \sup_{\tau \in \mathcal{T}_T} \inf_{x \in I} \mathbb{E}[\Phi^*(x + Y_\tau) - x].$$

Let $\mathcal{C}(I)$ denote the space of continuous real-valued mappings on I . This space will be equipped with the sup-norm $\|\cdot\|_\infty$, whereas the product $\mathcal{C}(I)^T$ is viewed to be endowed with the norm $\|\cdot\|_{\infty, T}$, defined by $\|(f_1, \dots, f_T)\|_{\infty, T} := \sum_{i=1}^T \|f_i\|_\infty$. The key in solving the maximization problem (3.5) is to show that

$$(7.17) \quad K := \{(G_{1, A_1}, \dots, G_{T, A_T}) | (A_1, \dots, A_T) \in \mathcal{P}_T\}$$

is a weakly compact subset of $\mathcal{C}(I)^T$ w.r.t. the norm $\|\cdot\|_{\infty, T}$. Here, \mathcal{P}_T stands for the set of all (A_1, \dots, A_T) satisfying $A_t \in \mathcal{F}_t$ for $t \in \{1, \dots, T\}$ as well as $P(A_t \cap A_s) = 0$ for $t \neq s$, and $P(\bigcup_{i=1}^T A_i) = 1$. Furthermore, define

$$G_{t, A_t} : I \rightarrow \mathbb{R}, \quad x \mapsto \mathbb{E}[\mathbb{1}_{A_t} \cdot (\Phi^*(x + Y_t) - x)]$$

for $t \in \{1, \dots, T\}$, $A_t \in \mathcal{F}_t$.

Notice that any mapping G_{t, A_t} is convex and, therefore, also continuous.

Before proceeding, we need some further notation, namely \mathcal{P}_T^∞ denoting the set of all (f_1, \dots, f_T) satisfying $f_t \in L^\infty(\Omega, \mathcal{F}_t, P|_{\mathcal{F}_t})$ with $f_t \geq 0$ P-a.s. for $t \in \{1, \dots, T\}$, and $\sum_{i=1}^T f_i = 1$ P-a.s. Obviously, the subset $\{(\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_T}) | (A_1, \dots, A_T) \in \mathcal{P}_T\}$ consists of extreme points of \mathcal{P}_T^∞ . Any $f_t \in L^\infty(\Omega, \mathcal{F}_t, P|_{\mathcal{F}_t})$ may be associated with the mapping

$$H_{t, f_t} : I \rightarrow \mathbb{R}, \quad x \mapsto \mathbb{E}[f_t \cdot (\Phi^*(x + Y_t) - x)] \quad (t \in \{1, \dots, T\}).$$

It is convex, and thus also continuous. Hence, the mapping

$$\Lambda : \prod_{i=1}^T L^\infty(\Omega, \mathcal{F}_i, P|_{\mathcal{F}_i}) \rightarrow \mathcal{C}(I)^T, \quad (f_1, \dots, f_T) \mapsto (H_{1, f_1}, \dots, H_{T, f_T})$$

is well-defined, and obviously linear. In addition, it satisfies the following convenient continuity property.

LEMMA 7.5. *Let $\prod_{i=1}^T \sigma(L_i^\infty, L_i^1)$ be the product topology of $\sigma(L_i^\infty, L_i^1)$ ($i = 1, \dots, T$) on $\prod_{i=1}^T L^\infty(\Omega, \mathcal{F}_i, P|_{\mathcal{F}_i})$, where $\sigma(L_i^\infty, L_i^1)$ denotes the weak* topology on $L^\infty(\Omega, \mathcal{F}_i, P|_{\mathcal{F}_i})$.*

Then \mathcal{P}_T^∞ is compact w.r.t. $\prod_{i=1}^T \sigma(L_i^\infty, L_i^1)$, and the mapping Λ is continuous w.r.t. $\prod_{i=1}^T \sigma(L_i^\infty, L_i^1)$ and the weak topology induced by $\|\cdot\|_{\infty, T}$. In particular, the image $\Lambda(\mathcal{P}_T^\infty)$ of \mathcal{P}_T^∞ under Λ is weakly compact w.r.t. $\|\cdot\|_{\infty, T}$.

PROOF. The continuity of Λ follows in nearly the same way as in the proof of Proposition 3.1 from [21]. Moreover, \mathcal{P}_T^∞ is obviously closed w.r.t. the product topology $\prod_{i=1}^T \sigma(L_i^\infty, L_i^1)$, and even compact due to Banach–Alaoglu theorem.

Then by continuity of Λ , the set $\Lambda(\mathcal{P}_T^\infty)$ is weakly compact w.r.t. $\|\cdot\|_{\infty, T}$. This completes the proof. \square

We need some further preparation to utilize Lemma 7.5.

LEMMA 7.6. *Let $s, t \in \{1, \dots, T\}$ with $t \leq s$, and let $A \in \mathcal{F}_T$. If $(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ is atomless and if $\{\mathbb{E}[\mathbb{1}_A \cdot \Phi^*(x + Y_s)|\mathcal{F}_t]|x \in \mathbb{R}\}$ is a thin subset of $L^1(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$, then $\{\mathbb{E}[\mathbb{1}_A \cdot (\Phi^*(x + Y_s) - x)|\mathcal{F}_t]|x \in \mathbb{R}\}$ is a thin subset of $L^1(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$.*

PROOF. Let $A \in \mathcal{F}_t$ with $\mathbb{P}(A) > 0$. Since $(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ is atomless, we may find disjoint $B_1, B_2 \in \mathcal{F}_t$ contained in A with $\mathbb{P}(B_1), \mathbb{P}(B_2) > 0$. Then by assumption there exist nonzero $f_1, f_2 \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ such that f_i vanishes outside B_i as well as $\mathbb{E}[f_i \cdot \mathbb{E}[\mathbb{1}_A \cdot \Phi^*(x + Y_s)|\mathcal{F}_t]] = 0$ for $x \in \mathbb{R}$ and $i \in \{1, 2\}$.

Moreover, we may choose $\lambda_1, \lambda_2 \in \mathbb{R}$ with $\lambda_i \neq 0$ for at least one $i \in \{1, 2\}$ and $\mathbb{E}[(\lambda_1 f_1 + \lambda_2 f_2) \cdot \mathbb{1}_A] = 0$. Finally, $\lambda_1 f_1 + \lambda_2 f_2 \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t}) \setminus \{0\}$, and, setting $f := \lambda_1 f_1 + \lambda_2 f_2$,

$$\begin{aligned} & \mathbb{E}[f \cdot \mathbb{E}[\mathbb{1}_A \cdot (\Phi^*(x + Y_s) - x)|\mathcal{F}_t]] \\ &= \sum_{i=1}^2 \lambda_i \mathbb{E}[f_i \cdot \mathbb{E}[\mathbb{1}_A \cdot \Phi^*(x + Y_s)|\mathcal{F}_t]] - x \mathbb{E}[(\lambda_1 f_1 + \lambda_2 f_2) \cdot \mathbb{1}_A] = 0 \end{aligned}$$

for $x \in \mathbb{R}$. This completes the proof. \square

The missing link in concluding the desired compactness of the set K from (7.17) is provided by the following auxiliary result.

LEMMA 7.7. *Let $(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ be atomless for $t \in \{1, \dots, T\}$, and furthermore let the subset $\{\mathbb{E}[\mathbb{1}_A \cdot \Phi^*(x + Y_s)|\mathcal{F}_t]|x \in \mathbb{R}\}$ of $L^1(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ be thin for arbitrary $s, t \in \{1, \dots, T\}$ with $t \leq s$ and $A \in \mathcal{F}_T$.*

Then for any $(f_1, \dots, f_T) \in \mathcal{P}_T^\infty$, there exist $(A_1, \dots, A_T) \in \mathcal{P}_T$ and mappings $g_t \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$ ($t = 1, \dots, T$) such that $\Lambda(g_1, \dots, g_T) \equiv 0$, and

$$(f_1, \dots, f_T) = (\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_T}) + (g_1, \dots, g_T) \quad \text{P-a.s.}$$

PROOF. Let $s, t \in \{1, \dots, T\}$ with $t \leq s$ and $A \in \mathcal{F}_T$. We may draw on Lemma 7.6 to observe that $\{\mathbb{E}[\mathbb{1}_A \cdot (\Phi^*(x + Y_s) - x)|\mathcal{F}_t]|x \in \mathbb{R}\}$ is a thin subset of $L^1(\Omega, \mathcal{F}_t, \mathbb{P}|_{\mathcal{F}_t})$. Then the statement of Lemma 7.7 follows immediately from Proposition C.3 (cf. Appendix C) applied to the sets M_t ($t = 1, \dots, T$), where $M_t := \{\Phi^*(x + Y_t) - x|x \in \mathbb{R}\}$. \square

Under the assumptions of Lemma 7.7, the set K defined in (7.17) coincides with $\Lambda(\mathcal{P}_T^\infty)$, which in turn is weakly compact w.r.t. $\|\cdot\|_{\infty, T}$ due to Lemma 7.5.

COROLLARY 7.8. *Under the assumptions of Lemma 7.7, the set K [cf. (7.17)] is weakly compact w.r.t. $\|\cdot\|_{\infty, T}$.*

Now we are ready to select a solution of the maximization problem (3.5).

Existence of a solution of maximization problem (3.5): Let the assumptions of Proposition 3.5 be fulfilled. In view of (7.16), it suffices to solve

$$\text{maximize} \quad \inf_{x \in I} \mathbb{E}[\Phi^*(x + Y_\tau) - x] \quad \text{over} \quad \tau \in \mathcal{T}_{\mathbb{T}}.$$

Let us assume that $\sup_{\tau \in \mathcal{T}_{\mathbb{T}}} \inf_{x \in I} \mathbb{E}[\Phi^*(x + Y_\tau) - x] > \inf_{x \in I} \mathbb{E}[\Phi^*(x + Y_0) - x]$ because otherwise $\tau \equiv 0$ would be optimal. Since $P(A) \in \{0, 1\}$ for $A \in \mathcal{F}_0$ by assumption, any stopping time $\tau \in \mathbb{T} \setminus \{0\}$ is concentrated on $\{1, \dots, T\}$.

By Corollary 7.8, the set K [cf. (7.17)] is weakly compact w.r.t. the norm $\|\cdot\|_{\infty, T}$. Furthermore, the concave mapping $L : \mathcal{C}(I)^T \rightarrow \mathbb{R}$, defined by $L(r_1, \dots, r_T) := \inf_{x \in I} \sum_{t=1}^T r_t(x)$, is continuous w.r.t. $\|\cdot\|_{\infty, T}$. This means that $-L$ is convex as well as $\|\cdot\|_{\infty, T}$ -continuous, and thus also weakly lower semicontinuous because $\|\cdot\|_{\infty, T}$ -closed convex subsets are also weakly closed. Hence, L is weakly upper semicontinuous and, therefore, its restriction to K attains a maximum. In particular, the set

$$\left\{ \inf_{x \in I} \mathbb{E}[\Phi^*(x + Y_\tau) - x] \mid \tau \in \mathcal{T}_{\mathbb{T}} \setminus \{0\} \right\} = L(K)$$

has a maximum. This shows that we may find a solution of (3.5).

Existence of a solution of problem (3.6): By $l(x) := \sup_{\tau \in \mathcal{T}_{\mathbb{T}}} \mathbb{E}[\Phi^*(x + Y_\tau) - x]$, we may define a convex and, therefore, also continuous mapping $l : \mathbb{R} \rightarrow \mathbb{R}$. Moreover, by Lemma A.1 (cf. Appendix A),

$$\lim_{x \rightarrow -\infty} l(x) \geq \lim_{x \rightarrow -\infty} (\Phi^*(x) - x) = \infty = \lim_{x \rightarrow \infty} (\Phi^*(x) - x) \leq \lim_{x \rightarrow \infty} l(x).$$

This means that $\inf_{x \in \mathbb{R}} l(x) = \inf_{x \in [-\varepsilon, \varepsilon]} l(x)$ for some $\varepsilon > 0$. Hence, l attains its minimum at some $x^* \in [-\varepsilon, \varepsilon]$ because l is continuous. Any such x^* is a solution of the problem (3.6).

APPENDIX A

LEMMA A.1. *Let $\Phi : [0, \infty[\rightarrow [0, \infty]$ be a lower semicontinuous, convex mapping satisfying $\inf_{x \geq 0} \Phi(x) = 0$, and $\lim_{x \rightarrow \infty} \frac{\Phi(x)}{x} = \infty$. Furthermore, let $\mathcal{Q}_{\Phi, 0}$ denotes the set of all probability measures Q on \mathcal{F} which are absolutely continuous w.r.t. P such that the Radon–Nikodym derivative $\frac{dQ}{dP}$ satisfies $\mathbb{E}[\Phi(\frac{dQ}{dP})] < \infty$. Then the following statements hold true.*

(i) *If $\Phi(x_0) < \infty$ for some $x_0 > 0$, then the Fenchel–Legendre transform $\Phi^* : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$ of Φ is a nondecreasing, convex finite mapping. In particular, its restriction $\Phi^*_{|[0, \infty[}$ to $[0, \infty[$ is a finite Young-function, which in addition satisfies the condition $\lim_{x \rightarrow \infty} (\Phi^*(x) - x) = \infty$ if $x_0 > 1$, and $\lim_{x \rightarrow -\infty} (\Phi^*(x) - x) = \infty$ in the case of $x_0 < 1$.*

(ii) If $\Phi(x_0), \Phi(x_1) < \infty$ for some $x_0 < 1 < x_1$, then for any X from $H^{\Phi^*} := H^{\Phi^*}_{[0, \infty[}$, we obtain

$$\sup_{Q \in \mathcal{Q}_{\Phi, 0}} \left(\mathbb{E}_Q[X] - \mathbb{E} \left[\Phi \left(\frac{dQ}{dP} \right) \right] \right) = \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x + X) - x],$$

where the supremum on the left-hand side of the equality is attained for some $Q \in \mathcal{Q}_{\Phi, 0}$.

PROOF. Obviously, Φ^* is a nondecreasing convex function satisfying the properties

$$(A.1) \quad \begin{aligned} \Phi^*(0) &= - \inf_{x \geq 0} \Phi(x) = 0 \quad \text{and} \\ \lim_{y \rightarrow \infty} \Phi^*(y) &\geq \lim_{y \rightarrow \infty} (x_0 y - \Phi(x_0)) = \infty. \end{aligned}$$

Next, we want to verify the finiteness of Φ^* . Since Φ^* is nondecreasing, and $\Phi^*(y) \geq x_0 y - \Phi(x_0) > -\infty$ holds for any $y \in \mathbb{R}$, it suffices to show that $\Phi^*(y) < \infty$ for every $y \geq 0$. For that purpose, consider the mapping

$$\beta : [0, \infty[\times [0, \infty[\rightarrow [-\infty, \infty[, \quad (y, x) \mapsto xy - \Phi(x).$$

By assumption on Φ , we have

$$\lim_{x \rightarrow \infty} \beta(y, x) = \lim_{x \rightarrow \infty} x \left(y - \frac{\Phi(x)}{x} \right) = -\infty < \beta(y, x_0) \quad \text{for } y \geq 0.$$

Hence for any $y \geq 0$, we may find some $z_y \in [x_0, \infty[$ such that we obtain $\Phi^*(y) = \sup_{0 \leq x \leq z_y} \beta(y, x)$. Moreover, Φ is nonnegative, $\beta(y, \cdot)$ is upper semi-continuous for $y \geq 0$. Hence, for every $y \geq 0$, there is some $x \in [0, z_y]$ with $\Phi^*(y) = \sup_{0 \leq x \leq z_y} \beta(y, x) = \beta(y, x) < \infty$.

As a finite convex function Φ^* is continuous. Since it is also nondecreasing, we may conclude from (A.1) that its restriction to $[0, \infty[$ is a finite Young function. Let us now assume that $x_0 > 1$. Then

$$\lim_{y \rightarrow \infty} (\Phi^*(y) - y) \geq \lim_{y \rightarrow \infty} ((x_0 - 1)y - \Phi(x_0)) = \infty.$$

Analogously, $\lim_{y \rightarrow -\infty} (\Phi^*(y) - y) = \infty$ may be derived in the case of $x_0 < 1$. Thus, we have proved the full statement (i).

Let us turn over to the proof of statement (ii), and let us consider the mapping

$$\rho : H^{\Phi^*} \rightarrow [-\infty, \infty[, \quad X \mapsto \inf_{x \in \mathbb{R}} \mathbb{E}[\Phi^*(x - X) - x].$$

Then, due to convexity of Φ^* , we may apply Jensen's inequality along with statement (i) to conclude

$$\lim_{x \rightarrow -\infty} \mathbb{E}[\Phi^*(x - X) - x] \geq \lim_{x \rightarrow -\infty} (\mathbb{E}[\Phi^*(x - \mathbb{E}[X])] - x) = \infty \quad \text{for } X \in H^{\Phi^*},$$

and

$$\lim_{x \rightarrow \infty} \mathbb{E}[\Phi^*(x - X) - x] \geq \lim_{x \rightarrow \infty} [\Phi^*(x - \mathbb{E}[X]) - x] = \infty \quad \text{for } X \in H^{\Phi^*}.$$

Thus, for any $X \in H^{\Phi^*}$, we find some $\delta_X > 0$ such that

$$\rho(X) = \inf_{x \in [-\delta_X, \delta_X]} \mathbb{E}[\Phi^*(x - X) - x].$$

In addition, for $X \in H^{\Phi^*}$, the mapping $x \mapsto \mathbb{E}[\Phi^*(x - X) - x]$ is a convex mapping on \mathbb{R} , hence its restriction to $[-\delta_X, \delta_X]$ is continuous. This implies that ρ is a real-valued function.

Moreover, it is easy to check that ρ is a so called convex risk measure, defined to mean that it satisfies the following properties.

- monotonicity: $\rho(X) \geq \rho(Y)$ for all $X, Y \in H^{\Phi^*}$ with $X \leq Y$,
- cash-invariance: $\rho(X + m) = \rho(X) - m$ for all $X \in H^{\Phi^*}$ and $m \in \mathbb{R}$,
- convexity: $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$ for all $X, Y \in H^{\Phi^*}$, $\lambda \in [0, 1]$.

Then we obtain from Theorem 4.3 in [16] that

$$\rho(X) = \max_{Q \in \mathcal{Q}_{\Phi,0}} (\mathbb{E}_Q[-X] - \rho^*(Q))$$

holds for all $X \in H^{\Phi^*}$, where

$$\rho^*(Q) := \sup_{X \in H^{\Phi^*}} (\mathbb{E}_Q[-X] - \rho(X)).$$

By routine procedures, we may verify

$$\rho^*(Q) = \sup_{X \in H^{\Phi^*}} (\mathbb{E}_Q[X] - \mathbb{E}[\Phi^*(X)])$$

for $Q \in \mathcal{Q}_{\Phi,0}$. Since $\lim_{x \rightarrow -\infty} [\Phi^*(x) - x] = \lim_{x \rightarrow \infty} [\Phi^*(x) - x] = \infty$ due to statement (i), we may conclude from [16]

$$\rho^*(Q) = \sup_{X \in H^{\Phi^*}} (\mathbb{E}_Q[X] - \mathbb{E}[\Phi^*(X)]) = \mathbb{E} \left[\Phi \left(\frac{dQ}{dP} \right) \right] \quad \text{for all } Q \in \mathcal{Q}_{\Phi,0}.$$

This completes the proof. \square

APPENDIX B

Let $(\bar{\Omega}, \bar{\mathcal{F}}, (\bar{\mathcal{F}})_{i \in \{1, \dots, m\}}, \bar{\mathbb{P}})$ be a filtered probability space, and let the product space $\times_{i=1}^m L^\infty(\bar{\Omega}, \bar{\mathcal{F}}_i, \bar{\mathbb{P}}|_{\bar{\mathcal{F}}_i})$ be endowed with the product topology $\times_{i=1}^m \sigma(L_i^\infty, L_i^1)$ of the weak* topologies $\sigma(L_i^\infty, L_i^1)$ on $L^\infty(\bar{\Omega}, \bar{\mathcal{F}}_i, \bar{\mathbb{P}}|_{\bar{\mathcal{F}}_i})$ (for $i = 1, \dots, m$).

PROPOSITION B.1. *Let $L^1(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ be separable w.r.t. the weak topology $\sigma(L_i^1, L_i^\infty)$ for $i \in \{1, \dots, m\}$, and let $\mathcal{A} \subseteq \times_{i=1}^m L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ be relatively compact w.r.t. $\times_{i=1}^m \sigma(L_i^\infty, L_i^1)$.*

Then for any X from the $\times_{i=1}^m \sigma(L_i^\infty, L_i^1)$ -closure of \mathcal{A} , we may find a sequence $(X_n)_{n \in \mathbb{N}}$ in \mathcal{A} which converges to X w.r.t. the $\times_{i=1}^m \sigma(L_i^\infty, L_i^1)$.

PROOF. Setting $E := \times_{i=1}^m L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$, we shall denote by E' the topological dual of E w.r.t. $\times_{i=1}^m \sigma(L_i^\infty, L_i^1)$. It is easy to check that

$$\Gamma(g_1, \dots, g_m)(f_1, \dots, f_m) := \sum_{i=1}^m \mathbb{E}[f_i \cdot g_i],$$

where $g_i \in L^1(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ and $f_i \in L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ (for $i = 1, \dots, m$) defines a linear operator from $\times_{i=1}^m L^1(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ onto E' which is continuous w.r.t. the product topology $\times_{i=1}^m \sigma(L_i^1, L_i^\infty)$ of the weak topologies $\sigma(L_1^1, L_1^\infty), \dots, \sigma(L_m^1, L_m^\infty)$ and the weak topology $\sigma(E', E)$.

Since $\times_{i=1}^m \sigma(L_i^1, L_i^\infty)$ is separable by assumption, we may conclude that $\sigma(E', E)$ is separable, too. Then the statement of the Proposition B.1 follows immediately from [23], page 30. \square

APPENDIX C

Let for $m \in \mathbb{N}$ denote by $(\overline{\Omega}, \overline{\mathcal{F}}, (\overline{\mathcal{F}}_i)_{i \in \{1, \dots, m\}}, \overline{\mathbb{P}})$ a filtered probability space, and let the set $\overline{\mathcal{P}}_m$ gather all sets (A_1, \dots, A_m) from $\times_{i=1}^m \overline{\mathcal{F}}_i$ satisfying $\overline{\mathbb{P}}(A_i \cap A_j) = 0$ for $i \neq j$ and $\overline{\mathbb{P}}(\bigcup_{i=1}^m A_i) = 1$. We shall endow, respectively, the product spaces $\times_{i=k}^m L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ with the product topologies $\times_{i=k}^m \sigma(L_i^\infty, L_i^1)$ of the weak* topologies $\sigma(L_i^\infty, L_i^1)$ on $L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ (for $k \in \{1, \dots, m\}$ and $i = k, \dots, m$). Fixing $k \in \{1, \dots, m\}$ and nonnegative $h \in L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_k, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_k})$, the subset $\overline{\mathcal{P}}_{mk}^\infty(h) \subseteq \times_{i=k}^m L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ is defined to consist of all $(f_k, \dots, f_m) \in \times_{i=k}^m L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ such that $f_i \geq 0$ $\overline{\mathbb{P}}$ -a.s. for any $i \in \{k, \dots, m\}$ and $\sum_{i=k}^m f_i = h$ $\overline{\mathbb{P}}$ -a.s. For abbreviation, we shall use notation $\overline{\mathcal{P}}_m^\infty := \overline{\mathcal{P}}_{m1}^\infty(1)$.

LEMMA C.1. *$\overline{\mathcal{P}}_{mk}^\infty(h)$ is a compact subset of $\times_{i=k}^m L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ w.r.t. the topology $\times_{i=k}^m \sigma(L_i^\infty, L_i^1)$ for $k \in \{1, \dots, m\}$ and arbitrary nonnegative $h \in L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_k, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_k})$.*

PROOF. The statement of Lemma C.1 is obvious in view of the Banach–Alaoglu theorem. \square

PROPOSITION C.2. Let $M_i \subseteq L^1(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ be nonvoid for $i = 1, \dots, m$ such that $\{\mathbb{E}[\mathbb{1}_A \cdot f | \overline{\mathcal{F}}_i] | f \in M_j\}$ is a thin subset of $L^1(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ for $i, j \in \{1, \dots, m\}$ with $i \leq j$ and any $A \in \overline{\mathcal{F}}_m$. Furthermore, let us fix $(f_1, \dots, f_m) \in \overline{\mathcal{P}}_m^\infty$ and consider the set N_1 consisting of all (h_1, \dots, h_m) from $\times_{i=1}^m L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ satisfying $\mathbb{E}[h_i \cdot \varphi_i] = \mathbb{E}[f_i \cdot \varphi_i]$ for any $\varphi_i \in M_i, i = 1, \dots, m$. Then the set $N_1 \cap \overline{\mathcal{P}}_m^\infty$ has extreme points, and for each extreme point (h_1^*, \dots, h_m^*) , there exists some $(A_1, \dots, A_m) \in \overline{\mathcal{P}}_m$ such that $h_i^* = \mathbb{1}_{A_i}$ $\overline{\mathbb{P}}$ -a.s. holds for $i = 1, \dots, m$.

PROOF. We shall use ideas from the proof of Proposition 6 in [31].

First, let us, for any $k \in \{1, \dots, m\}$, denote by N_k the set of all (h_k, \dots, h_m) from $\times_{i=k}^m L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ satisfying $\mathbb{E}[h_i \cdot \varphi_i] = \mathbb{E}[f_i \cdot \varphi_i]$ for $\varphi_i \in M_i$ and $i = k, \dots, m$. It is closed w.r.t. $\times_{i=k}^m \sigma(L_i^\infty, L_i^1)$. Hence, by Lemma C.1, the set $K_k(h) := N_k \cap \overline{\mathcal{P}}_{mk}^\infty(h)$ is compact w.r.t. $\times_{i=k}^m \sigma(L_i^\infty, L_i^1)$ for every nonnegative $h \in L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_k, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_k})$. Since it is also convex, we may use the Krein–Milman theorem to conclude that each set $K_k(h)$ has some extreme point if it is nonvoid. Notice that $K_1(1)$ contains at least (f_1, \dots, f_m) so that it has some extreme point. We shall now show by backward induction that for any $k \in \{1, \dots, m\}$ and any nonnegative $h \in L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_k, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_k})$ with nonvoid $K_k(h)$

- (★★) each of its extreme points (h_k^*, \dots, h_m^*) satisfies $h_i^* = h \cdot \mathbb{1}_{A_i}$ $\overline{\mathbb{P}}$ -a.s. ($i = k, \dots, m$) for some $(A_1, \dots, A_m) \in \overline{\mathcal{P}}_m$ with $A_i = \emptyset$ if $i < k$.

Obviously, this would imply the statement of Proposition C.2.

For $k = m$, the set $K_m(h)$ is nonvoid iff $\mathbb{E}[h \cdot \varphi_m] = \mathbb{E}[f_m \cdot \varphi_m]$ holds for every $\varphi_m \in M_m$. In this case, h is the only extreme point, which has trivial representation $h = h \cdot \mathbb{1}_\Omega$ corresponding to $(\emptyset, \dots, \emptyset, \Omega) \in \overline{\mathcal{P}}_m$.

Now let us assume that for some $k \in \{2, \dots, m\}$ and every nonvoid $K_k(h)$ statement (★★) is satisfied. Let $h \in L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_{k-1}, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_{k-1}})$ be nonnegative with $K_{k-1}(h) \neq \emptyset$, and select any extreme point $(h_{k-1}^*, \dots, h_m^*)$ of $K_{k-1}(h)$. Then $h - h_{k-1}^*$ belongs to $L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_{k-1}, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_{k-1}})$ and is nonnegative. Moreover, $(h_k^*, \dots, h_m^*) \in K_k(h - h_{k-1}^*)$, and it is easy to check that (h_k^*, \dots, h_m^*) is even an extreme point of $K_k(h - h_{k-1}^*)$. Hence, by assumption, there exists some $(A_1, \dots, A_m) \in \overline{\mathcal{P}}_m$ satisfying $A_i = \emptyset$ if $i \leq k - 1$ and $h_i^* = (h - h_{k-1}^*) \cdot \mathbb{1}_{A_i}$ $\overline{\mathbb{P}}$ -a.s. for $i = k, \dots, m$.

Setting $D := \{h_{k-1}^* > 0\} \cap \{h - h_{k-1}^* > 0\}$, we want to show $\overline{\mathbb{P}}(D) = 0$. This will be done by contradiction assuming $\overline{\mathbb{P}}(D) > 0$. Then $\overline{\mathbb{P}}(D_\varepsilon) > 0$ for some $\varepsilon > 0$, where $D_\varepsilon := \{h_{k-1}^* > \varepsilon\} \cap \{h - h_{k-1}^* > \varepsilon\}$.

We may observe by assumption that $\{\mathbb{E}[\mathbb{1}_{A_i} \cdot \varphi_i | \overline{\mathcal{F}}_{k-1}] | \varphi_i \in M_i\}$ (with $i = k, \dots, m$) as well as M_{k-1} are all thin subsets of $L^1(\overline{\Omega}, \overline{\mathcal{F}}_{k-1}, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_{k-1}})$. Since finite unions of thin subsets are thin subsets again (cf. [1], Proposition 2.1), we may find some nonzero $g \in L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_{k-1}, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_{k-1}})$ vanishing outside D_ε , and satisfying

$\mathbb{E}[g \cdot \varphi_{k-1}] = 0$ for $\varphi_{k-1} \in M_{k-1}$ as well as

$$\mathbb{E}[g \cdot \mathbb{1}_{A_i} \cdot \varphi_i] = \mathbb{E}[g \cdot \mathbb{E}[\mathbb{1}_{A_i} \cdot \varphi_i | \overline{\mathcal{F}}_{k-1}]] = 0 \quad (\varphi_i \in M_i, i \in \{k, \dots, m\}).$$

According to Theorem 2.4 in [1], we may choose g such that

$$\overline{\mathbb{P}}(\{|g| = 1\} \cap D_\varepsilon) = \overline{\mathbb{P}}(D_\varepsilon)$$

holds. Now, define $(\widehat{h}_{k-1}, \dots, \widehat{h}_m)$ and $(\overline{h}_{k-1}, \dots, \overline{h}_m)$ by

$$\widehat{h}_i := \begin{cases} h_i^* + \varepsilon g, & i = k - 1, \\ h_i^* - \varepsilon g \mathbb{1}_{A_i}, & \text{otherwise} \end{cases} \quad \text{and} \quad \overline{h}_i := \begin{cases} h_i^* - \varepsilon g, & i = k - 1, \\ h_i^* + \varepsilon g \mathbb{1}_{A_i}, & \text{otherwise.} \end{cases}$$

Since $\overline{\mathbb{P}}(A_i \cap A_j) = 0$ for $i \neq j$ and $\overline{\mathbb{P}}(\bigcup_{i=k}^m A_i) = 1$, we obtain $\sum_{i=k}^m g \cdot \mathbb{1}_{A_i} = g$ $\overline{\mathbb{P}}$ -a.s. So by construction, $(\widehat{h}_{k-1}, \dots, \widehat{h}_m)$, $(\overline{h}_{k-1}, \dots, \overline{h}_m)$ differ, and belong both to $K_{k-1}(h)$. Moreover, $h_i^* = \widehat{h}_i/2 + \overline{h}_i/2$ for $i = k - 1, \dots, m$. This contradicts the fact that $(h_{k-1}^*, \dots, h_m^*)$ is an extreme point of $K_{k-1}(h)$. Therefore, $\overline{\mathbb{P}}(D) = 0$.

Now define $(B_1, \dots, B_m) \in \times_{i=1}^m \overline{\mathcal{F}}_i$ by

$$B_i := \begin{cases} \{h_{k-1}^* > 0, h = h_{k-1}^*\}, & i = k - 1, \\ A_i \cap \{h_{k-1}^* = 0\}, & i \in \{k, \dots, m\}, \\ \emptyset, & \text{otherwise.} \end{cases}$$

Obviously, $\overline{\mathbb{P}}(B_i \cap B_j) = 0$ for $i \neq j$ follows from $\overline{\mathbb{P}}(A_i \cap A_j) = 0$ for $i \neq j$. Moreover, $\overline{\mathbb{P}}(\bigcup_{i=1}^m B_i) \geq \overline{\mathbb{P}}(\Omega \setminus D \cap \bigcup_{i=k}^m A_i) = 1$. In particular, $(B_1, \dots, B_m) \in \overline{\mathcal{P}}_m$. Finally, it may be verified easily that $h_i^* = h \cdot \mathbb{1}_{B_i}$ $\overline{\mathbb{P}}$ -a.s. holds for $i = k - 1, \dots, m$. Hence, $K_{k-1}(h)$ fulfills statement $(\star\star)$ completing the proof. \square

PROPOSITION C.3. *Let $M_i \subseteq L^1(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ be nonvoid for $i = 1, \dots, m$ such that $\{\mathbb{E}[\mathbb{1}_A \cdot f | \overline{\mathcal{F}}_i] | f \in M_j\}$ is a thin subset of $L^1(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ for $i, j \in \{1, \dots, m\}$ with $i \leq j$ and any $A \in \overline{\mathcal{F}}_m$.*

Then for any $(f_1, \dots, f_m) \in \overline{\mathcal{P}}_m^\infty$, there exist $(A_1, \dots, A_m) \in \overline{\mathcal{P}}_m$ and $g_i \in L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ ($i = 1, \dots, m$) such that

$$\mathbb{E}[g_i \cdot \varphi_i] = 0 \quad \text{for } \varphi_i \in M_i \text{ with } i = 1, \dots, m,$$

and

$$(f_1, \dots, f_m) = (\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_m}) + (g_1, \dots, g_m) \quad \overline{\mathbb{P}}\text{-a.s.}$$

PROOF. Let us fix any $(f_1, \dots, f_T) \in \overline{\mathcal{P}}_m^\infty$, and let N_1 denote the set consisting of all (h_1, \dots, h_m) , where $h_i \in L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ such that $\mathbb{E}[h_i \cdot \varphi_i] = \mathbb{E}[f_i \cdot \varphi_i]$ for $\varphi_i \in M_i$. By Proposition C.2, we may select an extreme point (h_1, \dots, h_m) of $N_1 \cap \overline{\mathcal{P}}_m^\infty$ and some $(A_1, \dots, A_m) \in \overline{\mathcal{P}}_m$ such that $h_i = \mathbb{1}_{A_i}$ $\overline{\mathbb{P}}$ -a.s. holds for $i = 1, \dots, m$. Then $(g_1, \dots, g_m) := (f_1 - h_1, \dots, f_m - h_m)$ and (A_1, \dots, A_m) are as required. \square

COROLLARY C.4. *If $(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ is atomless for every $i \in \{1, \dots, m\}$, then $\overline{\mathcal{P}}_m^\infty$ is the $\times_{i=1}^m \sigma(L_i^\infty, L_i^1)$ -closure of*

$$\{(\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_m}) | (A_1, \dots, A_m) \in \overline{\mathcal{P}}_m\}.$$

PROOF. Let $(f_1, \dots, f_m) \in \overline{\mathcal{P}}_m^\infty$ be arbitrary. Consider the subsets

$$\begin{aligned} U_{i\varepsilon}(M_i) &:= \{\varphi \in L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i}) | |\mathbb{E}[(f_i - \varphi) \cdot f]| < \varepsilon \text{ for } f \in M_i\} \\ &= \{f_i\} + \{\phi \in L^\infty(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i}) | |\mathbb{E}[\phi \cdot f]| < \varepsilon \text{ for } f \in M_i\}, \end{aligned}$$

where $\varepsilon > 0$, and M_i any nonvoid, finite subset of $L^1(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$. The sets $\times_{i=1}^m U_{i\varepsilon}(M_i)$ constitute a basis of the $\times_{i=1}^m \sigma(L_i^\infty, L_i^1)$ -neighbourhoods of (f_1, \dots, f_m) . So let us select any $\varepsilon > 0$ and nonvoid finite subsets M_i of $L^1(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ for $i = 1, \dots, m$.

Let $i, j \in \{1, \dots, m\}$ with $i \leq j$, and $A \in \overline{\mathcal{F}}_m$. Then the set consisting of all $\mathbb{E}[\mathbb{1}_A \cdot f | \overline{\mathcal{F}}_i]$ with $f \in M_j$ is a nonvoid finite subset of $L^1(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$, in particular it is thin because $(\overline{\Omega}, \overline{\mathcal{F}}_i, \overline{\mathbb{P}}|_{\overline{\mathcal{F}}_i})$ is assumed to be atomless (cf. [29], Lemma 2). Hence, we may apply Proposition C.3 to select some $(A_1, \dots, A_m) \in \overline{\mathcal{P}}_m$ satisfying $\mathbb{E}[(f_i - \mathbb{1}_{A_i}) \cdot f] = 0$ for $f \in M_i$ and $i \in \{1, \dots, m\}$. This means

$$(\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_m}) \in \times_{i=1}^m U_{i\varepsilon}(M_i),$$

and completes the proof. \square

Acknowledgements. The article was prepared within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

The authors would like to thank Alexander Schied and Mikhail Urusov for fruitful discussions and helpful remarks.

REFERENCES

- [1] ANANTHARAMAN, R. (2012). Thin subspaces of $L^1(\lambda)$. *Quaest. Math.* **35** 133–143. [MR2945717](#)
- [2] ANDERSEN, L. and BROADIE, M. (2004). A primal-dual simulation algorithm for pricing multidimensional American options. *Management Sciences* **50** 1222–1234.
- [3] ANKIRCHNER, S. and STRACK, P. (2011). Skorokhod embeddings in bounded time. *Stoch. Dyn.* **11** 215–226.
- [4] BAXTER, J. R. and CHACON, R. V. (1977). Compactness of stopping times. *Z. Wahrsch. Verw. Gebiete* **40** 169–181.
- [5] BAYRAKTAR, E., KARATZAS, I. and YAO, S. (2010). Optimal stopping for dynamic convex risk measures. *Illinois J. Math.* **54** 1025–1067. [MR2928345](#)
- [6] BAYRAKTAR, E. and YAO, S. (2011). Optimal stopping for non-linear expectations. *Stochastic Process. Appl.* **121** 185–211. [MR2746173](#)

- [7] BAYRAKTAR, E. and YAO, S. (2011). Optimal stopping for non-linear expectations. *Stochastic Process. Appl.* **121** 212–264.
- [8] BELOMESTNY, D. (2013). Solving optimal stopping problems by empirical dual optimization. *Ann. Appl. Probab.* **23** 1988–2019.
- [9] BEN-TAL, A. and TEBOULLE, M. (1987). Penalty functions and duality in stochastic programming via ϕ -divergence functionals. *Math. Oper. Res.* **12** 224–240. [MR0888972](#)
- [10] BEN-TAL, A. and TEBOULLE, M. (2007). An old-new concept of convex risk measures: The optimized certainty equivalent. *Math. Finance* **17** 449–476. [MR2332265](#)
- [11] BION-NADAL, J. (2008). Dynamic risk measures: Time consistency and risk measures from BMO martingales. *Finance Stoch.* **12** 219–244. [MR2390189](#)
- [12] BORWEIN, J. M. and ZHUANG, D. (1986). On Fan’s minimax theorem. *Math. Program.* **34** 232–234.
- [13] CHENG, X. and RIEDEL, F. (2013). Optimal stopping under ambiguity in continuous time. *Math. Financ. Econ.* **7** 29–68. [MR3023890](#)
- [14] CHERIDITO, P., DELBAEN, F. and KUPPER, M. (2004). Coherent and convex monetary risk measures for bounded càdlàg processes. *Stochastic Process. Appl.* **112** 1–22. [MR2062565](#)
- [15] CHERIDITO, P., DELBAEN, F. and KUPPER, M. (2006). Dynamic monetary risk measures for bounded discrete-time processes. *Electron. J. Probab.* **11** 57–106. [MR2199055](#)
- [16] CHERIDITO, P. and LI, T. (2009). Risk measures on Orlicz hearts. *Math. Finance* **19** 189–214. [MR2509268](#)
- [17] DELBAEN, F., PENG, S. and ROSAZZA GIANIN, E. (2010). Representation of the penalty term of dynamic concave utilities. *Finance Stoch.* **14** 449–472. [MR2670421](#)
- [18] DETLEFSEN, K. and SCANDOLO, G. (2005). Conditional and dynamic convex risk measures. *Finance & Stochastics* **9** 539–561.
- [19] EDGAR, G. A., MILLET, A. and SUCHESTON, L. (1981). *On Compactness and Optimality of Stopping Times. Lecture Notes in Math.* **939** 36–61. Springer, Berlin.
- [20] EDGAR, G. A. and SUCHESTON, L. (1992). *Stopping Times and Directed Processes*. Cambridge Univ. Press, Cambridge.
- [21] EDWARDS, D. A. (1987). On a theorem of Dvoretzky, Wald and Wolfowitz concerning Liapounov measures. *Glasg. Math. J.* **29** 205–220. [MR0901667](#)
- [22] FAN, K. (1953). Minimax theorems. *Proc. Natl. Acad. Sci. USA* **39** 42–47.
- [23] FLORET, K. (1980). *Weakly Compact Sets. Lecture Notes in Math.* **801**. Springer, Berlin. [MR0576235](#)
- [24] FÖLLMER, H. and PENNER, I. (2006). Convex risk measures and the dynamics of their penalty functions. *Statist. Decisions* **24** 61–96. [MR2323189](#)
- [25] FÖLLMER, H. and SCHIED, A. (2010). *Stochastic Finance*, 3rd. ed. de Gruyter, Berlin.
- [26] FRITTELLI, M. and ROSAZZA GIANIN, E. (2002). Putting order in risk measures. *J. Bank. Financ.* **26** 1473–1486.
- [27] FRITTELLI, M. and ROSAZZA GIANIN, E. (2004). Dynamic convex risk measures. In *Risk Measures for the 21st Century* (G. Szegö, ed.) 227–248. Wiley, New York.
- [28] KAINA, M. and RÜSCHENDORF, L. (2009). On convex risk measures on L^p -spaces. *Math. Methods Oper. Res.* **69** 475–495.
- [29] KINGMAN, J. F. C. and ROBERTSON, A. P. (1968). On a theorem of Lyapunov. *J. Lond. Math. Soc.* **43** 347–351.
- [30] KRÄTSCHMER, V. and SCHOENMAKERS, J. (2010). Representations for optimal stopping under dynamic monetary utility functionals. *SIAM J. Financial Math.* **1** 811–832.
- [31] KÜHN, Z. and RÖSLER, U. (1998). A generalization of Lyapunov’s convexity theorem with applications in optimal stopping. *Proc. Amer. Math. Soc.* **126** 769–777.
- [32] KUPPER, M. and SCHACHERMAYER, W. (2009). Representation results for law invariant time consistent functions. *Math. Financ. Econ.* **2** 189–210.

- [33] RIEDEL, F. (2009). Optimal stopping with multiple priors. *Econometrica* **77** 857–908.
- [34] ROCKAFELLAR, R. T. and WETS, J.-B. (1998). *Variational Analysis*. Springer, Berlin/Heidelberg.
- [35] ROGERS, L. C. G. (2002). Monte Carlo valuation of American options. *Math. Finance* **12** 271–286.
- [36] WITTING, H. and MÜLLER-FUNK, U. (1995). *Mathematische Statistik II*. Teubner, Stuttgart.
- [37] XU, Z. Q. and ZHOU, X. Y. (2013). Optimal stopping under probability distortion. *Ann. Appl. Probab.* **23** 251–282.

FACULTY OF MATHEMATICS
DUISBURG-ESSEN UNIVERSITY
THEA-LEYMANN-STR. 9
D-45127 ESSEN
GERMANY
AND
HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY
MOSCOW
RUSSIAN FEDERATION
E-MAIL: denis.belomestny@uni-due.de

FACULTY OF MATHEMATICS
DUISBURG-ESSEN UNIVERSITY
THEA-LEYMANN-STR. 9
D-45127 ESSEN
GERMANY
E-MAIL: volker.kraetschmer@uni-due.de