Supplementary Material to "Robust Multiple Stopping — A Duality Approach"

Roger J. A. Laeven*
University of Amsterdam
Mail: University of Amsterdam
Dept. of Quantitative Economics
PO Box 15867
1001 NJ Amsterdam, The Netherlands
Email: `R.J.A.Laeven@uva.nl`

John G. M. Schoenmakers
Weierstrass Institute Berlin
Mail: Weierstrass Institute Berlin
Stochastic Algorithms and Nonparametric Statistics
Mohrenstrasse 39
D-10117 Berlin, Germany
Email: `schoenma@wias-berlin.de`

Nikolaus Schweizer
Tilburg University
Mail: Tilburg University
Dept. of Econometrics and Operations Research
PO Box 90153
5000 LE Tilburg, The Netherlands
Email: `N.F.F.Schweizer@tilburguniversity.edu`

Mitja Stadje
Ulm University
Mail: Ulm University
Faculty of Mathematics and Economics
Helmholtzstrasse 20
D-89069 Ulm, Germany
Email: `Mitja.Stadje@uni-ulm.de`

This Version: November 13, 2023

Abstract

This online appendix provides: (i) the proofs of our results and several auxiliary results; (ii) six additional tables.

*Corresponding author.

ONLINE APPENDIX

A Proofs and Auxiliary Results for Section 2

A.1 An Auxiliary Lemma on Sensitivity (P2) and the Proofs of Lemmas 2.1 and 2.2

We state the following auxiliary lemma:

Lemma A.1 *(P2a) If subadditivity (P1) applies, then sensitivity (P2) of ρ implies*

$$[X \geq 0 \quad \text{and} \quad \rho_t(X) \leq 0] \Longrightarrow X = 0, \quad \text{for all } X \in \mathfrak{X}, \text{ and } t \in \{0, \dots, T\}. \tag{A.1}$$

Proof Let ρ be subadditive (P1) and sensitive (P2). Suppose $\rho_i(Y) \leq 0$ and $Y \geq 0$. Then, $-Y \leq 0$ and so, by subadditivity,

$$0 = \rho_t(Y - Y) \leq \rho_t(Y) + \rho_t(-Y) \leq \rho_t(-Y).$$

Hence, by (P2), $-Y = 0$, i.e., $Y = 0$ a.s. ■

Proof of Lemma 2.1. For $i = T$ the statement is trivial. Assuming that it holds for $0 < i \leq T$, we have

$$\begin{aligned}
\rho_{i-1}\left(M_{\tau_{i-1}}\right) &= \rho_{i-1}\left(1_{\{\tau_{i-1}=i-1\}}M_{i-1} + 1_{\{\tau_{i-1}>i-1\}}M_{\tau_{i-1}\vee i}\right)\\
\text{(by (C2) and (C4))} &= 1_{\{\tau_{i-1}=i-1\}}M_{i-1} + 1_{\{\tau_{i-1}>i-1\}}\rho_{i-1}\left(M_{\tau_{i-1}\vee i}\right)\\
\text{(by (C3))} &= 1_{\{\tau_{i-1}=i-1\}}M_{i-1} + 1_{\{\tau_{i-1}>i-1\}}\rho_{i-1}\circ\rho_i\left(M_{\tau_{i-1}\vee i}\right)\\
\text{(by induction)} &= 1_{\{\tau_{i-1}=i-1\}}M_{i-1} + 1_{\{\tau_{i-1}>i-1\}}\rho_{i-1}\left(M_i\right)\\
\text{(property of } \rho\text{-martingale)} &= M_{i-1}.
\end{aligned}$$

■

Proof of Lemma 2.2. For an arbitrary $X \in \mathfrak{X}$, and an arbitrary set $A \in \mathcal{B}(\mathbb{R})$, we have, for any $i \geq t$,

$$\begin{aligned}
\{\rho_\tau(X) \in A\} \cap \{\tau = i\} &= \left\{\sum_{j=t}^{T} 1_{\{\tau=j\}}\rho_j(X) \in A\right\} \cap \{\tau = i\}\\
&= \{\rho_i(X) \in A\} \cap \{\tau = i\} \in \mathcal{F}_i,
\end{aligned}$$

hence $\rho_\tau(X) \in \mathcal{F}_\tau$.

We next show (i) and (ii) in a joint induction. For $t = T$, both statements are trivial. Suppose they are true for $0 < t \leq T$. Now let $\tau \geq t - 1$ and define $\tau_1 := \tau \vee t$. Then,

$$\rho_\tau(X) = 1_{\{\tau=t-1\}}\rho_{t-1}(X) + 1_{\{\tau>t-1\}}\rho_{\tau_1}(X). \tag{A.2}$$

Thus, for any $X \in \mathfrak{X}$,

$$\begin{aligned}
\rho_{t-1}(X) =& 1_{\{\tau=t-1\}}\rho_{t-1}(X) + 1_{\{\tau>t-1\}}\rho_{t-1}(X) \\
\text{(by (C3) and induction)} =& 1_{\{\tau=t-1\}}\rho_{t-1}(X) + 1_{\{\tau>t-1\}}\rho_{t-1} \circ \rho_t \circ \rho_{\tau_1}(X) \\
\text{(by (C2) and (A.2))} =& 1_{\{\tau=t-1\}}\rho_{t-1}(X) + 1_{\{\tau>t-1\}}\rho_{t-1} \circ \rho_t \circ \left(\rho_\tau(X) - 1_{\{\tau=t-1\}}\rho_{t-1}(X)\right) \\
\text{(by (C4))} =& 1_{\{\tau=t-1\}}\rho_{t-1}(X) + 1_{\{\tau>t-1\}} \left(\rho_{t-1} \circ \rho_t \circ \rho_\tau(X) - 1_{\{\tau=t-1\}}\rho_{t-1}(X)\right) \\
\text{(by (C3) and (A.2))} =& 1_{\{\tau=t-1\}}\rho_\tau(X) + 1_{\{\tau>t-1\}}\rho_{t-1} \circ \rho_\tau(X) \\
\text{(by i) and (C6))} =& \rho_{t-1} \circ \rho_\tau(X).
\end{aligned}$$

This settles the induction step for (i). For the induction step of (ii), let $X \in \mathcal{F}_\tau$ and $Y \in \mathcal{F}_T$. Then, $1_{\{\tau=t-1\}}X \in \mathcal{F}_{t-1}$. Indeed, for any $A \in \mathcal{B}(\mathbb{R})$ one has

$$\begin{aligned}
\left\{1_{\{\tau=t-1\}}X \in A\right\} &= \left(\left\{1_{\{\tau=t-1\}}X \in A\right\} \cap \{\tau = t-1\}\right) \cup (\{0 \in A\} \cap \{\tau > t-1\}) \\
&= (\{X \in A\} \cap \{\tau = t-1\}) \cup (\{0 \in A\} \cap \{\tau > t-1\}) \in \mathcal{F}_{t-1}.
\end{aligned}$$

Furthermore, also $X \in \mathcal{F}_{\tau_1}$ since $\tau_1 \geq \tau$. Hence, we have by (A.2)

$$\begin{aligned}
\rho_\tau(X+Y) =& 1_{\{\tau=t-1\}}\rho_{t-1}(X+Y) + 1_{\{\tau>t-1\}}\rho_{\tau_1}(X+Y) \\
\text{((C2) and induction)} =& 1_{\{\tau=t-1\}}\rho_{t-1}(1_{\{\tau=t-1\}}X + 1_{\{\tau=t-1\}}Y) + 1_{\{\tau>t-1\}} \left(X + \rho_{\tau_1}(Y)\right) \\
\text{((C4) and above argument)} =& 1_{\{\tau=t-1\}} \left(1_{\{\tau=t-1\}}X + 1_{\{\tau=t-1\}}\rho_{t-1}(Y)\right) + 1_{\{\tau>t-1\}} \left(X + \rho_{\tau_1}(Y)\right) \\
=& X + 1_{\{\tau=t-1\}}\rho_{t-1}(Y) + 1_{\{\tau>t-1\}}\rho_{\tau_1}(Y) \\
\text{(by (A.2))} =& X + \rho_\tau(Y).
\end{aligned}$$

■

A.2 Auxiliary Results on the Robust Single Optimal Stopping Problem (2.3)

As is well-known, for the robust optimal single stopping problem, we may find an optimal stopping family $(\tau_t^*)_{t\in\{0,\ldots,T\}}$ satisfying

$$Y_t^* = \sup_{\tau\in\mathcal{T}_t} \rho_t(H_\tau) = \rho_t(H_{\tau_t^*}), \qquad t \in \{0,\ldots,T\},$$

and, furthermore, the Bellman principle

$$Y_t^* = \max\left(H_t, \rho_t\left(Y_{t+1}^*\right)\right), \qquad t \in \{0,\ldots,T-1\}, \tag{A.3}$$

is satisfied (see e.g., Krätschmer and Schoenmakers [2] and Krätschmer *et al.* [3] for details).

Let us briefly recall the already existing (non-pathwise) additive dual representation for the optimal single stopping problem (2.3) (cf. Krätschmer and Schoenmakers [2] and Krätschmer *et al.* [3]), but with a different proof adapted to the goals in this paper and exploited later.

Proposition A.2 *Let ρ be a DMU satisfying (C1)–(C4) and let $M^* = M^{*\rho} \in \mathcal{M}_0^\rho$ be the unique ρ-martingale in the ρ-Doob decomposition of $Y^* = (Y_t^*)_{0\leq t\leq T}$. Then the optimal single stopping problem* (2.3) *has an additive dual representation*

$$\begin{aligned}
Y_t^* &= \inf_{M\in\mathcal{M}_0^\rho} \rho_t\left(\max_{j\in\{t,\ldots,T\}} (H_j + M_T - M_j)\right) \qquad \text{(A.4)} \\
&= \rho_t\left(\max_{j\in\{t,\ldots,T\}} \left(H_j + M_T^* - M_j^*\right)\right), \quad t \in \{0,\ldots,T\}.
\end{aligned}$$

Proof of Proposition A.2. For any ρ-martingale M and any stopping time $\tau \geq t$, we have by Lemmas 2.1 and 2.2 that

$$\begin{aligned}\rho_t\Big(\max_{j\in\{t,\dots,T\}}(H_j+M_T-M_j)\Big) &\geq \rho_t\left(H_\tau+M_T-M_\tau\right)=\rho_t\circ\rho_\tau\left(H_\tau+M_T-M_\tau\right)\\ &=\rho_t\left(H_\tau-M_\tau+\rho_\tau\left(M_T\right)\right)=\rho_t\left(H_\tau\right),\end{aligned}$$

which implies

$$Y_t^*\leq \inf_{M\in\mathcal{M}_0^\rho}\rho_t\Big(\max_{j\in\{t,\dots,T\}}(H_j+M_T-M_j)\Big).$$

On the other hand, for the ρ-Doob martingale M^* it holds that

$$\begin{aligned}H_j+M_t^*-M_j^* &= H_j+\sum_{r=t}^{j-1}M_r^*-M_{r+1}^*\\ \text{(by (2.5))} &= H_j+\sum_{r=t}^{j-1}\rho_r\left(Y_{r+1}^*\right)-Y_{r+1}^*\\ \text{(Bellman)} &\leq H_j+\sum_{r=t}^{j-1}Y_r^*-Y_{r+1}^*=Y_t^*+H_j-Y_j^*\leq Y_t^*,\end{aligned}$$

whence

$$\begin{aligned}\rho_t\Big(\max_{j\in\{t,\dots,T\}}\left(H_j+M_T^*-M_j^*\right)\Big) &= \rho_t\Big(\max_{j\in\{t,\dots,T\}}\left(H_j+M_t^*-M_j^*\right)+M_T^*-M_t^*\Big)\\ &\leq Y_t^*+\rho_t\Big(M_T^*-M_t^*\Big)=Y_t^*.\end{aligned}$$

■

B Proofs and Auxiliary Results for Section 3

B.1 Proofs and Auxiliary Results for Section 3.1

We state the following lemma:

Lemma B.1 *Suppose that ρ satisfies (C1)–(C4) and (P1). Then, for any adapted process H, any ρ-martingale M, and any stopping τ, with $T\geq\tau\geq t$ a.s. it holds that*

$$\rho_t(H_\tau)\leq\rho_t(H_\tau+M_t-M_\tau),\qquad 0\leq t\leq T.$$

Proof of Lemma B.1. Using Lemma 2.2 and the proof of Proposition A.2 one has

$$\begin{aligned}\rho_t(H_\tau) &= \rho_t(H_\tau+M_T-M_\tau)\\ &= \rho_t(H_\tau+M_t-M_\tau+M_T-M_t)\\ \text{(by (P1))} &\leq \rho_t\left(H_\tau+M_t-M_\tau\right)+\rho_t(M_T-M_t)\\ \text{((C4) and }\rho\text{-mart. prop.)} &= \rho_t\left(H_\tau+M_t-M_\tau\right).\end{aligned}$$

■

Proof of Theorem 3.1. First we show that for any $t = 0, \ldots, T$, any sequence of stopping times with $t \le \tau_1 < \tau_2 < \cdots < \tau_L$ a.s., and any set of ρ-martingales $M^{(1)}, \ldots, M^{(L)}$, we have

$$\rho_t\left(\sum_{k=1}^{L} H_{\tau_k}\right) \le \rho_t\left(\sum_{k=1}^{L}\left(H_{\tau_k} + M^{(k)}_{\tau_{k-1}} - M^{(k)}_{\tau_k}\right)\right). \tag{B.1}$$

For $L = 1$, this statement boils down to Lemma B.1. Let us assume the statement is true for some $L \ge 1$. Take $0 \le t \le T$ and $t \le \tau_1 < \tau_2 < \cdots < \tau_{L+1}$ arbitrarily. Observe that (with $\rho_{T+1} := \rho_T$)

$$\begin{aligned}
&\rho_{\tau_1}\left(\sum_{k=2}^{L+1}\left(H_{\tau_k} + M^{(k)}_{\tau_{k-1}} - M^{(k)}_{\tau_k}\right)\right) \qquad \text{(B.2)}\\
&= \sum_{j=t}^{T} 1_{\{\tau_1 = j\}} \rho_j \circ \rho_{j+1}\left(1_{\{j+1 \le \tau_2 < \cdots < \tau_{L+1}\}} \sum_{k=2}^{L+1}\left(H_{\tau_k} + M^{(k)}_{\tau_{k-1}} - M^{(k)}_{\tau_k}\right)\right)\\
\text{(by induction)} &\ge \sum_{j=\tau_1}^{T} 1_{\{\tau_1 = j\}} \rho_j\left(1_{\{j < \tau_2 < \cdots < \tau_{L+1}\}} \sum_{k=2}^{L+1} H_{\tau_k}\right) = \rho_{\tau_1}\left(\sum_{k=2}^{L+1} H_{\tau_k}\right).
\end{aligned}$$

One may thus write, by Lemma 2.2,

$$\begin{aligned}
&\rho_t\left(\sum_{k=1}^{L+1}\left(H_{\tau_k} + M^{(k)}_{\tau_{k-1}} - M^{(k)}_{\tau_k}\right)\right)\\
&= \rho_t\left(H_{\tau_1} + M^{(1)}_t - M^{(1)}_{\tau_1} + \rho_{\tau_1}\left(\sum_{k=2}^{L+1}\left(H_{\tau_k} + M^{(k)}_{\tau_{k-1}} - M^{(k)}_{\tau_k}\right)\right)\right)\\
\text{(by Lemma B.1)} &\ge \rho_t\left(H_{\tau_1} + \rho_{\tau_1}\left(\sum_{k=2}^{L+1}\left(H_{\tau_k} + M^{(k)}_{\tau_{k-1}} - M^{(k)}_{\tau_k}\right)\right)\right)\\
\text{(by (B.2))} &\ge \rho_t\left(H_{\tau_1} + \rho_{\tau_1}\left(\sum_{k=2}^{L+1} H_{\tau_k}\right)\right)\\
\text{(Lemma 2.2)} &= \rho_t\left(\sum_{k=1}^{L+1} H_{\tau_k}\right),
\end{aligned}$$

which proves (B.1). As a corollary, we obtain

$$Y_t^{*,L} \le \rho_t\left(\max_{t \le j_1 < j_2 < \cdots < j_L} \sum_{k=1}^{L}\left(H_{j_k} + M^{(k)}_{j_{k-1}} - M^{(k)}_{j_k}\right)\right), \tag{B.3}$$

where we note that for any set A of probability one has

$$1_A \rho_t(X) = \rho_t(1_A X) = \rho_t(X),$$

due to monotonicity (C1). Since the ρ-martingales $M^{(k)}$ are arbitrary, we thus arrive at

$$Y_t^{*,L} \le \inf_{M^{(1)}, \ldots, M^{(L)} \in \mathcal{M}_0^\rho} \rho_t\left(\max_{t \le j_1 < j_2 < \cdots < j_L} \sum_{k=1}^{L}\left(H_{j_k} + M^{(k)}_{j_{k-1}} - M^{(k)}_{j_k}\right)\right). \tag{B.4}$$

On the other hand, for the ρ-Doob martingales $M^{*,L-k+1}$, i.e., the martingale parts of the Snell envelope due to $L-k+1$ exercise rights, we may write (with $j_0 = t$)

$$
\begin{aligned}
&\sum_{k=1}^{L}\left(H_{j_k} + M^{*,L-k+1}_{j_{k-1}} - M^{*,L-k+1}_{j_k}\right)\\
&= \sum_{k=1}^{L}\left(H_{j_k} + \sum_{r=j_{k-1}}^{j_k-1}\left(M^{*,L-k+1}_{r} - M^{*,L-k+1}_{r+1}\right)\right)\\
&= \sum_{k=1}^{L} H_{j_k} + \sum_{k=1}^{L}\sum_{r=j_{k-1}}^{j_k-1}\left(\rho_r\left(Y^{*,L-k+1}_{r+1}\right) - Y^{*,L-k+1}_{r+1}\right)\\
&= \sum_{k=1}^{L} H_{j_k} + \sum_{k=1}^{L}\sum_{r=j_{k-1}}^{j_k-1}\left(Y^{*,L-k+1}_{r} - Y^{*,L-k+1}_{r+1}\right)\\
&\quad + \sum_{k=1}^{L}\sum_{r=j_{k-1}}^{j_k-1}\left(\rho_r\left(Y^{*,L-k+1}_{r+1}\right) - Y^{*,L-k+1}_{r}\right)\\
&= \sum_{k=1}^{L} H_{j_k} + \sum_{k=1}^{L}\left(Y^{*,L-k+1}_{j_{k-1}} - Y^{*,L-k+1}_{j_k}\right)\\
&\quad + \sum_{k=1}^{L}\sum_{r=j_{k-1}}^{j_k-1}\left(\rho_r\left(Y^{*,L-k+1}_{r+1}\right) - Y^{*,L-k+1}_{r}\right)\\
&= Y^{*,L}_{j_0} + \underbrace{H_{j_L} - Y^{*,1}_{j_L}}_{\leq 0} + \sum_{k=1}^{L-1}\underbrace{\left(H_{j_k} + Y^{*,L-k}_{j_k} - Y^{*,L-k+1}_{j_k}\right)}_{\leq 0}\\
&\quad + \sum_{k=1}^{L}\sum_{r=j_{k-1}}^{j_k-1}\underbrace{\left(\rho_r\left(Y^{*,L-k+1}_{r+1}\right) - Y^{*,L-k+1}_{r}\right)}_{\leq 0} \leq Y^{*,L}_{j_0}.
\end{aligned}
$$

That is,

$$\max_{t\leq j_1<j_2<\cdots<j_L}\sum_{k=1}^{L}\left(H_{j_k} + M^{*,L-k+1}_{j_{k-1}} - M^{*,L-k+1}_{j_k}\right) \leq Y^{*,L}_{t},$$

while, due to (B.3),

$$\rho_t\left(\max_{t\leq j_1<j_2<\cdots<j_L}\sum_{k=1}^{L}\left(H_{j_k} + M^{*,L-k+1}_{j_{k-1}} - M^{*,L-k+1}_{j_k}\right)\right) \geq Y^{*,L}_{t}.$$

Thus, by monotonicity (C1) and $\mathcal{F}_t$-invariance (C6) we obtain (ii), and, by sensitivity (P2), we obtain (iii). Finally, (ii) combined with (B.4) yields (i). ■

B.2 Proofs and Auxiliary Results for Section 3.2

Proof of Theorem 3.4. Suppose that $\theta_i := \max_{i\leq j\leq T}(H_j - M_j + M_i) \in \mathcal{F}_i$ and define the stopping time

$$\tau_i := \inf\{j \geq i : H_j - M_j + M_i \geq \theta_i\}.$$

By the definition of θ_i, clearly $i \le \tau_i \le T$ a.s. Hence, we have on the one hand

$$Y_i^* \ge \rho_i(H_{\tau_i}) \ge \rho_i(M_{\tau_i} - M_i + \theta_i) = \theta_i,$$

by the fact that $-M_i + \theta_i \in \mathcal{F}_i$, translation invariance (C4), and Lemma 2.1. On the other hand, we have $\theta_i = \rho_i(\theta_i) \ge Y_i^*$ due to Theorem 3.2, Eqn. (3.6). ■

Proof of Lemma 3.5. By writing

$$\theta_{i+} = \underbrace{\max_{i+1 \le j \le T} (H_j - M_j + M_{i+1})}_{\in \mathcal{F}_{i+1}} + M_i - M_{i+1}, \tag{B.5}$$

and applying Theorem 3.4, we have

$$\theta_{i+} + M_{i+1} - M_i = Y_{i+1}^*. \tag{B.6}$$

Then, (i) follows by applying ρ_i on both sides, using conditional translation invariance (C4) and the martingale property. Next, (ii) is obvious from (B.6). ■

Proof of Proposition 3.6. It is sufficient to show that

$$\rho_i\left(1_{\{|Y| \ge \epsilon\}}\right) \le \frac{\rho_i\left(Y^2\right)}{\epsilon^2}. \tag{B.7}$$

Indeed, one has by monotonicity and positive homogeneity,

$$\rho_i\left(Y^2\right) = \rho_i\left(Y^2 1_{\{|Y| \ge \epsilon\}} + Y^2 1_{\{|Y| < \epsilon\}}\right) \ge \rho_i\left(Y^2 1_{\{|Y| \ge \epsilon\}}\right) \ge \rho_i\left(\epsilon^2 1_{\{|Y| \ge \epsilon\}}\right) = \epsilon^2 \rho_i\left(1_{\{|Y| \ge \epsilon\}}\right).$$

■

Proof of Lemma 3.7. Indeed, $\mathrm{Var}_{\rho_i}(X) = \rho_i\left((X - \rho_i(X))^2\right) = 0$ implies, by (A.1), $X - \rho_i(X) = 0$, hence $X \in \mathcal{F}_i$. The reverse direction is trivial. ■

Proof of Theorem 3.8. Suppose that the assumptions of the theorem are satisfied. Fix an $i \in \{0, \ldots, T\}$ and take an $\epsilon > 0$. Upon introducing an auxiliary time $\partial > T$ and setting $H_\partial = 0$, we next define the stopping time $\tau_i^{(n)} = \inf\{j \ge i : H_j - M_j^{(n)} + M_i^{(n)} \ge \rho_i(\theta_i^{(n)}) - \epsilon\} \wedge \partial$.

Then, with $M_\partial^{(n)} := M_T^{(n)}$, $n \ge 1$,

$$\begin{aligned}
Y_i^* &\ge \rho_i(H_{\tau_i^{(n)}}) = \rho_i(H_{\tau_i^{(n)}} 1_{\{\tau_i^{(n)} < \partial\}}) \\
&\ge \rho_i\Big((M_{\tau_i^{(n)}}^{(n)} - M_i^{(n)} + \rho_i(\theta_i^{(n)}) - \epsilon) 1_{\{\tau_i^{(n)} < \partial\}}\Big) \\
&\ge \rho_i\Big(M_{\tau_i^{(n)}}^{(n)} - M_i^{(n)} + \rho_i(\theta^{(n)}) - \epsilon\Big) - \rho_i\Big((M_T^{(n)} - M_i^{(n)} + \rho_i(\theta_i^{(n)}) - \epsilon) 1_{\{\tau_i^{(n)} = \partial\}}\Big) \\
&= \rho_i(\theta_i^{(n)}) - \epsilon - \rho_i\Big((M_T^{(n)} - M_i^{(n)} + \rho_i(\theta_i^{(n)}) - \epsilon) 1_{\{\tau_i^{(n)} = \partial\}}\Big), \qquad \text{almost surely,}
\end{aligned}$$

using subadditivity in the last inequality and translation invariance in the last equality. Hence,

$$\begin{aligned}
\rho_i\left(\theta_i^{(n)}\right) &\le Y_i^* + \epsilon + \rho_i\left(\left|M_T^{(n)} - M_i^{(n)} + \rho_i\left(\theta_i^{(n)}\right) - \epsilon\right| 1_{\{\tau_i^{(n)} = \partial\}}\right) \\
&=: Y_i^* + \epsilon + \rho_i\left(\left|U_i^{(n)}\right| 1_{\{\tau_i^{(n)} = \partial\}}\right), \qquad \text{almost surely.}
\end{aligned}$$

By (3.10),

$$\rho_i\left(1_{\left\{\tau_i^{(n)}=\partial\right\}}\right)=\rho_i\left(1_{\left\{\left|\theta_i^{(n)}-\rho_i\left(\theta_i^{(n)}\right)\right|\geq\epsilon\right\}}\right)\leq\frac{\mathrm{Var}_{\rho_i}\left(\theta_i^{(n)}\right)}{\epsilon^2}\overset{\mathrm{P}}{\to}0,$$

and since moreover by monotonicity $0\leq\rho_i\left(1_{\left\{\tau_i^{(n)}=\partial\right\}}\right)\leq\rho_i(1)=1$, it holds that

$$\rho_i\left(1_{\left\{\tau_i^{(n)}=\partial\right\}}\right)\overset{L^1}{\to}0. \tag{B.8}$$

Next, by subadditivity, monotonicity, and positive homogeneity, we have, for any $K>0$,

$$\begin{aligned}\mathbb{E}\rho_i\left(\left|U_i^{(n)}\right|1_{\left\{\tau_i^{(n)}=\partial\right\}}\right)&\leq\mathbb{E}\rho_i\left(\left|U_i^{(n)}\right|1_{\left\{\tau_i^{(n)}=\partial\right\}}1_{\left\{\left|U_i^{(n)}\right|\leq K\right\}}\right)+\mathbb{E}\rho_i\left(\left|U_i^{(n)}\right|1_{\left\{\tau_i^{(n)}=\partial\right\}}1_{\left\{\left|U_i^{(n)}\right|>K\right\}}\right)\\&\leq K\mathbb{E}\rho_i\left(1_{\left\{\tau_i^{(n)}=\partial\right\}}\right)+\mathbb{E}\rho_i\left(\left|U_i^{(n)}\right|1_{\left\{\left|U_i^{(n)}\right|>K\right\}}\right).\end{aligned}$$

Now Propositions B.2 and B.4 below imply that the family $\left(U_i^{(n)}\right)_{n\geq1}$ is also uniformly integrable in the sense of (3.11), i.e., there exists $K_{1,\epsilon}$ large enough such that

$$\sup_{n\geq1}\mathbb{E}\rho_i\left(\left|U_i^{(n)}\right|1_{\left\{\left|U_i^{(n)}\right|>K\right\}}\right)<\epsilon,$$

hence

$$\mathbb{E}\rho_i\left(\left|U_i^{(n)}\right|1_{\left\{\tau_i^{(n)}=\partial\right\}}\right)\leq K_\epsilon\underbrace{\mathbb{E}\rho_i\left(1_{\left\{\tau_i^{(n)}=\partial\right\}}\right)}_{\to0\text{ by (B.8)}}+\epsilon\leq2\epsilon,$$

for $n>N_{K_{1,\epsilon},\epsilon}$. Thus, since $\epsilon>0$ was arbitrary,

$$\overline{\lim}_{n\geq1}\mathbb{E}\rho_i\left(\theta_i^{(n)}\right)\leq\mathbb{E}Y_i^*+3\epsilon.$$

On the other hand, as a consequence of Corollary 3.2, i.e., by the duality theorem for optimal single stopping of functionals satisfying (C1)–(C4) and (P1),

$$\mathbb{E}\rho_i\left(\theta_i^{(n)}\right)\geq\mathbb{E}Y_i^*,$$

so it follows that

$$\lim_{n\to\infty}\mathbb{E}\rho_i\left(\theta_i^{(n)}\right)=\mathbb{E}Y_i^*.$$

■

Proposition B.2 *Suppose* $(A_n)_{n\geq1}$, *and* $(B_n)_{n\geq1}$ *satisfy (3.11), i.e.,*

$$\sup_{n\geq1}\mathbb{E}\rho_i\left(|A_n|1_{\{|A_n|>K_\epsilon\}}\right)<\epsilon\quad\textit{and}\quad\sup_{n\geq1}\mathbb{E}\rho_i\left(|B_n|1_{\{|B_n|>K_\epsilon\}}\right)<\epsilon,$$

for K_ϵ *large enough. If* ρ_i *is subadditive and positively homogeneous, then also* $(A_n+B_n)_{n\geq1}$ *satisfies (3.11).*

Proof of Proposition B.2. By (P1),

$$\begin{aligned}\mathbb{E}\rho_i\left(|A_n+B_n|\,1_{\{|A_n+B_n|>2K_\epsilon\}}\right) &\le \mathbb{E}\rho_i\left((|A_n|+|B_n|)\,1_{\{|A_n|+|B_n|>2K_\epsilon\}}\right)\\ &\le \mathbb{E}\rho_i\left(2\,|A_n|\,1_{\{|A_n|>K_\epsilon\}}+2\,|B_n|\,1_{\{|B_n|>K_\epsilon\}}\right)\\ &\le 2\mathbb{E}\rho_i\left(|A_n|\,1_{\{|A_n|>K_\epsilon\}}\right)+2\mathbb{E}\rho_i\left(|B_n|\,1_{\{|B_n|>K_\epsilon\}}\right)<4\epsilon.\end{aligned}$$

Hence, A_n+B_n satisfies (3.11) as well. ■

Lemma B.3 *Assume (P1) and (P3).* $(A_n)_{n\ge1}$ *(with w.l.o.g.* $A_n\ge0$*) satisfy (3.11) if and only if*

(i) $\sup_{n\ge1}\rho_i(A_n)<\infty$*;*

(ii) For every $\epsilon>0$ *there exists* $\delta>0$ *such that for all* $B\in\mathcal{F}$ *with* $\rho_i(1_B)<\delta$*, it holds that* $\sup_{n\ge1}\rho_i(A_n1_B)<\epsilon$.

Proof of Lemma B.3. ($\Longrightarrow$) Let $(A_n)_{n\ge1}$ (with w.l.o.g. $A_n\ge0$) satisfy (3.11). Then, for any $n\ge1$, by subadditivity, monotonicity, and positive homogeneity,

$$\rho_i(A_n)\le\rho_i\left(A_n1_{\{A_n\le K\}}\right)+\rho_i\left(A_n1_{\{A_n>K\}}\right)\le K\rho_i\left(1_{\{A_n\le K\}}\right)+1\le K+1,$$

for large enough K. So $\sup_{n\ge1}\rho_i(A_n)\le K+1$, whence (i). Now let $\epsilon>0$ and K be so large that

$$\sup_{n\ge1}\rho_i\left(A_n1_{\{A_n>K\}}\right)<\epsilon/2.$$

For any $B\in\mathcal{F}$ with $\rho_i(1_B)<\epsilon/(2K)=:\delta$ we then have

$$\rho_i(A_n1_B)\le\rho_i\left(A_n1_B1_{\{A_n\le K\}}\right)+\rho_i\left(A_n1_B1_{\{A_n>K\}}\right)\le K\rho_i(1_B)+\rho_i\left(A_n1_{\{A_n>K\}}\right)<\epsilon.$$

($\Longleftarrow$) Let $(A_n)_{n\ge1}$ satisfy (i) and (ii) for $\epsilon>0$ and $\delta>0$. For any $n\ge1$ we have

$$\rho_i(A_n)\ge\rho_i\left(A_n1_{\{A_n>K\}}\right)\ge K\rho_i\left(1_{\{A_n>K\}}\right),$$

so due to (i),

$$M:=\sup_{n\ge1}\rho_i(A_n)\ge K\sup_{n\ge1}\rho_i\left(1_{\{A_n>K\}}\right).$$

Hence,

$$\sup_{n\ge1}\rho_i\left(1_{\{A_n>K\}}\right)\le\frac{M}{K}<\delta,$$

if $K>M/\delta$. Thus, due to (ii), for all $n\ge1$, and $K>M/\delta$,

$$\rho_i\left(A_n1_{\{A_n>K\}}\right)<\epsilon.$$

■

Proposition B.4 *Let* ρ_i *be subadditive and positively homogeneous, and let* $(A_n)_{n\ge1}$ *satisfy (3.11). Then* $(\rho_i(A_n))_{n\ge1}$ *also satisfy (3.11).*

Proof of Proposition B.4. Due to Lemma B.3, (i) and (ii) apply for $(A_n)_{n\geq 1}$. Let $\epsilon > 0$ and take $\delta > 0$ such that (ii) holds for $(A_n)_{n\geq 1}$. Observe that

$$\rho_i(A_n) = \rho_i(\rho_i(A_n)) \geq \rho_i\left(\rho_i(A_n) 1_{\{\rho_i(A_n)>K\}}\right) \geq K\rho_i\left(1_{\{\rho_i(A_n)>K\}}\right).$$

Hence,

$$\sup_{n\geq 1} \rho_i\left(1_{\{\rho_i(A_n)>K\}}\right) \leq \frac{1}{K} \sup_{n\geq 1} \rho_i(A_n) =: \frac{M}{K}.$$

Take K such that $M/K < \delta$. Then, for all $n \geq 1$,

$$\rho_i\left(\rho_i(A_n) 1_{\{\rho_i(A_n)>K\}}\right) = \rho_i\left(A_n 1_{\{\rho_i(A_n)>K\}}\right) < \epsilon,$$

since $\rho_i\left(1_{\{\rho_i(A_n)>K\}}\right) < \delta$. That is, $(\rho_i(A_n))_{n\geq 1}$ satisfy (3.11). ■

Proof of Proposition 3.9. We have

$$\begin{aligned}
\sup_{n\geq 1} \mathbb{E}\rho_i\left(\left|M_i^{(n)}\right| 1_{\left\{\left|M_i^{(n)}\right|>K\right\}}\right) &= \sup_{n\geq 1} \mathbb{E}\rho_i\left(\frac{1}{\left|M_i^{(n)}\right|^{\eta}} \left|M_i^{(n)}\right|^{1+\eta} 1_{\left\{\left|M_i^{(n)}\right|>K\right\}}\right) \\
&\leq \frac{1}{K^{\eta}} \sup_{n\geq 1} \mathbb{E}\rho_i\left(\left|M_i^{(n)}\right|^{1+\eta}\right) \to 0 \quad \text{for} \quad K\to\infty.
\end{aligned}$$

■

Proof of Theorem 3.10. For $L = 1$, this follows from Lemma 3.5. Now let us suppose that

$$\Theta_{i+}^q \in \mathcal{F}_i, \quad \text{for} \quad q = 1, \ldots, L+1, \qquad 0 \leq i < T,$$

and that the theorem has been proved for $L \geq 1$. Then, by induction, we have (i) and (ii), and, with $j_0' = j_1$,

$$\begin{aligned}
\Theta_{i+}^{L+1} &= \max_{i<j_1<j_2<\cdots<j_{L+1}} \sum_{k=1}^{L+1} \left(H_{j_k} + M_{j_{k-1}}^{(L+2-k)} - M_{j_k}^{(L+2-k)}\right) \\
&= \max_{i<j_1} \Bigg(H_{j_1} + M_i^{(L+1)} - M_{j_1}^{(L+1)} \\
&\quad + \max_{j_1<j_2<\cdots<j_{L+1}} \sum_{k=2}^{L+1} \left(H_{j_k} + M_{j_{k-1}}^{(L+1-k+1)} - M_{j_k}^{(L+1-k+1)}\right) \Bigg) \\
&= \max_{i<j_1} \Bigg(H_{j_1} + M_i^{(L+1)} - M_{j_1}^{(L+1)} \\
&\quad + \max_{j_1<j_1'<\cdots<j_L'} \sum_{k=1}^{L} \left(H_{j_k'} + M_{j_{k-1}'}^{(L-k+1)} - M_{j_k'}^{(L-k+1)}\right) \Bigg) \\
&= \max_{i<j_1} \left(H_{j_1} + \Theta_{j_1+}^L + M_i^{(L+1)} - M_{j_1}^{(L+1)}\right) \\
&= \max_{i<j_1} \left(H_{j_1} + \rho_{j_1}\left(Y_{j_1+1}^{*,L}\right) + M_i^{(L+1)} - M_{j_1}^{(L+1)}\right).
\end{aligned}$$

Next, since $\Theta_{i+}^{L+1} \in \mathcal{F}_i$, Lemma 3.5 implies

$$\Theta_{i+}^{L+1} = \rho_i\left(Y_{i+1}^{*,L+1}\right) \quad \text{and} \quad M_{i+1}^{(L+1)} - M_i^{(L+1)} = Y_{i+1}^{*,L+1} - \rho_i\left(Y_{i+1}^{*,L+1}\right).$$

Since (ii) holds for L by induction, i.e., $M_{i+1}^{(q)} - M_i^{(q)} = Y_{i+1}^{*,q} - \rho_i\left(Y_{i+1}^{*,q}\right)$, $q = 1, \ldots, L$, and we have shown that $M_{i+1}^{(L+1)} - M_i^{(L+1)} = Y_{i+1}^{*,L+1} - \rho_i\left(Y_{i+1}^{*,L+1}\right)$, (ii) follows for $L+1$. Thus, although (ii) (for L) is not used in the derivations above, (ii) (for L) is needed to conclude the statement (ii) for $L+1$. ■

C Proofs of Section 4

We provide the following auxiliary lemma:

Lemma C.1 *A subadditive functional ρ_j satisfying* (3.12) *is Lipschitz continuous in L^p.*

Proof of Lemma C.1. For arbitrary $Z_1, Z_2 \in \mathfrak{X}$ one has that $Z_2 - Z_1 \in \mathfrak{X}$ and $Z_1 - Z_2 \in \mathfrak{X}$, since $\mathfrak{X}$ is a linear (sub-)space by assumption. It then follows, by subadditivity, that

$$\rho_j(Z_1) - \rho_j(Z_2) \leq \rho_j(Z_1 - Z_2) \quad \text{and} \quad \rho_j(Z_2) - \rho_j(Z_1) \leq \rho_j(Z_2 - Z_1).$$

Hence,

$$\begin{aligned} |\rho_j(Z_1) - \rho_j(Z_2)| &\leq \max\left(\rho_j(Z_1 - Z_2), \rho_j(Z_2 - Z_1)\right) \\ &\leq |\rho_j(Z_1 - Z_2)| + |\rho_j(Z_2 - Z_1)|, \end{aligned}$$

and so we have by (3.12),

$$\begin{aligned} \mathbb{E}\left[|\rho_j(Z_1) - \rho_j(Z_2)|^p\right]^{1/p} &\leq \mathbb{E}\left[|\rho_j(Z_1 - Z_2)|^p\right]^{1/p} + \mathbb{E}\left[|\rho_j(Z_2 - Z_1)|^p\right]^{1/p} \\ &\leq 2C_p^{1/p}\mathbb{E}\left[|Z_1 - Z_2|^p\right]^{1/p}. \end{aligned}$$

■

Proof of Lemma 4.1. We may write $\mathcal{C}^N = \mathcal{C}^N + \rho_j\left(\mathfrak{m}^N\right) = \rho_j\left(\mathcal{C}^N + \mathfrak{m}^N\right)$. Then, (4.2) follows from (4.1) and the L^2-Lipschitz continuity of ρ_j due to Lemma C.1. ■

Proof of Proposition 4.3. Let $\{\mathcal{B}_n : n \in \mathbb{N}\}$ be a dense subset of $L^2(\Omega, \mathcal{F}_{j+1}, \mathbb{P})$ by separability. Take an arbitrary $\mathcal{E} \in L^2_{j+1,0}$. There thus exists a sequence $(\mathcal{B}_{n_k})$ with $\mathcal{B}_{n_k} \overset{L^2}{\to} \mathcal{E}$, for $k \to \infty$. Next, consider $\mathcal{E}_k := \mathcal{B}_{n_k} - \rho(\mathcal{B}_{n_k})$. Hence, by translation invariance, $\rho_j(\mathcal{E}_k) = 0$, i.e., $\mathcal{E}_k \in L^2_{j+1,0}$ for all k. By L^2-continuity of ρ_j due to Lemma C.1 for $p = 2$, it follows that $\rho_j(\mathcal{B}_{n_k}) \overset{L^2}{\to} \rho_j(\mathcal{E}) = 0$, for $k \to \infty$, and so

$$\mathcal{E}_k = \mathcal{B}_{n_k} - \rho(\mathcal{B}_{n_k}) \overset{L^2}{\to} \mathcal{E}.$$

■

Proof of Theorem 4.4. Define $\overline{P}_j^{l,K,N} = \overline{M}_j^{l,K,N} - \overline{M}_T^{l,K,N} = -\sum_{r=j+1}^{T} \overline{\mathfrak{m}}_r^{l,K,N}$ and $P_j^{*,l} = M_j^{*,l} - M_T^{*,l}$. We first prove by induction that (4.13)–(4.14) hold together with

$$\lim_{K\to\infty} \lim_{N\to\infty} \overline{P}_j^{l,K,N} = P_j^{*,l} \text{ in } L^2, \tag{C.1}$$

for all $l = 1, \dots, L$ and $j = T, \dots, 0$. (4.12) then follows by noting that $\overline{P}_0^{l,K,N} = \overline{M}_0^{l,K,N} - \overline{M}_T^{l,K,N} = -\overline{M}_T^{l,K,N}$ so $\overline{M}_j^{l,K,N} = \overline{P}_j^{l,K,N} - \overline{P}_0^{l,K,N}$ and, analogously, $M_j^{*,l} = P_j^{*,l} - P_0^{*,l}$.

We perform a first, outer induction over $l = 1, \dots, L$ and, for each given l, a second, inner induction over $j = T, T-1, \dots, 0$. For the induction start, note that, for all $l = 1, \dots, L$, we have $\overline{Y}_T^{l,K,N} = H_T = Y_T^{*,l}$, $\overline{c}_T^{l,K,N} = 0 = c_T^{*,l}$ and, by construction, $\overline{P}_T^{l,K,N} = 0 = P_T^{*,l}$.

Assume now that (4.13)–(4.14) and (C.1) hold for a given l and $j+1 \leq T$. If $l > 1$, assume also that they hold for $l-1$ and all j. It follows from the Law of Large Numbers and, if $l > 1$, from the induction assumption for $l-1$ and $j+1$ that $\overline{\mathfrak{m}}_{j+1}^{l,K,N}$ and $\overline{c}_{j+1}^{l,K,N}$ converge a.s. to the projections of $\mathfrak{m}_{j+1}^{*,l}$ and $c_{j+1}^{*,l}$ on the spaces $\{\mathcal{E}_j^{(\beta_1,\dots,\beta_{K'})} | (\beta_1, \dots, \beta_{K'}) \in \mathbb{R}^{K'}\}$ and $\{\sum_{k=1}^{K''} \gamma_k \psi_k(X_j) | \gamma_k \in \mathbb{R}, k = 1, \dots, K''\}$. Letting $K = \min(K', K'')$ tend to infinity and using that both spaces form a basis, we can use Corollary 4.2 to conclude that (4.13) and (C.1) hold for j and l. It then follows directly from (4.6) (with l replaced by $l-1$) that (4.14) holds as well. This completes the double induction.

For simplicity, we drop the indexes K, N, n in the sequel. So we write $\overline{c}_j^l = \overline{c}_j^{l,K,N}(X_j^n)$. We let $\overline{c}_j^l$ and $c_j^{*,l}$ be a set of approximate and true continuation functions, respectively, let

$$\overline{U}_j^l = f_j(X_j) + \overline{c}_j^{l-1}(X_j), \qquad U_j^{*,l} = f_j(X_j) + c_j^{*,l-1}(X_j),$$

let $\overline{M}_j^l$ and $M_j^{*,l}$ be a set of approximate and true ρ-Doob martingales, and let $\overline{Y}_j^l$ and $Y_j^{*,l}$ be a set of approximate and true upper Snell envelopes. Then consider

$$\begin{aligned}
&\max_{j\leq r\leq T} \left(\overline{U}_r^l - \overline{M}_r^l\right) - Y_j^{*,l} = \max_{j\leq r\leq T} \left(\overline{U}_r^l - \overline{M}_r^l\right) - \max_{j\leq r\leq T} \left(U_r^{*,l} - M_r^{*,l}\right) \\
&= \max_{j\leq r\leq T} \left(\overline{U}_r^l - \overline{M}_r^l\right) - \max_{j\leq r\leq T} \left(\overline{U}_r^l - M_r^{*,l}\right) \\
&\quad + \max_{j\leq r\leq T} \left(\overline{U}_r^l - M_r^{*,l}\right) - \max_{j\leq r\leq T} \left(U_r^{*,l} - M_r^{*,l}\right) \\
&\leq \max_{j\leq r\leq T} \left(\overline{U}_r^l - \overline{M}_r^l - \left(\overline{U}_r^l - M_r^{*,l}\right)\right) \\
&\quad + \max_{j\leq r\leq T} \left(\overline{U}_r^l - M_r^{*,l} - \left(U_r^{*,l} - M_r^{*,l}\right)\right) \\
&= \max_{j\leq r\leq T} \left(M_r^{*,l} - \overline{M}_r^l\right) + \max_{j\leq r\leq T} \left(\overline{c}_r^{l-1} - c_r^{*,l-1}\right) \\
&\leq \max_{j\leq r\leq T} \left|M_r^{*,l} - \overline{M}_r^l\right| + \max_{j\leq r\leq T} \left|\overline{c}_r^{l-1} - c_r^{*,l-1}\right|.
\end{aligned}$$

Similarly,

$$\begin{aligned}
&Y_j^{*,l} - \max_{j\le r\le T}\left(\overline{U}_r^l - \overline{M}_r^l\right) = \max_{j\le r\le T}\left(U_r^{*,l} - M_r^{*,l}\right) - \max_{j\le r\le T}\left(\overline{U}_r^l - \overline{M}_r^l\right)\\
&= \max_{j\le r\le T}\left(U_r^{*,l} - M_r^{*,l}\right) - \max_{j\le r\le T}\left(\overline{U}_r^l - M_r^{*,l}\right)\\
&\qquad + \max_{j\le r\le T}\left(\overline{U}_r^l - M_r^{*,l}\right) - \max_{j\le r\le T}\left(\overline{U}_r^l - \overline{M}_r^l\right)\\
&= \max_{j\le r\le T}\left(U_r^{*,l} - M_r^{*,l} - \max_{j\le r'\le T}\left(\overline{U}_{r'}^l - M_{r'}^{*,l}\right)\right)\\
&\qquad + \max_{j\le r\le T}\left(\overline{U}_r^l - M_r^{*,l} - \max_{j\le r'\le T}\left(\overline{U}_{r'}^l - \overline{M}_{r'}^l\right)\right)\\
&\le \max_{j\le r\le T}\left(c_r^{*,l-1} - \overline{c}_r^{l-1}\right) + \max_{j\le r\le T}\left(\overline{M}_r^l - M_r^{*,l}\right)\\
&\le \max_{j\le r\le T}\left|M_r^{*,l} - \overline{M}_r^l\right| + \max_{j\le r\le T}\left|\overline{c}_r^{l-1} - c_r^{*,l-1}\right|,
\end{aligned}$$

whence

$$\left|Y_j^{*,l} - \max_{j\le r\le T}\left(\overline{U}_r^l - \overline{M}_r^l\right)\right| \le \max_{j\le r\le T}\left|M_r^{*,l} - \overline{M}_r^l\right| + \max_{j\le r\le T}\left|\overline{c}_r^{l-1} - c_r^{*,l-1}\right|.$$

That is, by monotonicity and subadditivity,

$$\begin{aligned}
&\rho_j\left(\left|Y_j^{*,l} - \max_{j\le r\le T}\left(\overline{U}_r^l - \overline{M}_r^l\right)\right|\right)\\
&\qquad \le \rho_j\left(\max_{j\le r\le T}\left|M_r^{*,l} - \overline{M}_r^l\right|\right) + \rho_j\left(\max_{j\le r\le T}\left|\overline{c}_r^{l-1} - c_r^{*,l-1}\right|\right).
\end{aligned} \tag{C.2}$$

By the first part of the theorem, the right-hand side in (C.2) goes to zero. Thus,

$$\begin{aligned}
&\left|Y_j^{*,l} - \rho_j\left(\max_{j\le r\le T}\left(\overline{U}_r^l - \overline{M}_r^l\right)\right)\right| = \left|\rho_j\left(Y_j^{*,l}\right) - \rho_j\left(\max_{j\le r\le T}\left(\overline{U}_r^l - \overline{M}_r^l\right)\right)\right|\\
&\le \rho_j\left(\left|Y_j^{*,l} - \max_{j\le r\le T}\left(\overline{U}_r^l - \overline{M}_r^l\right)\right|\right)\\
&\le \rho_j\left(\max_{j\le r\le T}\left|M_r^{*,l} - \overline{M}_r^l\right|\right) + \rho_j\left(\max_{j\le r\le T}\left|\overline{c}_r^{l-1} - c_r^{*,l-1}\right|\right)
\end{aligned}$$

tends to zero as well. (Here, the first inequality follows as, by monotoniciy and subadditivity, $\rho(X) \le \rho(Y + |X-Y|)) \le \rho(Y) + \rho(|X-Y|)$ yielding $\rho(X) - \rho(Y) \le \rho(|X-Y|)$, and switching the roles of X and Y then gives the desired inequality.) ■

Proof of Proposition 4.5. We write, with $j_1' := i$,

$$\begin{aligned}
\Theta_i^q &= \max\Bigg[\max_{i<j_2'<\cdots<j_q'}\Bigg(f_i(X_i) + \sum_{l=2}^{q}\Big(f_{j_l'}(X_{j_l'}) - \overline{M}_{j_l'}^{q-l+1} + \overline{M}_{j_{l-1}'}^{q-l+1}\Big)\Bigg), \\
&\qquad \max_{i<j_1<j_2<\cdots<j_q}\sum_{l=1}^{q}\Big(f_{j_l}(X_{j_l}) - \overline{M}_{j_l}^{q-l+1} + \overline{M}_{j_{l-1}}^{q-l+1}\Big)\Bigg] \\
&= \max\Bigg[\max_{i+1\le j_2'<\cdots<j_q'}\Bigg(f_i(X_i) + \overline{M}_i^{q-1} - \overline{M}_{i+1}^{q-1} \\
&\qquad + \sum_{l=2}^{q}\Big(f_{j_l'}(X_{j_l'}) - \overline{M}_{j_l'}^{q-l+1} + \overline{M}_{j_{l-1}'\vee i+1}^{q-l+1}\Big)\Bigg), \\
&\qquad \max_{i+1\le j_1<j_2<\cdots<j_q}\overline{M}_i^q - \overline{M}_{i+1}^q + \sum_{l=1}^{q}\Big(f_{j_l}(X_{j_l}) - \overline{M}_{j_l}^{q-l+1} + \overline{M}_{j_{l-1}\vee i+1}^{q-l+1}\Big)\Bigg] \\
&= \max\Bigg[f_i(X_i) + \overline{M}_i^{q-1} - \overline{M}_{i+1}^{q-1} \\
&\qquad + \max_{i+1\le j_2'<\cdots<j_q'}\sum_{l=2}^{q}\Big(f_{j_l'}(X_{j_l'}) - \overline{M}_{j_l'}^{q-l+1} + \overline{M}_{j_{l-1}'\vee i+1}^{q-l+1}\Big), \\
&\qquad \overline{M}_i^q - \overline{M}_{i+1}^q + \max_{i+1\le j_1<j_2<\cdots<j_q}\sum_{l=1}^{q}\Big(f_{j_l}(X_{j_l}) - \overline{M}_{j_l}^{q-l+1} + \overline{M}_{j_{l-1}\vee i+1}^{q-l+1}\Big)\Bigg],
\end{aligned}$$

which is equal to (4.15). ∎

D Proofs of Section 5

Proof of Proposition 5.3. The result represents a generalization of Theorem 21 in Krätschmer *et al.* [3]. Following the exact same lines of reasoning as in the proof of Theorem 4.4, we perform a first, outer induction over the exercise rights and a second, inner and backward induction over the exercise dates. This yields, from the Law of Large Numbers and the induction assumption, the a.s. convergence of the time-discretized upper Snell envelope and the corresponding continuation values and ρ-martingale increments to suitable projections on spaces that form a basis, as in Theorem 4.4.

Next, their L^2-convergence when the time discretization $\Delta \to 0$ then follows from existing results on the convergence of solutions of discrete-time BSDEs (i.e., BSΔEs) to solutions of continuous-time BSDEs; see Briand *et al.* [1]. This proves the stated result. ∎

Proof of Theorem 5.4. First, $\mathbb{E}\left[\widetilde{Y}_0^{\text{low},L}\right] \le Y_0^{*,L}$ follows by (5.14). Second, the convergence statement follows by applying Proposition 5.3 and Theorem 4.4 three times. ∎

Proof of Proposition 5.5. We write

$$
\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left[U-\overline{M}_T^N\right] &= \mathbb{E}_{\mathbb{Q}}\left[U\right]-\mathbb{E}_{\mathbb{Q}}\left[\overline{M}_T^N\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[U\right]-\mathbb{E}_{\mathbb{Q}}\bigg[\int_0^T \mathcal{Z}_s^N dW_s^{\mathbb{Q}}+\int_0^T \tilde{\mathcal{Z}}_s^N d\tilde{N}_s^{\mathbb{Q}} \\
&\qquad +\int_0^T \left\{\mathcal{Z}_s^N q_s+\tilde{\mathcal{Z}}_s^N(\lambda_s-\lambda_{\mathbb{P}})-g(s,\mathcal{Z}_s^N,\tilde{\mathcal{Z}}_s^N)\right\}ds\bigg] \\
&= \mathbb{E}_{\mathbb{Q}}\left[U\right]+0+0=\mathbb{E}_{\mathbb{Q}}\left[U\right],
\end{aligned}
$$

where we used in the one but last equality that the convex conjugate satisfies

$$
\sup_{q,\lambda}\{zq+\tilde{z}(\lambda-\lambda_{\mathbb{P}})-g(t,z,\tilde{z})\}=0,
$$

as g is positively homogeneous. Moreover, this equality is attained above in $(q_s,\lambda_s-\lambda_{\mathbb{P}})\in\partial g(s,\mathcal{Z}_s^N,\tilde{\mathcal{Z}}_s^N)$. ■

Proof of Theorem 5.6. First, $\mathbb{E}\left[\widetilde{Y}_0^{\mathrm{upp},L}\right]\geq Y_0^{*,L}$ follows by (5.19). Second, the two convergence statements follow by applying Proposition 5.3 and Theorem 4.4 three times and two times, respectively. ■

Proof of Proposition 5.7. Fix $\alpha\in\mathbb{R}$. Applying Itô's generalized formula yields

$$
\begin{aligned}
&e^{\alpha t}|\delta Y_t|^2+\int_t^T e^{\alpha s}|\delta\mathcal{Z}_s^N|^2ds+\sum_{s\geq t:\ \delta Y\text{jumps at } s} e^{\alpha s}(|\delta Y_s|^2-|\delta Y_{s-}|^2-2\delta Y_{s-}\delta\tilde{\mathcal{Z}}_s^N) \\
&= e^{\alpha T}|\delta\xi|^2+\int_t^T e^{\alpha s}\left\{2\delta Y_s(g(\mathcal{Z}_s^N,\tilde{\mathcal{Z}}_s^N)-g(Z_s',\tilde{Z}_s'))-\alpha|\delta Y_s|^2\right\}ds \\
&\quad -2\int_t^T e^{\alpha s}\delta Y_s\delta\mathcal{Z}_s^N dW_s-2\int_t^T e^{\alpha s}\delta Y_s\delta\tilde{\mathcal{Z}}_s^N d\tilde{N}_s \\
&\leq e^{\alpha T}|\delta\xi|^2+\int_t^T e^{\alpha s}\left\{\mathcal{L}^2|\delta Y_s|^2+|\delta\mathcal{Z}_s^N|^2+|\delta\tilde{\mathcal{Z}}_s^N|^2-\alpha|\delta Y_s|^2\right\}ds \\
&\quad -2\int_t^T e^{\alpha s}\delta Y_s\delta\mathcal{Z}_s^N dW_s-2\int_t^T e^{\alpha s}\delta Y_s\delta\tilde{\mathcal{Z}}_s^N d\tilde{N}_s,
\end{aligned}
$$

using the Lipschitz continuity of g in the equality, and that $2ab\leq\mathcal{L}^2a^2+\frac{b^2}{\mathcal{L}^2}$ where $\mathcal{L}$ is the Lipschitz constant of g in the inequality. Choosing $\alpha=\mathcal{L}^2$ and observing that

$$
\sum_{s\geq t:\ \delta Y\text{ jumps at } s} e^{\alpha s}(|\delta Y_s|^2-|\delta Y_{s-}|^2-2\delta Y_{s-}\delta\tilde{\mathcal{Z}}_s^N)=\sum_{s\geq t:\ \delta Y\text{ jumps at } s} e^{\alpha s}|\delta\tilde{\mathcal{Z}}_s^N|^2,
$$

(which is the quadratic variation of the jump part of $e^{\alpha s/2}Y_s$) we obtain, for $t=0$,

$$
\begin{aligned}
&|\delta Y_0|^2 + \int_0^T e^{\alpha s}|\delta \mathcal{Z}_s^N|^2 ds + \sum_{s\geq 0:\ \delta Y \text{ jumps at } s} e^{\alpha s}|\delta \tilde{\mathcal{Z}}_s^N|^2 \\
&\leq e^{\mathcal{L}^2 T}|\delta\xi|^2 + \int_0^T e^{\alpha s}\Big\{|\delta \mathcal{Z}_s^N|^2 + |\delta \tilde{\mathcal{Z}}_s^N|^2\Big\}ds \\
&\quad - 2\int_0^T e^{\alpha s}\delta Y_s \delta \mathcal{Z}_s^N dW_s - 2\int_0^T e^{\alpha s}\delta Y_s \delta \tilde{\mathcal{Z}}_s^N d\tilde{N}_s.
\end{aligned}
$$

Taking expectations on both sides and cancelling the $\delta\mathcal{Z}$ and $\delta\tilde{\mathcal{Z}}$ terms corresponding to the quadratic variation yields the proposition. ■

E Additional Tables

L	1	2	3	4	5
LB	0.9722	1.7389	2.3707	2.8969	3.3350
s.e.	0.0012	0.0018	0.0023	0.0027	0.0031
$\overline{Y}_0^{N_4,L}$	0.9869	1.7603	2.3953	2.9235	3.3635
TE	0.0682	0.1046	0.1352	0.1631	0.1849
UB	1.0550	1.8649	2.5306	3.0866	3.5484

Table E.1: Bounds for $\delta_1 = 0$, $\delta_2 = 0$ and $J = 0.06$

L	1	2	3	4	5
LB	0.9884	1.7727	2.4173	2.9625	3.4123
s.e.	0.0015	0.0025	0.0032	0.0040	0.0045
$\overline{Y}_0^{N_4,L}$	1.0120	1.8091	2.4672	3.0184	3.4806
TE	0.0764	0.1102	0.1365	0.1586	0.1791
UB	1.0903	1.9221	2.6071	3.1810	3.6642

Table E.2: Bounds for $\delta_1 = \frac{1}{10}$, $\delta_2 = 0$ and $J = 0.06$

L	1	2	3	4	5
LB	0.9999	1.8066	2.4640	3.0228	3.4831
s.e.	0.0021	0.0037	0.0049	0.0061	0.0071
$\overline{Y}_0^{N_4,L}$	1.0370	1.8587	2.5407	3.1150	3.5996
TE	0.0747	0.1095	0.1413	0.1695	0.1942
UB	1.1196	1.9797	2.6969	3.3023	3.8142

Table E.3: Bounds for $\delta_1 = \frac{1}{5}$, $\delta_2 = 0$ and $J = 0.06$

L	1	2	3	4	5
LB	0.9776	1.7484	2.3814	2.9123	3.3529
s.e.	0.0020	0.0033	0.0044	0.0054	0.0062
$\overline{Y}_0^{N_4,L}$	1.0028	1.7907	2.4386	2.9790	3.4302
TE	0.0657	0.1003	0.1301	0.1569	0.1807
UB	1.0754	1.9015	2.5823	3.1524	3.6300

Table E.4: Bounds for $\delta_1 = 0$, $\delta_2 = \frac{1}{5}$ and $J = 0.06$

L	1	2	3	4	5
LB	0.9929	1.7811	2.4296	2.9688	3.4244
s.e.	0.0023	0.0038	0.0052	0.0063	0.0072
$\overline{Y}_0^{N_4,L}$	1.0275	1.8398	2.5113	3.0748	3.5483
TE	0.0740	0.1106	0.1395	0.1654	0.1875
UB	1.1093	1.9620	2.6655	3.2576	3.7555

Table E.5: Bounds for $\delta_1 = \frac{1}{10}$, $\delta_2 = \frac{1}{5}$ and $J = 0.06$

L	1	2	3	4	5
LB	1.0026	1.8031	2.4755	3.0091	3.4702
s.e.	0.0028	0.0049	0.0067	0.0081	0.0094
$\overline{Y}_0^{N_4,L}$	1.0536	1.8896	2.5848	3.1715	3.6676
TE	0.0749	0.1097	0.1420	0.1699	0.1947
UB	1.1365	2.0109	2.7418	3.3592	3.8828

Table E.6: Bounds for $\delta_1 = \frac{1}{5}$, $\delta_2 = \frac{1}{5}$ and $J = 0.06$

References

[1] Briand, P., Delyon, B., and Mémin, J. (2002). On the robustness of backward stochastic differential equations. *Stochastic Processes and their Applications* 97, 229–253.

[2] Krätschmer, V., and Schoenmakers, J. (2010). Representations for optimal stopping under dynamic monetary utility functionals. *SIAM Journal on Financial Mathematics* 1, 811–832.

[3] Krätschmer, V., Ladkau, M., Laeven, R.J.A., Schoenmakers, J.G.M., and Stadje, M. (2018). Optimal stopping under uncertainty in drift and jump intensity. *Mathematics of Operations Research* 43, 1177–1209.