Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Optimal stopping with randomly arriving opportunities to stop

Josha A. Dekker[1], Roger J.A. Laeven[1], John G. M. Schoenmakers[2], Michel H. Vellekoop[1]

submitted: November 20, 2023

[1] University of Amsterdam
P.O. Box 15867
1001 NJ Amsterdam
The Netherlands
E-Mail: J.A.Dekker@uva.nl
R.J.A.Laeven@uva.nl
M.H.Vellekoop@uva.nl

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: john.schoenmakers@wias-berlin.de

No. 3056
Berlin 2023

2020 *Mathematics Subject Classification.* 60G40, 65C05, 93E24, 91G60.

Key words and phrases. Optimal stopping on random times, infinite horizon, duality, least squares regression, policy improvement.

This research was funded in part by the Netherlands Organization for Scientific Research (NWO) under grant NWO-Vici (Laeven) and by the DFG Excellence Cluster Math+ Berlin, project AA4-2 (Schoenmakers).

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Optimal stopping with randomly arriving opportunities to stop

Josha A. Dekker, Roger J.A. Laeven, John G. M. Schoenmakers, Michel H. Vellekoop

Abstract

We develop methods to solve general optimal stopping problems with opportunities to stop that arrive randomly. Such problems occur naturally in applications with market frictions. Pivotal to our approach is that our methods operate on random rather than deterministic time scales. This enables us to convert the original problem into an equivalent discrete-time optimal stopping problem with $\mathbb{N}_0$-valued stopping times and a possibly infinite horizon. To numerically solve this problem, we design a random times least squares Monte Carlo method. We also analyze an iterative policy improvement procedure in this setting. We illustrate the efficiency of our methods and the relevance of randomly arriving opportunities in a few examples.

1 Introduction

Since the early work of Wald [25] and Snell [24], a vast literature has contributed to the theory of optimal stopping and stochastic control. Applications of this theory have been analyzed in a wide variety of fields including statistics, operations research, and financial and insurance mathematics.

A standard assumption in this literature is that opportunities to stop arise deterministically — either at pre-specified time instances in a discrete-time (Bermudan) setting or at any time within a pre-specified time interval in a continuous-time (American) setting. In recent years, motivated by market frictions, a few papers have studied optimal stopping problems in which the stopping times are restricted to take values in a random set, typically generated by an independent Poisson process ([8, 18, 17, 12, 13]). The development of numerical methods to solve general optimal stopping problems with randomly arriving opportunities may, however, be considered in its infancy.

In this paper, we develop methods to solve, theoretically and numerically, general optimal stopping problems in which opportunities to stop arrive randomly. The key to our approach is to consider random rather than deterministic time scales on which our methods operate. That is, instead of working on a possibly fine deterministic time grid, our approach consists of proceeding over random time scales from one stochastic opportunity to the next, which brings significant computational advantages.

More specifically, on the new, random time scale that we consider, the order of the arrivals is encoded in the set of natural numbers, rather than designated by the actual arrival times on the positive real line. This enables us to convert the problem into a discrete-time optimal stopping problem with respect to $\mathbb{N}_0$-valued stopping times with filtration generated by the consecutively arriving random times, on a possibly infinite horizon. We establish the theoretical properties of this problem. In order to numerically solve this problem, we adapt the well-known Longstaff-Schwartz [19] method via a suitable truncation of the horizon if necessary. Furthermore, we revisit the policy improvement procedure in [15], which was developed in a finite horizon setting. Since the theory in [15] is entirely based on backward induction from a finite horizon, we generalize their main statements to an infinite horizon setting and give new (essentially different) proofs. In fact, this generalization may be considered of interest in its own right.

This paper is organized as follows. In Section 2, we introduce the setting and notation, and establish pivotal equivalence results. In Section 3, we develop *random times least squares Monte Carlo*. In Section 4, we

DOI 10.20347/WIAS.PREPRINT.3056 Berlin 2023

analyze iterative policy improvement in an infinite horizon setting. Section 5 provides extensive numerical illustrations to assess the accuracy and efficiency of our algorithms and to illustrate the relevance of randomly arriving opportunities for optimal stopping. Conclusions are in Section 6.

2 Optimal Stopping on Random Times

2.1 Setting and notation

We consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with filtration $\mathbb{F} := (\mathcal{F}_t)_{0\leq t\leq\infty}$ satisfying "the usual conditions". We denote by $\mathbb{N}$ the (strictly positive) natural numbers and by $\mathbb{N}_0$ the natural numbers including zero.

We suppose that we are given a random sequence of $\mathbb{F}$-stopping times

$$\tau_0 = 0,\ \tau_i < \tau_{i+1} < \infty \ \text{ for } \ i \in \mathbb{N}_0, \text{ and } \lim_{i\to\infty} \tau_i = \infty \ \text{ a.s.} \tag{2.1}$$

We are also given a nonnegative $\mathbb{F}$-adapted càdlàg (on $[0, T]$) reward process $(Z_t)_{0\leq t<\infty}$. We assume that $Z_t = 0$ a.s. for $t > T$, for a given finite time horizon $T > 0$, and that

$$\mathbb{E}\left[\sup_{i\in\mathbb{N}_0} Z^2_{\tau_i}\right] \leq B^2, \quad \text{for some } B > 0. \tag{2.2}$$

We now consider the problem of optimal stopping of Z on the first K stopping times in the sequence (2.1), where $K \in \mathbb{N} \cup \{\infty\}$. That is, we consider the stopping problem

$$Y_0^{(K)} := \sup_{\mathbb{F}\text{-stopping time } \mathfrak{t},\ \mathfrak{t}\in\mathcal{T}_K} \mathbb{E}\left[Z_{\mathfrak{t}}\right], \tag{2.3}$$

where on each event ω, $\mathcal{T}_K(\omega) := \{\tau_i(\omega) : i \in \mathbb{N}_0,\ i \leq K\}$ denotes the set of realized stopping times (2.1) up to and including number $K \leq \infty$.

2.2 An equivalent discrete-time problem with infinite horizon

We next recast problem (2.3) into an infinite horizon discrete-time stopping problem in which the opportunities to stop emerge as discrete stopping times in the set $\mathbb{N}_0$. Due to the order of the stopping times (2.1), the corresponding stop σ-algebras

$$\mathcal{F}_{\tau_i} := \{A \in \mathcal{F}_\infty : A \cap \{\tau_i \leq t\} \in \mathcal{F}_t,\ \ t \geq 0\}, \tag{2.4}$$

are analogously ordered,

$$\mathcal{F}_{\tau_0} \subset \mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2} \subset \ldots.$$

We thus obtain, with $\mathcal{G}_i := \mathcal{F}_{\tau_i} \subset \mathcal{F}$, $i \in \mathbb{N}_0$, a discrete filtration $\mathbb{G} := (\mathcal{G}_i)_{i\in\mathbb{N}_0}$. Furthermore, the discrete reward process $(Z_{\tau_\mathfrak{n}})_{n\in\mathbb{N}_0}$ is $\mathbb{G}$-adapted (see, e.g., [21, Thm. 6] or [9]). Now let $\mathcal{N}$ be the set of discrete $\mathbb{G}$-stopping times, i.e.,

$$\mathcal{N} := \{\mathfrak{n} : \Omega \to \mathbb{N}_0 \ \text{ such that } \ \{\mathfrak{n} = i\} \in \mathcal{G}_i \text{ for all } i \in \mathbb{N}_0\}.$$

We then state the following proposition:

Proposition 2.1 *Recall* (2.3)*. We have that*

$$Y_0^{(K)} = \sup_{\mathfrak{n}\in\mathcal{N},\,\mathfrak{n}\le K} \mathbb{E}\left[Z_{\tau_\mathfrak{n}}\right]. \tag{2.5}$$

Proof. (i) It holds that $\sup_{\mathfrak{n}\in\mathcal{N},\,\mathfrak{n}\le K} \mathbb{E}\left[Z_{\tau_\mathfrak{n}}\right] \le Y_0^{(K)}$ since, for every fixed $\mathfrak{n}\in\mathcal{N}$ with $\mathfrak{n}\le K$ a.s., we have that $\mathfrak{t}:=\tau_\mathfrak{n}$ is an $\mathbb{F}$-stopping time with $\mathfrak{t}\in\mathcal{T}_K$. Indeed, one trivially has that $\tau_{\mathfrak{n}(\omega)}(\omega)\in\mathcal{T}_K(\omega)$, and moreover,

$$\{\mathfrak{t}\le t\} = \{\tau_\mathfrak{n}\le t\} = \bigcup_{i=0}^{K}\{\tau_\mathfrak{i}\le t\}\cap\{\mathfrak{n}=i\}\in\mathcal{F}_t,$$

since, for any $i\in\mathbb{N}_0$, $\{\tau_i\le t\}\in\mathcal{F}_t$ and $\{\mathfrak{n}=i\}\in\mathcal{F}_{\tau_\mathfrak{i}}$, and so $\{\tau_\mathfrak{i}\le t\}\cap\{\mathfrak{n}=i\}\in\mathcal{F}_t$, due to (2.4).

(ii) Now let $\mathfrak{t}$ be any $\mathbb{F}$-stopping time with a.s. $\mathfrak{t}(\omega)\in\mathcal{T}_K(\omega)$. Define $\mathfrak{n}:\Omega\to\mathbb{N}_0$ by

$$\mathfrak{n}(\omega) = \sum_{i=0}^{K} i 1_{\{\mathfrak{t}(\omega)=\tau_i(\omega)\}}.$$

Then, for any j, since both $\mathfrak{t}$ and τ_j are $\mathbb{F}$-stopping times, it holds that

$$\{\mathfrak{t}(\omega)=\tau_j(\omega)\}\in\mathcal{F}_\mathfrak{t}\cap\mathcal{F}_{\tau_j};$$

see e.g., [14, Lemma 2.16]. Hence,

$$\{\mathfrak{n}(\omega)=j\} = \{\mathfrak{t}(\omega)=\tau_j(\omega)\}\in\mathcal{F}_{\tau_j}=\mathcal{G}_j,$$

i.e., $\mathfrak{n}$ is a discrete $\mathbb{G}$-stopping time with $\mathfrak{n}\le K$ a.s., and moreover

$$\mathbb{E}\left[Z_\mathfrak{t}\right] = \mathbb{E}\left[\sum_{i=1}^{K} Z_{\tau_\mathfrak{i}} 1_{\{\mathfrak{t}=\tau_\mathfrak{i}\}}\right] = \mathbb{E}\left[\sum_{i=1}^{K} Z_{\tau_i} 1_{\{\mathfrak{n}=i\}}\right] = \mathbb{E}\left[Z_{\tau_\mathfrak{n}}\right].$$

Thus, $Y_0^{(K)} \le \sup_{\mathfrak{n}\in\mathcal{N},\,\mathfrak{n}\le K}\mathbb{E}\left[Z_{\tau_\mathfrak{n}}\right]$. ■

In order to study the stopping problem (2.3), we introduce, in view of Proposition 2.1, the discrete-time reward process $(U_i)_{i\in\mathbb{N}_0}$ with

$$U_i := Z_{\tau_i}, \quad i\in\mathbb{N}_0, \tag{2.6}$$

which is $\mathcal{G}$-adapted and satisfies

$$\mathbb{E}\left[\sup_{i\in\mathbb{N}_0} U_i^2\right] \le B^2, \tag{2.7}$$

due to (2.2). We now consider the discrete stopping problem as seen from a generic point in discrete time $i\in\mathbb{N}_0$,

$$Y_i^{(K)} := \operatorname*{ess\,sup}_{\mathfrak{n}\in\mathcal{N},\,i\le\mathfrak{n}\le K} \mathbb{E}_{\mathcal{G}_i}\left[U_\mathfrak{n}\right], \quad i\in\mathbb{N}_0. \tag{2.8}$$

Remark 2.2 *In practice, the bound* (2.2)*, hence* (2.7)*, may be determined by simulation. For an analytic alternative, let us observe that condition* (2.2) *is implied by the somewhat stronger condition*

$$\mathbb{E}\left[\sup_{0\le t\le T} Z_t^2\right] \le B^2, \quad \text{for some } B>0.$$

In the case in which the reward can be written in the form $Z_t = g(t, M_t)$ for some d-dimensional martingale M, and the function g satisfies

$$|g(t, M_t)| \leq c_1 \|M_t\|_\infty^q + c_2, \quad 0 \leq t \leq T, \quad q \geq 1,$$

it follows that

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\leq t\leq T} Z_t^2\right] &\leq \mathbb{E}\left[\sup_{0\leq t\leq T} \left(c_1 \|M_t\|_\infty^q + c_2\right)^2\right] \leq \mathbb{E}\left[\left(c_1 \sup_{0\leq t\leq T} \|M_t\|_\infty^q + c_2\right)^2\right] \\
&\leq 2c_1^2 \mathbb{E}\left[\sup_{0\leq t\leq T} \|M_t\|_\infty^{2q}\right] + 2c_2^2 \\
\text{(with } M = (M^i)_{i=1,\ldots,d}\text{)} &\leq 2c_1^2 \sum_{i=1}^d \mathbb{E}\left[\sup_{0\leq t\leq T} \left|M_t^i\right|^{2q}\right] + 2c_2^2 \\
&\leq 2c_1^2 C_{2q}^{BDG} \sum_{i=1}^d \mathbb{E}\left[\langle M_T^i, M_T^i\rangle^q\right] + 2c_2^2,
\end{aligned}$$

using the Burkholder-Davis-Gundy inequality. The constant C_{2q}^{BDG} is universal and explicitly known for continuous martingales and jump martingales, respectively. Further, if $q = 1$ for example, $\mathbb{E}\left[\langle M_T^i, M_T^i\rangle\right]$ may usually be estimated explicitly for a given specific model.

2.3 Bellman principle and optimal stopping time

The ultimate goal in problem (2.8) is finding the value of $Y_0^{(K)}$. For running i, $Y_i^{(K)}$ is called the Snell envelope for the reward process $(U_i)_{i\in\mathbb{N}_0}$. By

$$\begin{aligned}
C_i^{(K)} &:= \operatorname*{ess\,sup}_{\mathfrak{n}\in\mathcal{N},\, i+1\leq\mathfrak{n}\leq K} \mathbb{E}_{\mathcal{G}_i}\left[U_\mathfrak{n}\right], \quad i \in \mathbb{N}_0, \quad 0 \leq i < K, \\
C_K^{(K)} &= 0 \quad \text{if } K < \infty,
\end{aligned} \tag{2.9}$$

we denote the discrete-time continuation value process. The following proposition collects a few facts about optimal stopping in discrete time with finite time horizon $K < \infty$, or with infinite time horizon $K = \infty$.

Proposition 2.3 *In our present setting, in particular under assumption* (2.2)*, we have for $K \in \mathbb{N} \cup \{\infty\}$ and for all $i \in \mathbb{N}_0$ with $0 \leq i \leq K$ that:*

(i) The continuation value process (2.9) *is connected with* (2.8) *via the Bellman principle*

$$Y_i^{(K)} = \max\left(U_i, C_i^{(K)}\right).$$

(ii) The process $Y^{(K)}$ is a $\mathcal{G}$-supermartingale, i.e.,

$$Y_i^{(K)} \geq \mathbb{E}_{\mathcal{G}_i}\left[Y_{i+1}^{(K)}\right], \tag{2.10}$$

and it is the smallest supermartingale that dominates U.

(iii) The random time $\mathfrak{n}_i^{K,*}$, defined by

$$\mathfrak{n}_i^{K,*} := \inf \left\{ i \leq j \leq K : \; Y_j^{(K)} = U_j \right\},$$

is an optimal stopping time for $\mathcal{G}$, that is,

$$Y_i^{(K)} = \mathbb{E}_{\mathcal{G}_i} \left[U_{\mathfrak{n}_i^{K,*}} \right].$$

Proof. We note that (2.7), which follows from assumption (2.2), implies that condition [20, Ch. 1, (1.1.3))] is satisfied and as a consequence for details and proof of the stated results one may consult [20, Ch. 1]. ■

Corollary 2.4 *There also exists an optimal stopping time for Problem* (2.3)*, which moreover may be obtained by taking $\mathfrak{t}^* := \mathfrak{t}_0^{K,*} := \tau_{\mathfrak{n}^*} := \tau_{\mathfrak{n}_0^{K,*}}$.*

Proof. Indeed, from item (i) in the proof of Proposition 2.1, we see that $\mathfrak{t}^*$ is an $\mathbb{F}$-stopping time which satisfies $\mathfrak{t}^* \in \mathcal{T}_K$ and which moreover satisfies

$$Y_0^{(K)} = \sup_{\mathfrak{n} \in \mathcal{N}, \, \mathfrak{n} \leq K} \mathbb{E}\left[Z_{\tau_\mathfrak{n}} \right] = \mathbb{E}\left[Z_{\tau_{\mathfrak{n}^*}} \right] = \mathbb{E}\left[Z_{\mathfrak{t}^*} \right].$$

■

2.4 Duality

In this subsection, we concisely recall the dual representation of the discrete optimal stopping problem in our present setting, which essentially goes back to [22] and [11] (see also the early [7]). We consider here the infinite horizon case, i.e., we take $K = \infty$ in Proposition 2.1. The case $K < \infty$ goes analogously, and moreover does not require an extra uniform integrability argument (see Remark 2.6 below).

Proposition 2.5 *Let $\mathcal{M}_0^{\mathsf{UI}}$ be the collection of uniformly integrable $\mathbb{G}$-martingales $(M_i)_{i \in \mathbb{N}_0}$ with $M_0 = 0$. In addition, assume that* (2.2) *holds and that there exists some $0 < \alpha < 1/2$ such that*

$$\sum_{j=1}^{\infty} \mathbb{P}\left(\tau_j \leq T \right)^{\alpha} < \infty. \tag{2.11}$$

We then have the following duality results.

(i) Weak duality: It holds that

$$Y_0 \equiv Y_0^{(\infty)} = \inf_{M \in \mathcal{M}_0^{\mathsf{UI}}} \mathbb{E}\left[\sup_{i \geq 0} \left(Z_{\tau_i} - M_i \right) \right].$$

(ii) Strong duality: Let us consider, for $Y_i \equiv Y_i^{(\infty)}$ given in (2.8)*, the $\mathcal{G}$-martingale, a.k.a. the Doob martingale of the Snell envelope,*

$$M_i^\circ := \sum_{j=1}^{i} Y_j - \mathbb{E}_{\mathcal{G}_{j-1}}\left[Y_j \right], \; i \geq 1, \quad M_0^\circ = 0.$$

Then it holds that $M^\circ \in \mathcal{M}_0^{\mathsf{UI}}$ and that

$$Y_0 = \sup_{i \geq 0} \left(U_i - M_i^\circ \right), \quad \textit{almost surely.}$$

Proof. By Proposition 2.1 and Doob's sampling theorem we have, for any $M \in \mathcal{M}_0^{\mathsf{UI}}$,

$$\begin{aligned} Y_0 &= \sup_{\mathfrak{n}\in\mathcal{N}} \mathbb{E}\left[U_{\mathfrak{n}}\right] = \sup_{\mathfrak{n}\in\mathcal{N}} \mathbb{E}\left[U_{\mathfrak{n}} - M_{\mathfrak{n}}\right] \\ &\leq \mathbb{E}\left[\sup_{i\geq 0}\left(U_i - M_i\right)\right] = \mathbb{E}\left[\sup_{i\geq 0}\left(Z_{\tau_i} - M_i\right)\right]. \end{aligned} \tag{2.12}$$

Note that for any $i \geq 0$ (empty sums being zero),

$$\begin{aligned} U_i - M_i^{\circ} &= U_i - \sum_{j=1}^{i} Y_j + \sum_{j=1}^{i} \mathbb{E}_{\mathcal{G}_{j-1}}\left[Y_j\right] \qquad (2.13) \\ &= Y_0 + U_i - Y_i - \sum_{j=0}^{i-1} Y_j + \sum_{j=0}^{i-1} \mathbb{E}_{\mathcal{G}_j}\left[Y_{j+1}\right] \\ &\leq Y_0, \end{aligned}$$

due to (2.10). Hence, we have

$$\sup_{i\geq 0}\left(U_i - M_i^{\circ}\right) \leq Y_0, \tag{2.14}$$

by (2.13). We now only need to show that $M^{\circ} \in \mathcal{M}_0^{\mathsf{UI}}$. Indeed, if $M^{\circ} \in \mathcal{M}_0^{\mathsf{UI}}$, we have

$$Y_0 \leq \mathbb{E}\left[\sup_{i\geq 0}\left(U_i - M_i^{\circ}\right)\right],$$

by inequality (2.12), and then by (2.14) we obtain statement (ii) due to the sandwich property. Finally, (ii) combined with (2.12) yields statement (i). Let us show that $M^{\circ} \in \mathcal{M}_0^{\mathsf{UI}}$. By the well-known lemma of de la Vallée-Poussin, it is enough to show that

$$\sup_{i\geq 0} \mathbb{E}\left[\left|M_i^{\circ}\right|^p\right] < \infty,$$

for some $p > 1$. Let us take $p = 2/(2\alpha+1)$, hence $1 < p < 2$. Since $Z_t \geq 0$ for $t \geq 0$, and $Z_t = 0$ for $t > T$, we have that

$$\begin{aligned} M_i^{\circ} &= \sum_{j=1}^{i}\left(Y_j - \mathbb{E}_{\mathcal{G}_{j-1}}\left[Y_j\right]\right) 1_{\{\tau_j < T\}}, \quad \text{hence} \\ \text{(Minkowski) } \mathbb{E}\left[\left|M_i^{\circ}\right|^p\right]^{1/p} &\leq \sum_{j=1}^{i} \mathbb{E}\left[\left(Y_j + \mathbb{E}_{\mathcal{G}_{j-1}}\left[Y_j\right]\right)^p 1_{\{\tau_j < T\}}\right]^{1/p}. \end{aligned} \tag{2.15}$$

Furthermore, we have, since also $1 < 2/p < 2$,

$$\begin{aligned} &\mathbb{E}\left[\left(Y_j + \mathbb{E}_{\mathcal{G}_{j-1}}\left[Y_j\right]\right)^p 1_{\{\tau_j \leq T\}}\right] \\ \text{(convexity) } &\leq 2^{p-1}\mathbb{E}\left[\left(Y_j^p + \mathbb{E}_{\mathcal{G}_{j-1}}\left[Y_j\right]^p\right) 1_{\{\tau_j \leq T\}}\right] \\ \text{(Jensen) } &\leq 2^{p-1}\mathbb{E}\left[\left(Y_j^p + \mathbb{E}_{\mathcal{G}_{j-1}}\left[Y_j^p\right]\right) 1_{\{\tau_j \leq T\}}\right] \\ \text{(Hölder) } &\leq 2^{p-1}\mathbb{E}\left[\left(Y_j^p + \mathbb{E}_{\mathcal{G}_{j-1}}\left[Y_j^p\right]\right)^{2/p}\right]^{p/2} \mathbb{E}\left[1_{\{\tau_j \leq T\}}\right]^{1-p/2}, \end{aligned} \tag{2.16}$$

and

$$\begin{aligned} \mathbb{E}\left[\left(Y_j^p + \mathbb{E}_{\mathcal{G}_{j-1}}\left[Y_j^p\right]\right)^{2/p}\right] &\leq 2^{2/p-1}\mathbb{E}\left[\left(Y_j^2 + \mathbb{E}_{\mathcal{G}_{j-1}}\left[Y_j^p\right]^{2/p}\right)\right] \\ &\leq 2^{2/p}\mathbb{E}\left[Y_j^2\right] \leq 2^{2/p}\mathbb{E}\left[\sup_{i\in\mathbb{N}_0} Z_{\tau_i}^2\right] \leq 2^{2/p}B^2, \end{aligned} \tag{2.17}$$

due to (2.2). With (2.15), (2.16), (2.17) it follows that for any $i \geq 0$,

$$\mathbb{E}\left[|M_i^\circ|^p\right]^{1/p} \leq \sum_{j=1}^{i} 2B\mathbb{P}\left(\tau_j \leq T\right)^{\frac{1}{p}-\frac{1}{2}} = \sum_{j=1}^{i} 2B\mathbb{P}\left(\tau_j \leq T\right)^{\alpha} \leq 2B\sum_{j=1}^{\infty} \mathbb{P}\left(\tau_j \leq T\right)^{\alpha}.$$

■

Remark 2.6 *(a) For $K < \infty$ and $Y_i \equiv Y_i^{(K)}$, we may replace in the statements (i) and (ii) of Proposition 2.5, $\sup_{i\geq 0}$ with $\max_{0\leq i\leq K}$. Moreover, the condition* (2.11) *can be removed, since any integrable martingale $(M_i)_{1\leq i\leq K}$ is uniformly integrable.*

(b) If $K = \infty$ but one has for some $L \in \mathbb{N}$ that $\mathbb{P}\left(\tau_L > T\right) = 1$, condition (2.11) *is trivially fulfilled. However, we are then dealing with problem* (2.3) *for $K = L$ in fact.*

(c) It should be noted that the condition (2.11) *for uniform integrability of M° is a sufficient condition. It requires that, loosely speaking, the probability that τ_j did (still) not pass T at the stopping time with rank j converges to zero fast enough, when j tends to infinity.*

(d) More formally, condition (2.11) *holds if and only if the distribution of $N \equiv N_T := \sup\{j : \tau_j \leq T\}$ has tails that decay faster than j^{-2}. Indeed, the series*

$$\sum_{j=1}^{\infty} \mathbb{P}\left(\tau_j \leq T\right)^{\alpha} = \sum_{j=1}^{\infty} \mathbb{P}\left(N \geq j\right)^{\alpha} \sim \sum_{j=1}^{\infty} j^{-\gamma\alpha},$$

with γ being the tail exponent, converges for $\gamma\alpha > 1$ and diverges for $\gamma\alpha \leq 1$. Hence, when $\gamma > 2$, we may consider $1/\gamma < \alpha < 1/2$. Conversely, for $0 < \alpha < 1/2$, we must have $\gamma > 2$.

2.5 Structural assumption to facilitate numerical approaches

To develop feasible simulation-based numerical approaches, we consider a basic structural assumption on our setup. Recall that the (augmented) filtration $\mathbb{F}$ describes the information flow and that Z determines the reward upon stopping at one of the $\mathbb{F}$-stopping times $(\tau_k)_{k\in\mathbb{N}_0}$, with $\tau_0 = 0$ and $\lim_{k\to\infty} \tau_k = \infty$ a.s. Henceforth, we assume that there is an underlying process X such that $Z = Z(\cdot, X_\cdot)$. The dynamics of X may also depend on another, auxiliary process Θ that does not directly influence the reward process Z.

Assumption 2.7 *There exists an auxiliary càdlàg process Θ in $\mathbb{R}^q$ such that the process (X, Θ) in $\mathbb{R}^d \times \mathbb{R}^q$ is adapted to the (augmented) filtration $\mathbb{F}$ and is such that the $\mathbb{G}$-adapted discrete-time process*

$$(\tau_k, X_{\tau_k}, \Theta_{\tau_k})_{k\in\mathbb{N}_0}, \quad (\tau_0, X_{\tau_0}, \Theta_{\tau_0}) = (0, X_0, \Theta_0),$$

is a Markov chain in the state space $\mathbb{R}_{\geq 0} \times \mathbb{R}^d \times \mathbb{R}^q$.

The interpretation of Assumption 2.7 is as follows. We think of (X, Θ) as an observable process that may generate the filtration $\mathbb{F}$, whereas X is the underlying process that determines the reward Z_{τ_i} upon stopping at time τ_i. Assumption 2.7 paves the way for simulation-based numerical approaches. The following results are readily obtained.

Lemma 2.8 *If Assumption 2.7 is fulfilled, then for any non-negative Borel measurable $g : (t, x, \theta) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^q \to g(t, x, \theta)$, we have that*

$$\begin{aligned}\mathbb{E}_{\mathcal{F}_{\tau_k}}\left[g(\tau_{k+1}, X_{\tau_{k+1}}, \Theta_{\tau_{k+1}})\right] &= \mathbb{E}_{(\tau_k, X_{\tau_k}, \Theta_{\tau_k})}\left[g(\tau_{k+1}, X_{\tau_{k+1}}, \Theta_{\tau_{k+1}})\right] \\ &= c(\tau_k, X_{\tau_k}, \Theta_{\tau_k}),\end{aligned}$$

for some non-negative Borel function $c(\cdot, \cdot, \cdot)$.

Lemma 2.9 *Suppose that Assumption 2.7 applies, X and Θ are independent, and $\tau_{k+1}-\tau_k$ is independent of Θ_{τ_k}. Then we have that*

$$\mathbb{E}_{\mathcal{F}_{\tau_k}}\left[g(\tau_{k+1}, X_{\tau_{k+1}})\right] = c(\tau_k, X_{\tau_k}),$$

for some non-negative Borel function $c(\cdot,\cdot)$.

Proof. By Assumption 2.7,

$$\begin{aligned}\mathbb{E}_{\mathcal{F}_{\tau_k}}\left[g(\tau_{k+1}, X_{\tau_{k+1}})\right] &= \mathbb{E}_{\left(\tau_k, X_{\tau_k}, \Theta_{\tau_k}\right)}\left[g(\tau_k+\tau_{k+1}-\tau_k, X_{\tau_k+\tau_{k+1}-\tau_k})\right] \\ &= \mathbb{E}_{\left(\tau_k, X_{\tau_k}\right)}\left[g(\tau_{k+1}, X_{\tau_{k+1}})\right] =: c(\tau_k, X_{\tau_k}).\end{aligned}$$

■

Example 2.10 *Take X independent of Θ, where Θ is a continuous-time Markov chain with state space $\mathbb{N}_0$, generator*

$$\begin{bmatrix} -\lambda_1 & \lambda_1 & 0 & \dots \\ 0 & -\lambda_2 & \lambda_2 & \dots \\ 0 & 0 & -\lambda_3 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}, \quad \lambda_i > 0, \ \ i = 1, 2, \dots,$$

and $\Theta_0 = 0$. One thus has that $\tau_k = \inf\{s \geq 0 : \Theta_s = k\}$.

Example 2.11 *As a more general example, consider (X, Θ) to be given by a jump-diffusion, more specifically, a strong Markov process of the form*

$$\begin{aligned}\mathrm{d}X_t &= \sigma(t, X_t) X_t \,\mathrm{d}W, \quad X_0 = x_0, \\ \mathrm{d}\Theta_t &= \int_{\mathbb{R}^q} z\mathcal{N}\left(X_t, \,\mathrm{d}t, \,\mathrm{d}z\right), \quad \Theta_0 = 0,\end{aligned}$$

where $\mathcal{N}\left(x, \,\mathrm{d}t, \,\mathrm{d}z, \omega\right)$ is some Poisson random measure on $\mathbb{N}_0$, independent of W, with

$$\mathbb{P}\left(\mathcal{N}\left(x, (s,t], B\right) = k\right) := \exp(-v(x,B)\,(t-s))\frac{v(x,B)^k\,(t-s)^k}{k!}, \quad k \in \mathbb{N}_0,$$

where for simplicity $v(x, B)$ is of finite activity in the sense that

$$v(x, \{0\}) = 0 \ \text{ and } \ \int v(x, \,\mathrm{d}z) < \infty, \quad x \in \mathbb{R}^d.$$

The stopping times are determined by $\tau_0 = 0$ and $\tau_k = \inf\{s > 0 : \Theta_s = k\}$, $k > 0$.

(i): Simple case with $d = q = 1$ and $v(x, B) = \lambda\delta_1(B) := \lambda 1_B(1)$: We then have that

$$\mathbb{P}\left(\mathcal{N}\left(x, (0,t], B\right) = k\right) = \exp\left(-\lambda t \delta_1(B)\right)\frac{\lambda^k t^k \delta_1(B)^k}{k!},$$

with $0^0 := 1$, and hence we get

$$\begin{aligned}\Theta_t &= \int_{(0,t]}\int_{\mathbb{R}} z\mathcal{N}\left(\,\mathrm{d}s, \,\mathrm{d}z\right) = \mathcal{N}\left((0,t], \{1\}\right), \quad \text{where} \\ &\mathbb{P}\left(\mathcal{N}\left((0,t], \{1\}\right) = k\right) = \exp\left(-\lambda t\right)\frac{\lambda^k t^k}{k!},\end{aligned}$$

i.e., $\Theta_t = \sharp$ *jumps in the interval* $(0,t]$ *observed by the agent. With* $\tau_0 = 0$ *and* $\tau_k = \inf\{s > 0 : \Theta_s = k\}$, $k > 0$, *one has that* $\tau_k - \tau_{k-1}$, $k \geq 1$, *is exp(*λ*) distributed and independent of* τ_{k-1}. *In particular, the conclusion of Lemma 2.8 applies.*

(ii): Suppose $d = q = 1$ *and that the Poisson random measure is only active during times that the process* X *is visiting a certain interval* $I^{\pm} := [c_-, c_+]$ *with* $0 < c_- < c_+$. *That is, we consider*

$$v(x,B) := \lambda 1_{[c_-,c_+]}(x)\delta_1(B), \quad \textit{hence}$$
$$\mathbb{P}\left(\mathcal{N}\left(x,(0,t],B\right) = k\right) = \exp\left(-\lambda t 1_{[c_-,c_+]}(x)\delta_1(B)\right)\frac{\lambda^k t^k 1_{[c_-,c_+]}(x)^k \delta_1(B)^k}{k!}$$

with $0^0 := 1$, *and*

$$\begin{aligned}\Theta_t &= \int_{(0,t]} 1_{[c_-,c_+]}(X_{s-}) \int_{\mathbb{R}} z\mathcal{N}\left(X_{s-},\, \mathrm{d}s,\, \mathrm{d}z\right)\\ &= \int_{(0,t]} 1_{[c_-,c_+]}(X_s)\mathcal{N}_0\left(\,\mathrm{d}s, \{1\}\right),\end{aligned}$$

with $\mathcal{N}_0$ *as in (i). Note that* X *is continuous. So* $\Theta_t = \sharp$ *jumps in the interval* $(0,t]$ *while* X *is in the interval* $[c_-, c_+]$, *which is observable by the agent. Let* $\tau_0 = 0$, *and* $\tau_k = \inf\{s > 0 : \Theta_s = k\}$, $k > 0$. *Then, obviously,* $\Theta_{\tau_k} = k$, $k \geq 0$, *and*

$$\begin{aligned}\tau_{k+1} - \tau_k &\in \sigma\left\{\left(X_s, \mathcal{N}_0\left(\left(\tau_k, s\right], \{1\}\right)\right) 1_{\{s\geq\tau_k\}}\right\}\\ &\subset \sigma\left\{\left(X_s, \Theta_s\right) 1_{\{s\geq\tau_k\}}\right\}.\end{aligned}$$

Hence, in particular Assumption 2.7 applies, i.e., by the Markov property,

$$\begin{aligned}&\mathbb{E}_{\mathcal{F}_{\tau_k}}\left[g(\tau_{k+1}, X_{\tau_{k+1}}, \Theta_{\tau_{k+1}})\right]\\ &= \mathbb{E}_{\mathcal{F}_{\tau_k}}\left[g(\tau_k + \tau_{k+1} - \tau_k, X_{\tau_k+\tau_{k+1}-\tau_k}, \Theta_{\tau_k+\tau_{k+1}-\tau_k})\right]\\ &= \mathbb{E}_{\mathcal{F}_{\tau_k}}\left[g(\tau_k + \tau_{k+1} - \tau_k, X_{\tau_k+\tau_{k+1}-\tau_k}, \Theta_{\tau_k+\tau_{k+1}-\tau_k})\right]\\ &= \mathbb{E}_{\left(\tau_k, X_{\tau_k}, \Theta_{\tau_k}\right)}\left[g(\tau_k + \tau_{k+1} - \tau_k, X_{\tau_k+\tau_{k+1}-\tau_k}, \Theta_{\tau_k+\tau_{k+1}-\tau_k})\right]\\ &= \mathbb{E}_{\left(\tau_k, X_{\tau_k}, \Theta_{\tau_k}\right)}\left[g(\tau_{k+1}, X_{\tau_{k+1}}, \Theta_{\tau_{k+1}})\right]\\ &=: c(\tau_k, X_{\tau_k}, \Theta_{\tau_k}),\end{aligned}$$

and so this is a simple example in which the continuation functions are dependent on the evaluation of both X *and* Θ.

3 Least Squares Monte Carlo over Random Times

3.1 Finite horizon approximation

Any backward recursive approach such as the Longstaff-Schwartz (LS, [19]) method is initialized at some (discrete) finite time horizon. In case $K = \infty$, we have an infinite horizon problem due to Proposition 2.1. Thus, in this case, a backward recursive approach is not directly applicable. However, as we show below, we may approximate the infinite horizon problem by a finite horizon counterpart with arbitrary accuracy, and then, if additional structural assumptions are fulfilled (see Section 2.5), apply the LS method to the finite horizon problem.

Since, by (2.1), $\tau_k \to \infty$ a.s. if $k \to \infty$, we have by dominated convergence that

$$\mathbb{P}(\tau_K \leq T) \to 0, \quad K \to \infty, \quad \text{for } K \in \mathbb{N}.$$

Let us fix some tolerance level $\varepsilon > 0$ and choose K such that $\mathbb{P}(\tau_K \leq T) < \varepsilon^2/B^2$ with B specified in (2.2). The value of the infinite horizon stopping problem is denoted by $Y_0^{(\infty)}$, and the value of the corresponding "truncated" stopping problem is given by

$$Y_0^{(K)} = \sup_{\mathfrak{n}\in\mathcal{N},\, 0\leq\mathfrak{n}\leq K} \mathbb{E}\left[U_{\mathfrak{n}}\right].$$

Obviously, one has $Y_0^{(K)} \leq Y_0^{(\infty)}$. On the other hand,

$$\begin{aligned} Y_0^{(\infty)} &= \sup_{\mathfrak{n}\in\mathcal{N}} \mathbb{E}\left[U_{\mathfrak{n}} 1_{\{\mathfrak{n}\leq K\}} + U_{\mathfrak{n}} 1_{\{\mathfrak{n}>K\}}\right] \\ &\leq \sup_{\mathfrak{n}\in\mathcal{N}} \mathbb{E}\left[U_{\mathfrak{n}\wedge K} 1_{\{\mathfrak{n}\leq K\}}\right] + \sup_{\mathfrak{n}\in\mathcal{N}} \mathbb{E}\left[U_{\mathfrak{n}\vee K} 1_{\{\mathfrak{n}>K\}}\right] \\ &\leq \sup_{\mathfrak{n}\in\mathcal{N},\, 0\leq\mathfrak{n}\leq K} \mathbb{E}\left[U_{\mathfrak{n}}\right] + \sup_{\mathfrak{n}\in\mathcal{N},\, \mathfrak{n}>K} \mathbb{E}\left[U_{\mathfrak{n}}\right] \\ &\leq Y_0^{(K)} + \varepsilon, \end{aligned}$$

since for any $\mathfrak{n}$ with $\mathfrak{n} > K$ one has,

$$\begin{aligned} \mathbb{E}\left[U_{\mathfrak{n}}\right] &= \mathbb{E}\left[Z_{\tau_{\mathfrak{n}}} 1_{\{\tau_{\mathfrak{n}}\leq T\}}\right] + \mathbb{E}\left[Z_{\tau_{\mathfrak{n}}} 1_{\{\tau_{\mathfrak{n}}>T\}}\right] \\ &\leq \mathbb{E}\left[Z_{\tau_{\mathfrak{n}}} 1_{\{\tau_K\leq T\}}\right] + 0 \\ &\leq \mathbb{E}\left[1_{\{\tau_K\leq T\}} \sup_{i\in\mathbb{N}_0} Z_{\tau_i}\right] \leq B\mathbb{P}\left(\tau_K \leq T\right)^{1/2} < \varepsilon, \end{aligned} \tag{3.1}$$

using (2.2). Thus, for any pre-specified accuracy ε, $Y_0^{(K)}$ yields a lower bound that is "ε-close" to $Y_0^{(\infty)}$.

3.2 Pseudo-algorithm

Let us suppose we are in a Markovian environment $(X_t, \Theta_t)_{t\geq 0}$, under Assumption 2.7, with a reward of the form

$$Z_t := Z(t, X_t).$$

Then, $Y_0^{(K)}$ may be numerically approximated by regression, based on a given set of measurable basis functions

$$\psi_l : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}, \quad l \in \mathbb{N}. \tag{3.2}$$

We propose the following pseudo-algorithm for $n = 1, \ldots, N$:

(i) Simulate $\left(X^{(n)}_{\tau_1^{(n)}}, \Theta^{(n)}_{\tau_1^{(n)}}\right), \ldots, \left(X^{(n)}_{\tau_K^{(n)}}, \Theta^{(n)}_{\tau_K^{(n)}}\right)$.

(ii) Initialize $\widehat{Y}_K^{K,n} = U_K^{(n)} = Z(\tau_K^{(n)}, X^{(n)}_{\tau_K^{(n)}})$, $\widehat{\mathfrak{n}}_K^{(n)} = K$.

(iii) If, for $0 < k \leq K$, the values $\widehat{Y}_k^{K,n}$ and $\widehat{\mathfrak{n}}_k^{(n)}$ have been constructed, then solve the regression problem

$$(\widehat{\alpha}_{k-1,l})_{l=1,\ldots,L} := \arg\min_{\alpha\in\mathbb{R}^L} \sum_{n=1}^{N} \left(\widehat{Y}_k^{K,n} - \sum_{l=1}^{L} \alpha_l \psi_l\left(\tau_{k-1}^{(n)}, X^{(n)}_{\tau_{k-1}^{(n)}}, \Theta^{(n)}_{\tau_{k-1}^{(n)}}\right)\right)^2,$$

and define

$$\widehat{C}_{k-1}(t,x,\theta) := \sum_{l=1}^{L} \widehat{\alpha}_{k-1,l}\psi_l\left(t,x,\theta\right).$$

Next, set

$$\widehat{\mathfrak{n}}_{k-1}^{(n)} = \begin{cases} k-1 & \text{if} \quad Z(\tau_{k-1}^{(n)}, X_{\tau_{k-1}^{(n)}}^{(n)}) > \widehat{C}_{k-1}(\tau_{k-1}^{(n)}, X_{\tau_{k-1}^{(n)}}^{(n)}, \Theta_{\tau_{k-1}^{(n)}}^{(n)}) \\ \widehat{\mathfrak{n}}_k^{(n)} & \text{else.} \end{cases}$$

Having thus constructed the functions $\widehat{C}_0(t,x,\theta),\ldots,\widehat{C}_{K-1}(t,x,\theta)$ (with $\widehat{C}_K \equiv 0$), one may generate an independent simulation

$$\left(\widetilde{\tau}_1^{(n)}, \widetilde{X}_{\widetilde{\tau}_1^{(n)}}^{(n)}, \widetilde{\Theta}_{\widetilde{\tau}_1^{(n)}}^{(n)}\right),\ldots,\left(\widetilde{\tau}_K^{(n)}, \widetilde{X}_{\widetilde{\tau}_K^{(n)}}^{(n)}, \widetilde{\Theta}_{\widetilde{\tau}_K^{(n)}}^{(n)}\right), \quad \text{for} \quad n=1,\ldots,\widetilde{N},$$

and construct a lower biased estimate $\widetilde{Y}_0$ for Y_0 due to the (path-wise) policy

$$\widetilde{\mathfrak{n}}^{(n)} := \min\left\{k : Z(\widetilde{\tau}_k^{(n)}, \widetilde{X}_{\widetilde{\tau}_k^{(n)}}^{(n)}) \geq \widehat{C}_k(\widetilde{\tau}_k^{(n)}, \widetilde{X}_{\widetilde{\tau}_k^{(n)}}^{(n)}, \widetilde{\Theta}_{\widetilde{\tau}_k^{(n)}}^{(n)})\right\}, \tag{3.3}$$

by defining

$$\widetilde{Y}_0 := \frac{1}{\widetilde{N}}\sum_{n=1}^{\widetilde{N}} Z(\widetilde{\tau}_{\widetilde{\mathfrak{n}}^{(n)}}^{(n)}, \widetilde{X}_{\widetilde{\tau}_{\widetilde{\mathfrak{n}}^{(n)}}^{(n)}}^{(n)}).$$

The policy (3.3) can and will be used (in Section 5) as input policy for the well-known Andersen-Broadie algorithm [2], in order to obtain a dual estimate of an upper bound on Y_0 (in mean). As the Andersen-Broadie method for computing dual upper bounds is widely known, we refrain from a description here and refer to [2], and also [10], for further details.

3.3 Convergence

To obtain convergence results for the algorithm we propose, we need to formulate assumptions for the set of basis functions that are used in our regression, in addition to the structural assumption on the state dynamics formulated in Assumption 2.7.

Assumption 3.1 *For the collection of measurable functions in* (3.2)*, let* P_k^L *denote the* L^2*-projection of* $\sigma\left\{(\tau_k, X_{\tau_k}, \Theta_{\tau_k})\right\}$*-measurable random variables onto the subspace*

$$\mathcal{S}_L = \overline{\text{span}\left\{\psi_l\left(\tau_k, X_{\tau_k}, \Theta_{\tau_k}\right), \quad l=1,\ldots,L\right\}}.$$

We assume the following. For each $k=1,\ldots,K-1$*:*

1. *The sequence* $\left(\psi_l\left(\tau_k, X_{\tau_k}, \Theta_{\tau_k}\right)\right)_{l\in\mathbb{N}}$ *is total in* $L^2\left(\Omega, \sigma\left\{(\tau_k, X_{\tau_k}, \Theta_{\tau_k})\right\}, \mathbb{P}\right)$*, i.e.,* $\mathcal{S}_\infty = L^2\left(\Omega, \sigma\left\{(\tau_k, X_{\tau_k}, \Theta_{\tau_k})\right\}, \mathbb{P}\right)$*.*
2. *For* $L\in\mathbb{N}$*, if* $\sum_{l=1}^L \lambda_l\psi_l\left(\tau_k, X_{\tau_k}, \Theta_{\tau_k}\right) = 0$ *a.s., then* $\lambda_l = 0$ *for* $l=1,\ldots L$*.*

DOI 10.20347/WIAS.PREPRINT.3056

3. *We have*

$$\mathbb{P}\left(\sum_{l=1}^{L} \alpha_{k,l}\psi_l\left(\tau_k, X_{\tau_k}, \Theta_{\tau_k}\right) = Z\left(\tau_k, X_{\tau_k}\right)\right) = 0,$$

where

$$\sum_{l=1}^{L} \alpha_{k,l}\psi_l\left(\tau_k, X_{\tau_k}, \Theta_{\tau_k}\right) = P_k^L\left(Z(\tau_{\mathfrak{n}_k^L}, X_{\tau_{\mathfrak{n}_k^L}})\right),$$

and $\left(\mathfrak{n}_k^L\right)_{k=0,\ldots,K}$ *is a sequence of discrete stopping times with values in* $\{1,\ldots,K\}$ *solving the recursion* $\mathfrak{n}_K^L = K$,

$$\mathfrak{n}_k^L = \begin{cases} k \quad \text{if} \quad Z\left(\tau_k, X_{\tau_k}\right) \geq P_k^L\left(Z(\tau_{\mathfrak{n}_k^L}, X_{\tau_{\mathfrak{n}_k^L}})\right) \\ \mathfrak{n}_{k+1}^L \quad \text{if} \quad Z_k < P_k^L\left(Z(\tau_{\mathfrak{n}_k^L}, X_{\tau_{\mathfrak{n}_k^L}})\right) \end{cases}, \quad 1 \leq k < K.$$

Under Assumptions 2.7 and 3.1, we have the following convergence result.

Theorem 3.2 *([6, Theorem 3.2]) Let* $\mathfrak{n}_k^*$ *be an optimal stopping index after the (physical) times* $\tau_1,\ldots,\tau_{k-1}$ *have passed. It then holds almost surely that*

$$\lim_{L\to\infty}\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} Z(\widehat{\tau}^{(n)}_{\widehat{\mathfrak{n}}_k^{L,n}}, X^{(n)}_{\widehat{\tau}^{(n)}_{\widehat{\mathfrak{n}}_k^{L,n}}}) = \mathbb{E}\left[Z(\tau_{\mathfrak{n}_k^*}, X_{\tau_{\mathfrak{n}_k^*}})\right].$$

[6] further implies that one may obtain a sequence $\widehat{C}_k^{(r)}(t,x,\theta),\ r\in\mathbb{N}$, of approximative continuation functions from a sequence of LS-training procedures due to L_r basis functions and N_r training trajectories, such that

$$\widehat{C}_k^{(r)}(\tau_k, X_{\tau_k}, \Theta_{\tau_k}) \xrightarrow{P} C_k^*(\tau_k, X_{\tau_k}, \Theta_{\tau_k}), \quad k = 1,\ldots,K, \tag{3.4}$$

for $r \to \infty$, where C_k^* denotes the optimal continuation function at time k, and the process $(\tau_k, X_{\tau_k}, \Theta_{\tau_k})_{k=1,\ldots,K}$ is independent of the training procedures. One next has the following convergence result on stopping times.

Theorem 3.3 *For each* $r\in\mathbb{N}$ *there exists an optimal stopping time* $\left(\mathfrak{n}^{*,r}\right)$ *such that*

$$\mathbb{P}(\mathfrak{n}^{(r)} \neq \mathfrak{n}^{*,r}) \to 0 \ \text{ for } r\to\infty,$$

where

$$\mathfrak{n}^{(r)} := \min\left\{k\geq 0 : Z(\tau_k, X_{\tau_k}) \geq \widehat{C}_k^{(r)}(\tau_k, X_{\tau_k}, \Theta_{\tau_k})\right\}, \quad r\in\mathbb{N}.$$

If in particular $\mathfrak{n}^*$ *is unique, one has that* $\mathfrak{n}^{(r)} \xrightarrow{P} \mathfrak{n}^*$.

Proof. Follows, in principle, from the convergence (3.4) of the continuation values. For a detailed proof in the context of robust optimal stopping we refer to [16, Section 6]. ■

4 Optimal Stopping via Policy Improvement

In this section, we revisit the iterative policy improvement procedure for optimal stopping, which was developed in [15] and extended to multiple stopping in [3]. In these works a deterministic, fixed and finite number of opportunities to stop was assumed, and all proofs exploited the backward recursive nature with initialization at the last opportunity to stop.

In this paper, we convert our initial problem with randomly arriving opportunities into a discrete-time optimal stopping problem with $\mathbb{N}_0$-valued stopping times and an infinite horizon. The infinite horizon entails that the results in [15] are not applicable in the present context. However, we will show that, by applying Theorem 3.1 of that paper to a truncated setting with horizon N, and next letting N tend to infinity, a similar improvement result for an infinite horizon can be established. Subsequently, in an infinite horizon setting, we will prove results similar to Propositions 4.1, 4.3 and 4.4 in [15].

4.1 Iterative construction of the optimal stopping time

Consider a family of stopping times $(\sigma_i)_{i\in\mathbb{N}_0}$ with respect to the discrete filtration $(\mathcal{G}_i)_{i\in\mathbb{N}_0}$ that satisfies the consistency condition

$$\begin{aligned} &i \leq \sigma_i < \infty, \\ &\sigma_i > i \Longrightarrow \sigma_i = \sigma_{i+1}, \quad i \in \mathbb{N}_0, \end{aligned} \tag{4.1}$$

and define the process

$$Y_i^\sigma := \mathbb{E}_{\mathcal{G}_i}\left[U_{\sigma_i}\right].$$

We now fix a window parameter $\kappa \in \mathbb{N} \cup \{\infty\}$ and introduce a new stopping family $(\widehat{\sigma}_i)_{i\in\mathbb{N}_0}$ by

$$\widehat{\sigma}_i := \inf\left\{ j \geq i : \ U_j \geq \widetilde{Y}_j^\sigma \right\}, \tag{4.2}$$

with

$$\widetilde{Y}_j^\sigma = \max_{j\leq k\leq j+\kappa} \mathbb{E}_{\mathcal{G}_j}\left[U_{\sigma_k}\right].$$

Obviously, the family $(\widehat{\sigma}_i)$ also satisfies (4.1). In view of the next proposition, $(\widehat{\sigma}_i)$ may be considered an improvement of (σ_i).

Proposition 4.1 *Let (σ_i) satisfy* (4.1) *and $(\widehat{\sigma}_i)$ be given by* (4.2)*. Then it holds that*

$$Y_i^\sigma \leq \widetilde{Y}_i^\sigma \leq Y_i^{\widehat{\sigma}} \leq Y_i.$$

Proof. Consider for $N \in \mathbb{N}$ the truncated reward $U_i^N := U_i 1_{i\leq N}$ and the stopped policy $\sigma_i^N := \sigma_i \wedge N$. It is easy to see that, for $0 \leq i \leq N$,

$$i \leq \sigma_i^N \leq N, \quad \text{and} \quad \sigma_i^N > i \quad \Longrightarrow \quad \sigma_i^N = \sigma_{i+1}^N,$$

i.e., σ_i^N is consistent in the sense of (4.1). We next consider the iteration step

$$\widehat{\sigma}_i^N := \inf\left\{ j \geq i : U_j^N \geq \max_{j\leq k\leq j+\kappa} \mathbb{E}_{\mathcal{G}_j}\left[U^N_{\sigma_k^N}\right]\right\}$$

and define

$$Y_i^{\sigma^N} := \mathbb{E}_{\mathcal{G}_i}\left[U^N_{\sigma_i^N}\right],$$
$$\widetilde{Y}_i^{\sigma^N} := \max_{i\le k\le i+\kappa} \mathbb{E}_{\mathcal{G}_i}\left[U^N_{\sigma_k^N}\right],$$
$$Y_i^{\widehat{\sigma}^N} := \mathbb{E}_{\mathcal{G}_i}\left[U^N_{\widehat{\sigma}_i^N}\right].$$

Note that $i \le \widehat{\sigma}_i^N \le N$. Due to [15, Theorem 3.1], it holds that

$$Y_i^{\sigma^N} \le \widetilde{Y}_i^{\sigma^N} \le Y_i^{\widehat{\sigma}^N} \le Y_i^{(N)}, \quad \text{for} \quad 0 \le i \le N, \tag{4.3}$$

where $\left(Y_i^{(N)}\right)$ is the Snell envelope for the truncated reward $\left(U_i^N\right)$. We now fix i and show that for $N \to \infty$, a.s.,

(i) $Y_i^{\sigma^N} \to Y_i^{\sigma}$,

(ii) $\widetilde{Y}_i^{\sigma^N} \to \widetilde{Y}_i^{\sigma}$,

(iii) $Y_i^{\widehat{\sigma}^N} \to Y_i^{\widehat{\sigma}}$,

from which the statement follows by (4.3) and the obvious fact that $Y_i^{(N)} \le Y_i$.

Ad (i)+(ii): For any fixed k one has $\sigma_k^N \to \sigma_k < \infty$ a.s. for $N \to \infty$, hence

$$U^N_{\sigma_k^N} = U_{\sigma_k^N} 1_{\{\sigma_k^N \le N\}} = U_{\sigma_k} 1_{\{\sigma_k \le N\}} \to U_{\sigma_k} \ \text{ a.s.}$$

It then follows by conditional dominated convergence that

$$Y_i^{\sigma^N} = \mathbb{E}_{\mathcal{G}_i}\left[U^N_{\sigma_i^N}\right] \to \mathbb{E}_{\mathcal{G}_i}\left[U_{\sigma_k}\right] = Y_i^{\sigma}, \quad \text{and}$$
$$\widetilde{Y}_i^{\sigma^N} = \max_{i\le k\le i+\kappa} \mathbb{E}_{\mathcal{G}_i}\left[U^N_{\sigma_k^N}\right] \to \max_{i\le k\le i+\kappa} \mathbb{E}_{\mathcal{G}_i}\left[U_{\sigma_k}\right] = \widetilde{Y}_i^{\sigma}.$$

Ad (iii): Note that by the definition of U_i^N,

$$\widehat{\sigma}_i^N := \inf\left\{ j \ge i : U_j 1_{\{j\le N\}} \ge \max_{j\le k\le j+\kappa} \mathbb{E}_{\mathcal{G}_j}\left[U_{\sigma_k} 1_{\{\sigma_k \le N\}}\right]\right\}.$$

Let us suppose that we are on the set $\{\omega : \widehat{\sigma}_i(\omega) = l\}$. For $N \ge l$, one then has

$$U_l(\omega) 1_{\{l\le N\}} = U_l(\omega) \ge \max_{l\le k\le l+\kappa} \mathbb{E}_{\mathcal{G}_l}\left[U_{\sigma_k}\right](\omega) \ge \max_{l\le k\le l+\kappa} \mathbb{E}_{\mathcal{G}_l}\left[U_{\sigma_k} 1_{\{\sigma_k\le N\}}\right](\omega)$$

(recall that by assumption $U_\cdot \ge 0$). Hence, for $N \ge l$, one has $\widehat{\sigma}_i^N(\omega) \le l$. Since l was arbitrary, we conclude that

$$\widehat{\sigma}_i^N \le \widehat{\sigma}_i \ \text{ a.s. for } N \ge \mathcal{N}(\omega) \text{ say.}$$

Now assume that $\widehat{\sigma}_i^N \nrightarrow \widehat{\sigma}_i$ with positive probability. Then with positive probability, there must exist numbers r and l with $0 \le r < l$ and a sequence $(N_m)_{m\in\mathbb{N}}$ with $N_m \uparrow \infty$ for $m \uparrow \infty$, such that $\widehat{\sigma}_i = l$ and $\widehat{\sigma}_i^{N_m} = r < l$ for $m \in \mathbb{N}$. This implies

$$U_r 1_{\{r\le N_m\}} \ge \max_{r\le k\le r+\kappa} \mathbb{E}_{\mathcal{G}_r}\left[U_{\sigma_k} 1_{\{\sigma_k \le N_m\}}\right] \quad \text{for all } m \in \mathbb{N},$$

on a set of positive probability. By conditional dominated convergence it then follows that

$$U_r \ge \max_{r\le k\le r+\kappa} \mathbb{E}_{\mathcal{G}_r}\left[U_{\sigma_k}\right]$$

on this set, which contradicts the definition of $\widehat{\sigma}_i$. That is, $\widehat{\sigma}_i^N \to \widehat{\sigma}_i$ a.s., and then (iii) follows by conditional dominated convergence. ■

The next corollary states simple implications of Proposition 4.1.

Corollary 4.2

(i) It holds that $Y_i^{\widehat{\sigma}} \geq U_i$ *for all* $i \in \mathbb{N}_0$.

(ii) Let $\widehat{\sigma}_i$ *and* $\widehat{\sigma}_i'$ *be improvements of* σ_i *due to Proposition 4.1 that correspond to* κ' *and* κ*, respectively. Then, if* $\kappa' > \kappa$*, one has* $Y_i^{\widehat{\sigma}'} \geq Y_i^{\widehat{\sigma}}$.

Proof. Ad (i): By Proposition 4.1 we may write

$$\begin{aligned} 1_{\left\{Y_i^{\widehat{\sigma}} < U_i\right\}} Y_i^{\widehat{\sigma}} &= 1_{\left\{Y_i^{\widehat{\sigma}} < U_i\right\}} 1_{\left\{\widetilde{Y}_i^{\sigma} \leq U_i\right\}} \mathbb{E}_{\mathcal{G}_i}\left[U_{\widehat{\sigma}_i}\right] \\ \text{(definition of } \widehat{\sigma}_i) \;&= 1_{\left\{Y_i^{\widehat{\sigma}} < U_i\right\}} 1_{\left\{\widetilde{Y}_i^{\sigma} \leq U_i\right\}} U_i = 1_{\left\{Y_i^{\widehat{\sigma}} < U_i\right\}} U_i \text{ a.s.} \end{aligned}$$

That is, on the set $\left\{Y_i^{\widehat{\sigma}} < U_i\right\}$ one has $Y_i^{\widehat{\sigma}} = U_i$, a.s., which implies $\mathbb{P}(Y_i^{\widehat{\sigma}} < U_i) = 0$.

Ad (ii): First note that $\widehat{\sigma}_i{}' \geq \widehat{\sigma}_i$ by construction. Hence, we may write

$$\begin{aligned} Y_i^{\widehat{\sigma}'} &= \mathbb{E}_{\mathcal{G}_i}\left[U_{\widehat{\sigma}_i'}\right] = \sum_{l=i}^{\infty} \mathbb{E}_{\mathcal{G}_i}\left[1_{\{\widehat{\sigma}_i = l\}} U_{\widehat{\sigma}_i'}\right] \\ \text{(using (4.1))} &= \sum_{l=i}^{\infty} \mathbb{E}_{\mathcal{G}_i}\left[1_{\{\widehat{\sigma}_i = l\}} U_{\widehat{\sigma}_l'}\right] = \sum_{l=i}^{\infty} \mathbb{E}_{\mathcal{G}_i}\left[1_{\{\widehat{\sigma}_i = l\}} Y_l^{\widehat{\sigma}'}\right] \\ &\geq \sum_{l=i}^{\infty} \mathbb{E}_{\mathcal{G}_i}\left[1_{\{\widehat{\sigma}_i = l\}} U_l\right] = \mathbb{E}_{\mathcal{G}_i}\left[U_{\widehat{\sigma}_i}\right] = Y_i^{\widehat{\sigma}}, \end{aligned}$$

where the already proved statement (i) is used. ■

Corollary 4.2-(ii) thus explains the role of κ: the larger κ, the better the improvement. Proposition 4.1 suggests the following iterative procedure. Define, for $m \in \mathbb{N}_0$, the $\mathcal{G}$-stopping family $\left(\sigma_i^{(m)}\right)_{i \in \mathbb{N}_0}$ as follows: Initialize $\sigma_i^{(0)} = i$, $i \in \mathbb{N}_0$. Suppose $\left(\sigma_i^{(m)}\right)_{i \in \mathbb{N}_0}$ has been defined. Then,

$$\sigma_i^{(m+1)} := \inf\left\{ j \geq i : \; U_j \geq \max_{j \leq k \leq j+\kappa} \mathbb{E}_{\mathcal{G}_j}\left[U_{\sigma_k^{(m)}}\right]\right\}. \tag{4.4}$$

Proposition 4.3 *Let* (σ_i^*) *denote the family of (first) optimal stopping times. That is,*

$$\sigma_i^* := \inf\left\{ j \geq i : \; U_j \geq \mathbb{E}_{\mathcal{G}_j}\left[Y_{j+1}\right]\right\}.$$

For the sequence $\left(\sigma_i^{(m)}\right)$ *defined in* (4.4)*, we have*

$$\sigma_i^{(m)} \leq \sigma_i^{(m+1)} \leq \sigma_i^* < \infty, \quad m \in \mathbb{N}_0, \tag{4.5}$$

and moreover it holds that

$$\sigma_i^* = \uparrow \lim_{m \to \infty} \sigma_i^{(m)} \quad \textit{and} \quad Y_i = \uparrow \lim_{m \to \infty} Y_i^{(m)}, \tag{4.6}$$

with $Y_i^{(m)} \equiv Y_i^{\sigma^{(m)}}$*. The up-arrows indicate that the respective sequences are non-decreasing.*

Proof. The fact that $\sigma_i^* < \infty$ follows from Proposition 2.3. Suppose that $\sigma_i^{(m)} > \sigma_i^*$ for some $m \in \mathbb{N}_0$. Then we have

$$\begin{aligned} U_{\sigma_i^*} &< \max_{\sigma_i^* \le k \le \sigma_i^* + \kappa} \mathbb{E}_{\mathcal{G}_{\sigma_i^*}} \left[U_{\sigma_k^{(m)}} \right] = \sum_{j=i}^{\infty} 1_{\{\sigma_i^* = j\}} \max_{j \le k \le j+\kappa} \mathbb{E}_{\mathcal{G}_j} \left[U_{\sigma_k^{(m)}} \right] \\ \text{(by Proposition (4.1))} &\le \sum_{j=i}^{\infty} 1_{\{\sigma_i^* = j\}} Y_j = Y_{\sigma_i^*}, \end{aligned}$$

which contradicts the optimality of σ_i^*. Suppose next that $\sigma_i^{(m+1)} < \sigma_i^{(m)}$ for some $m \in \mathbb{N}_0$. Then, since $\sigma_i^{(0)} = i$, we must have that $m \ge 1$, and due to the definition of $\sigma_i^{(m)}$ that

$$\begin{aligned} U_{\sigma_i^{(m+1)}} &< \max_{\sigma_i^{(m+1)} \le k \le \sigma_i^{(m+1)} + \kappa} \mathbb{E}_{\mathcal{G}_{\sigma_i^{(m+1)}}} \left[U_{\sigma_k^{(m-1)}} \right] \\ &= \sum_{j=i}^{\infty} 1_{\{\sigma_i^{(m+1)} = j\}} \max_{j \le k \le j+\kappa} \mathbb{E}_{\mathcal{G}_j} \left[U_{\sigma_k^{(m-1)}} \right]. \end{aligned}$$

On the other hand, due to the definition of $\sigma_i^{(m+1)}$ we must have that

$$\begin{aligned} U_{\sigma_i^{(m+1)}} &\ge \max_{\sigma_i^{(m+1)} \le k \le \sigma_i^{(m+1)} + \kappa} \mathbb{E}_{\mathcal{G}_{\sigma_i^{(m+1)}}} \left[U_{\sigma_k^{(m)}} \right] \\ &= \sum_{j=i}^{\infty} 1_{\{\sigma_i^{(m+1)} = j\}} \max_{j \le k \le j+\kappa} \mathbb{E}_{\mathcal{G}_j} \left[U_{\sigma_k^{(m)}} \right], \end{aligned}$$

which in turn yields

$$\sum_{j=i}^{\infty} 1_{\{\sigma_i^{(m+1)} = j\}} \max_{j \le k \le j+\kappa} \mathbb{E}_{\mathcal{G}_j} \left[U_{\sigma_k^{(m)}} \right] < \sum_{j=i}^{\infty} 1_{\{\sigma_i^{(m+1)} = j\}} \max_{j \le k \le j+\kappa} \mathbb{E}_{\mathcal{G}_j} \left[U_{\sigma_k^{(m-1)}} \right].$$

However, since for all $j \ge i$,

$$\max_{j \le k \le j+\kappa} \mathbb{E}_{\mathcal{G}_j} \left[U_{\sigma_k^{(m)}} \right] \ge \mathbb{E}_{\mathcal{G}_j} \left[U_{\sigma_j^{(m)}} \right] = Y_j^{(m)} \ge \max_{j \le k \le j+\kappa} \mathbb{E}_{\mathcal{G}_j} \left[U_{\sigma_k^{(m-1)}} \right] \quad \text{a.s.,}$$

by Proposition (4.1), we get a contradiction. Thus, (4.5) is proved.

Due to (4.5) and Proposition 4.1 we may define

$$\sigma_i^{\infty} =\uparrow \lim_{m \to \infty} \sigma_i^{(m)} \quad \text{and} \quad Y_i^{\infty} =\uparrow \lim_{m \to \infty} Y_i^{(m)} \quad \text{a.s.}$$

Note that $\sigma_i^{\infty} \le \sigma_i^* < \infty$, and that, since $\sigma_i^{(m)}$ is integer valued, one must have $\sigma_i^{(m)}(\omega) = \sigma_i^{(\infty)}(\omega)$ for $m > N(i, \omega)$ say. Hence, $U_{\sigma_i^{(m)}} \to U_{\sigma_i^{\infty}}$ a.s., and then by dominated convergence we get

$$Y_i^{\infty} =\uparrow \lim_{m \to \infty} Y_i^{(m)} =\uparrow \lim_{m \to \infty} \mathbb{E}_{\mathcal{G}_i} \left[U_{\sigma_i^{(m)}} \right] = \mathbb{E}_{\mathcal{G}_i} \left[U_{\sigma_i^{\infty}} \right].$$

Obviously, we have that $Y_i \ge Y_i^{\infty} \ge U_i$ a.s. Since (Y_i) is the smallest supermartingale that dominates the reward (U_i) by Proposition 2.3, we can prove that $\sigma_i^{\infty} = \sigma_i^*$ and $Y_i^{\infty} = Y_i$, and hence (4.6), by showing that (Y_i^∞) is a supermartingale. It is easy to see that (σ_i^∞) is a consistent $\mathcal{G}$-stopping family, i.e., (σ_i^∞) satisfies (4.1). So for any $i \geq 0$ one has

$$\begin{aligned}\mathbb{E}_{\mathcal{G}_i}\left[Y_{i+1}^\infty\right] &= \mathbb{E}_{\mathcal{G}_i}\left[U_{\sigma_{i+1}^\infty}\right] = Y_i^\infty + \mathbb{E}_{\mathcal{G}_i}\left[\left(U_{\sigma_{i+1}^\infty} - U_{\sigma_i^\infty}\right)\right] \\ &= Y_i^\infty + \mathbb{E}_{\mathcal{G}_i}\left[1_{\{\sigma_i^\infty = i\}}\left(U_{\sigma_{i+1}^\infty} - U_{\sigma_i^\infty}\right)\right] \\ \text{(by (4.1))} \quad &= Y_i^\infty + 1_{\{\sigma_i^\infty = i\}}\mathbb{E}_{\mathcal{G}_i}\left[U_{\sigma_{i+1}^\infty}\right] - 1_{\{\sigma_i^\infty = i\}}U_i. \end{aligned} \tag{4.7}$$

Note that by (4.5) and the definition of σ_i^∞ one has $\{\sigma_i^\infty = i\} = \bigcap_{m=0}^{\infty}\left\{\sigma_i^{(m)} = i\right\}$. Hence, by the definition of $\sigma_i^{(m')}$, for all $m' \geq 1$,

$$\begin{aligned}1_{\{\sigma_i^\infty = i\}}U_i &\geq 1_{\{\sigma_i^\infty = i\}} \max_{i \leq k \leq i+\kappa} \mathbb{E}_{\mathcal{G}_i}\left[U_{\sigma_k^{(m'-1)}}\right] \\ &\geq 1_{\{\sigma_i^\infty = i\}}\mathbb{E}_{\mathcal{G}_i}\left[U_{\sigma_i^{(m'-1)}}\right]\end{aligned}$$

Since $\uparrow \lim_{m' \uparrow \infty} \mathbb{E}_{\mathcal{G}_i}\left[U_{\sigma_i^{(m'-1)}}\right] = \mathbb{E}_{\mathcal{G}_i}\left[U_{\sigma_i^\infty}\right]$ a.s., it then follows that

$$1_{\{\sigma_i^\infty = i\}}U_i \geq 1_{\{\sigma_i^\infty = i\}}\mathbb{E}_{\mathcal{G}_i}\left[U_{\sigma_i^\infty}\right],$$

hence by (4.7) we get $\mathbb{E}_{\mathcal{G}_i}\left[Y_{i+1}^\infty\right] \leq Y_i^\infty$. That is, (Y_i^∞) is a supermartingale. ■

4.2 Pseudo-algorithm

Let us assume we are in a Markovian setting in which Assumption 2.7 is satisfied. Further assume that we are given an input stopping family (σ_i) satisfying (4.1). Such a family may be obtained, for example, by the least squares method in Section 3 choosing some truncation level K if need be. That is, (σ_i) may be given by

$$\sigma_i := \inf\left\{j \geq i : Z(\tau_j, X_{\tau_j}) \geq \widehat{C}_j(\tau_j, X_{\tau_j}, \Theta_{\tau_j})\right\}. \tag{4.8}$$

We then choose some window parameter $\kappa \geq 1$ and improve this policy due to Proposition 4.1. We thus have to compute on a particular trajectory at time τ_j in X_{τ_j}, $j \geq 0$,

$$\mathbb{E}_{\mathcal{G}_j}\left[U_{\sigma_k}\right] = \mathbb{E}_{X_{\tau_j}}\left[Z(\tau_{\sigma_k}, X_{\tau_{\sigma_k}})\right], \quad j \leq k \leq j+\kappa,$$

which can be done by sub-simulations: Starting at $\tau_j^{(n)}$ in $X^{n,j} \equiv X_{\tau_j^{(n)}}^{(n)}$ on an outer trajectory with number n, $n = 1, \ldots, N$, simulate, for $m = 1, \ldots, M$, sub-trajectories

$$\left(X_s^{n,m,X^{n,j}}\right)_{s \geq \tau_j^{(n)}},$$

and stopping times $\sigma_k^{n,m}$, $j \leq k \leq j+\kappa$, according to (4.8). Then construct

$$\widehat{\sigma}_0^{(n)} := \inf\left\{j \geq 0 : Z(\tau_j^{(n)}, X^{n,j}) \geq \max_{j \leq k \leq j+\kappa} \frac{1}{M}\sum_{m=1}^{M} Z(\tau_{\sigma_k^{n,m}}, X_{\tau_{\sigma_k^{n,m}}}^{n,m,X^{n,j}})\right\},$$

and compute

$$\frac{1}{N}\sum_{n=1}^{N} Z(\widehat{\sigma}_0^{(n)}, X_{\widehat{\sigma}_0^{(n)}}^{(n)}) \approx \mathbb{E}\left[U_{\widehat{\sigma}_0}\right] = Y_0^{\widehat{\sigma}} \geq Y_0^{\sigma}.$$

5 Numerical Test Cases and Illustrations

In this section, we analyze our numerical methods in several examples. In our numerical analyses based on random times LSMC, we consider only the first $\bar{K}$ random times for each problem, where $\bar{K}$ may depend on the initial values and parameters chosen. From Section 3.1, we know that for the finite horizon approximation to be ε-close, we need to ensure that $\sqrt{\mathbb{P}(\tau_{\bar{K}} \leq T)}B < \varepsilon$, where $\mathbb{E}[\sup_{i\in\mathbb{N}_0} Z^2_{\tau_i}] \leq B^2$. We would like to choose $\bar{K}$ as small as possible. From (3.1) one easily verifies that, for any $\mathfrak{n}$ with $\mathfrak{n} > K$,

$$\begin{aligned}\mathbb{E}\left[U_{\mathfrak{n}}\right] &= \mathbb{E}\left[Z_{\tau_{\mathfrak{n}}} 1_{\{\tau_{\mathfrak{n}}\leq T\}}\right] + \mathbb{E}\left[Z_{\tau_{\mathfrak{n}}} 1_{\{\tau_{\mathfrak{n}}> T\}}\right] \leq \mathbb{E}\left[Z_{\tau_{\mathfrak{n}}} 1_{\{\tau_K\leq T\}}\right] \\ &\leq \mathbb{E}\left[1_{\{\tau_K\leq T\}} \sup_{i>K} Z_{\tau_i}\right] \leq \mathbb{E}\left[1_{\{\tau_K\leq T\}} \sup_{i\in\mathbb{N}_0} Z_{\tau_i}\right] \leq \sqrt{\mathbb{P}\left(\tau_K \leq T\right)}B < \varepsilon.\end{aligned}$$

Hence, when relying on numerical simulation (rather than analytic computation), we observe that it suffices to pick $\bar{K}$ such that $\mathbb{E}[1_{\{\tau_{\bar{K}}\leq T\}} \sup_{i>\bar{K}} Z_{\tau_i}] < \varepsilon$, which can lead to values of $\bar{K}$ that are substantially smaller than would be selected from $\sqrt{\mathbb{P}(\tau_{\bar{K}} \leq T)}B < \varepsilon$. To this end, we obtain a simulation estimate of $\mathbb{E}[1_{\{\tau_{\bar{K}}\leq T\}} \sup_{i>\bar{K}} Z_{\tau_i}]$ by using $100{,}000$ paths ($50{,}000$ antithetic) and add two standard errors on top of the estimated mean to be on the safe side. We then pick the smallest $\bar{K}$ for which the resulting value is smaller than $\varepsilon = 0.001$. In our numerical analysis using policy iteration, we take the window parameter equal to $\bar{K}$ in view of Corollary 4.2-(ii). Using the same value of $\bar{K}$ across the different methods (primal, dual, policy iteration) within the same example and the same parameter combination aids in ensuring a fair comparison.

5.1 A benchmark example

For our first example, we consider a geometric Brownian motion with jumps, given for all $t \geq 0$ by

$$X_t = X_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t + N_t \ln(1+j)},$$

with $X_0 > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$ and $j > -1$ known constants. The process W is a standard Brownian motion and N is an (independent) Poisson process with intensity $\tilde{\lambda}$. To illustrate the approach proposed in this paper, we will analyze an optimal stopping problem with payoff $e^{-r(s-t)}(X_s)^\eta$ at time $s \in [t, T]$, for a given power parameter $\eta > 0$ and discount rate $r > 0$. We suppose that stopping is only allowed at the first K jump times of N, i.e., at times $s \in]t, T]$ that satisfy $N_{s-} < N_s \leq K$.

To investigate the performance of our method, we first derive the optimal strategy in closed form. In this example, for ease of exposition the value function for the problem is denoted by $v_K(t, x, d)$ since the current time t and x, the current value of X, are both components of Z_t and it turns out to be convenient to express the value in terms of K, the (remaining) number of opportunities to stop. This includes the current opportunity if stopping is allowed at the current time; the value of d indicates whether we can stop at a particular moment $(d = 1)$ or not $(d = 0)$. The value at time zero, $Y_0^{(K)}$, is then given by $v_K(0, x, 0)$.

These definitions imply that $v_0 \equiv 0$, since the value function is equal to zero when there are no opportunities to stop left. We have for all $K \geq 1$ that

$$v_K(t, x, 1) = x^\eta 1_{\{t\leq T\}} \vee v_{K-1}(t, x, 0), \tag{5.1}$$

because at a point in time where it is possible to stop, we either choose to receive the payoff at that time or choose not to, in which case there will be one opportunity to stop less going forward. At moments where it is not possible to stop, the value function will equal the discounted version of the value at the following possible stopping time, provided this next opportunity occurs before the time of maturity T, so for $K \geq 1$,

$$v_K(t,x,0) = \int_0^{T-t} \tilde{\lambda} e^{-(r+\tilde{\lambda})u} \mathbb{E}\left[v_K\left(t+u, x(1+j)e^{(\mu-\frac{1}{2}\sigma^2)u+\xi\sigma\sqrt{u}}, 1\right)\right] \,\mathrm{d}u, \tag{5.2}$$

with $\xi \sim N(0,1)$. To solve equations (5.1)–(5.2), we substitute $v_K(t,x,0) = x^\eta e^{\alpha(T-t)} f_K(T-t)$ with $\alpha = \mu\eta + \frac{1}{2}\sigma^2\eta(\eta-1) - (r+\tilde{\lambda})$. We impose that $\alpha > 0$, which is assumed from now on. The substitution implies that we have found a solution for the value function if for all $K \geq 0$

$$f_{K+1}(s) = \lambda \int_0^s (e^{-\alpha u} \vee f_K(u)) \,\mathrm{d}u,$$

with $\lambda = \tilde{\lambda}(1+j)^\eta$ and $f_0 \equiv 0$, which immediately gives that $f_1(s) = \frac{\lambda}{\alpha}(1-e^{-\alpha s})$. Let s^* be the solution to $f_1(s) = e^{-\alpha s}$ so $s^* = \ln(1+\alpha/\lambda)/\alpha$ and define $\zeta := f_1(s^*) = e^{-\alpha s^*} = \lambda/(\lambda+\alpha)$. We use a further Ansatz $f_{K+1}(s) = f_1(s)1_{\{s\leq s^*\}} + z_{K+1}(s-s^*)1_{\{s>s^*\}}$ and conclude that this is a solution for $K > 0$ if

$$\partial_s z_{K+1}(s) = \partial_s f_{K+1}(s+s^*) = \lambda f_K(s+s^*) = \lambda z_K(s), \qquad z_{K+1}(0) = \zeta,$$

while for $K = 0$ we have

$$z_1(s) = f_1(s+s^*) = \tfrac{\lambda}{\alpha}(1-e^{-\alpha s}e^{-\alpha s^*}) = \tfrac{\lambda}{\alpha}(1-\zeta e^{-\alpha s}) = \zeta + \zeta\tfrac{\lambda}{\alpha}(1-e^{-\alpha s}).$$

A power series expansion gives

$$\begin{aligned} z_{K+1}(s) &= \zeta \sum_{m=0}^{K} \tfrac{(\lambda s)^m}{m!} + \zeta(\tfrac{-\lambda}{\alpha})^{K+1} \sum_{m=K+1}^{\infty} \tfrac{(-\alpha s)^m}{m!} \\ &= \zeta \tfrac{\Gamma_{K+1}(\lambda s)}{K!} e^{\lambda s} + \zeta \tfrac{\gamma_{K+1}(-\alpha s)}{K!} (\tfrac{-\lambda}{\alpha})^{K+1} e^{-\alpha s}, \end{aligned}$$

with Γ_{K+1} and γ_{K+1} the upper and lower Gamma functions of order $K+1$, respectively. The value function must therefore be, for $K \geq 1$,

$$\begin{aligned} v_K(t,x,0) = {} & 1_{\{t>T-s^*\}} x^\eta \tfrac{\lambda}{\alpha}(e^{\alpha(T-t)}-1) \\ & + 1_{\{t\leq T-s^*\}} x^\eta \tfrac{\lambda}{\alpha+\lambda} e^{\alpha(T-t)} \tfrac{\Gamma_{K+1}(\lambda(T-t-s^*))}{K!} e^{\lambda(T-t-s^*)} \\ & + 1_{\{t\leq T-s^*\}} x^\eta \tfrac{\lambda}{\alpha+\lambda} e^{\alpha(T-t)} \tfrac{\gamma_{K+1}(-\alpha(T-t-s^*))}{K!} (\tfrac{-\lambda}{\alpha})^{K+1} e^{-\alpha(T-t-s^*)}. \end{aligned} \tag{5.3}$$

Since $v_K(t,x,1) = x^\eta 1_{\{t<T\}} \vee v_{K-1}(t,x,0)$ stopping is optimal whenever $t \geq T - s^* =: t^*$ or when we reach the last possible moment on which a payoff can be received $(K=1)$. The optimal stopping strategy does not depend on K for $K \geq 2$ but the value function does, since the number of exercise times we can ignore before t^* matters.

For $K \to \infty$, we find the value function for the case where we may exercise as often as we want:

$$\begin{aligned} v_\infty(t,x,0) &= 1_{\{t>t^*\}} x^\eta \tfrac{\lambda}{\alpha}(e^{\alpha(T-t)}-1) \;+\; 1_{\{t\leq t^*\}} x^\eta \tfrac{\lambda}{\alpha+\lambda} e^{\alpha(T-t)} e^{\lambda(T-t-s^*)} \\ &= 1_{\{t>t^*\}} x^\eta \tfrac{\lambda}{\alpha}((1+\tfrac{\alpha}{\lambda})e^{-\alpha(t-t^*)}-1) \;+\; 1_{\{t\leq t^*\}} x^\eta e^{(\lambda+\alpha)(t^*-t)}. \end{aligned} \tag{5.4}$$

Figure 1 shows scaled versions of the value function $v_K(t,x,0)/x^\eta$ for different characteristics of the arrival process for exercise opportunities and different parameters of the geometric Brownian motion with jumps. We notice that for the parameter values used here, the effect of allowing for at most $1, 3$ or an unlimited number of randomly arriving exercise opportunities is much smaller than the effect of changing the size of the downward

T	$\tilde{\lambda}$	μ	K	j	True	Primal	(s.e.)	Dual	(s.e.)
3	1	0.2	1	-0.05	1.3112	1.3103	0.0007	1.3103	0.0007
3	1	0.2	3	-0.05	1.5250	1.5263	0.0010	1.5265	0.0010
3	1	0.2	∞	-0.05	1.5448	1.5443	0.0010	1.5448	0.0010
1	1	0.2	∞	-0.05	0.6910	0.6911	0.0004	0.6911	0.0004
3	2	0.5	∞	-0.05	6.4685	6.4686	0.0048	6.4728	0.0052
3	1	0.2	∞	0.00	1.9639	1.9644	0.0013	1.9658	0.0014

Table 1: Randomized stopping of geometric Brownian motion with jumps.

Notes. This table displays the value of $Y_0^{(K)}$ for the square of a geometric Brownian motion with jumps stopped at random times. We suppose that random stopping opportunities for the payoff $(X_t)^\eta$ with $\eta = 2$ coincide with jump times for X, which arrive according to a homogeneous Poisson process with rate $\tilde{\lambda}$. The process X starts in $X_0 = 1$ and has drift parameter μ, volatility parameter $\sigma = 20\%$ and relative jump size j. Exercising is restricted to the first K Poisson events and to times before the maturity T. If $K = \infty$, we restrict our attention to the first $\bar{K}$ random times for each of the algorithms, where $\bar{K}$ is chosen to ensure that the truncation error ε is at most 0.001. The sixth column (marked "True") shows the value function $v_K(0, x, 0)$ as calculated in (5.3)–(5.4). The next two columns contain the value and standard error for the lower bound based on our random times LSMC algorithm. The regression coefficients for the exercise policy are estimated using $200{,}000$ paths. The primal, lower bound estimate is then determined by evaluating this policy along $2{,}000{,}000$ (new) paths. The last two columns display the value and standard error for the upper bound, which is based on the dual algorithm, using the policy from the primal method. We use $1{,}500$ paths and $10{,}000$ sub-simulations along each path for this algorithm. The continuation values are estimated by regressing on the variables $\phi_i(t)(x_t)^j$, where ϕ_i denotes the i-th Laguerre polynomial, with $i \in \{0, 1, 2, 3, 4, 5\}$, $j \in \{0, 1, 2, 3\}$.

jumps or the expected number of exercise opportunities before maturity. The optimal stopping times in terms of the remaining time to maturity, s^*, have been indicated by circles in Figure 1. They correspond to the point in time at which the scaled value function equals one. As expected, stopping tends to take place later (hence, s^* is smaller) when the expected number of remaining exercise opportunities is larger and/or when future values of the geometric Brownian motion with jumps have a higher expected value.

The explicit form of the value function makes this optimal stopping problem a suitable benchmark to test our random times LSMC approach. Table 1 displays numerical results for $Y_0^{(K)} = v_K(0, x, 0)$ using different values of the maturity T, jump intensity $\tilde{\lambda}$, drift parameter of the geometric Brownian motion μ, the allowed number of moments to stop K and the relative jump size j. To facilitate comparison with examples in [2], we take the same number of paths ($200{,}000$ from $100{,}000$ antithetic samples) to estimate regression coefficients and then estimate a lower bound for the value function by applying the stopping strategy based on these coefficients to $2{,}000{,}000$ (new) paths ($1{,}000{,}000$ antithetic). For our version of the dual method we use $1{,}500$ paths (750 antithetic) and $10{,}000$ sub-simulations per path ($5{,}000$ antithetic). This is one tenth of the number of sub-simulations in [2], since we found that this reduces the time for our computations by a factor five, whereas standard errors are hardly effected by this modification.

We use 24 basis functions in the regression and the results in Table 1 show that for this benchmark example with known theoretical values, this leads to lower and upper bound estimates that generate quite narrow confidence intervals. They coincide in the first row of the table, since stopping at the earliest opportunity is optimal.

We note that the standard errors are mainly determined by the lower bound computations since we take the regression coefficients, and hence our approximated optimal policy, fixed when calculating the upper bound estimates. The computation times are dominated by the calculations for the upper bound, since these require many sub-simulations.

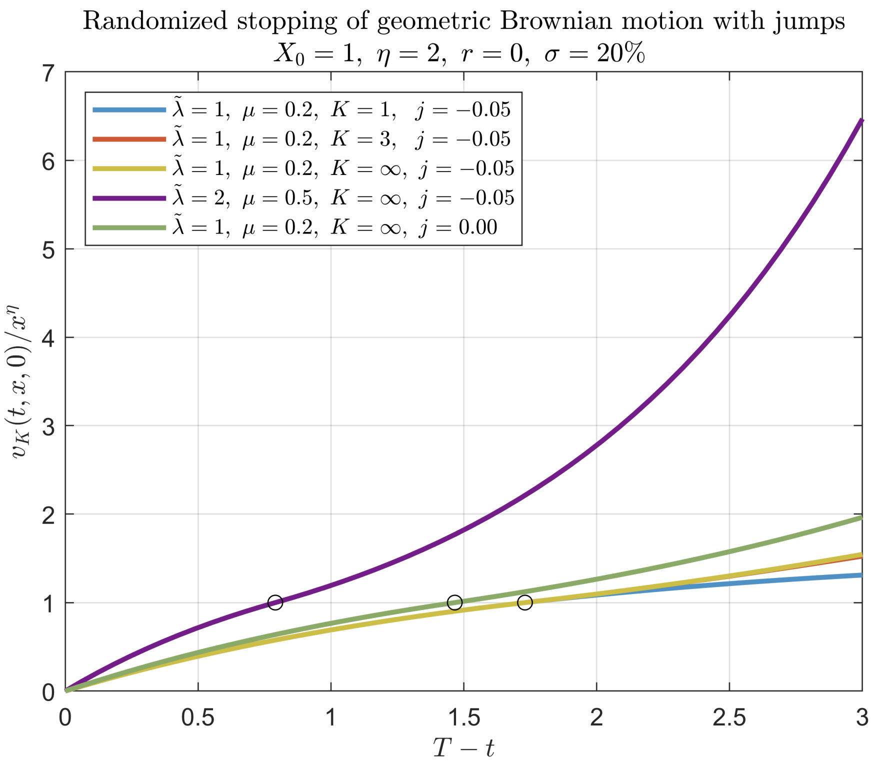

Figure 1: Randomized stopping of geometric Brownian motions with jumps.

Notes. The scaled value function $v_K(t,x,0)/x^\eta$ is displayed for different remaining times to maturity $T-t$. The different lines correspond to the cases shown in Table 1; see there for a description of the numerical procedure and assumptions. The black circles indicate the remaining time to maturities $s^* = T - t^*$ where it becomes optimal to stop, i.e., the times where $v_K(T - s^*, x, 0) = x^\eta$. We notice that the second and third cases (the red and yellow lines) are almost indistinguishable.

5.2 A max-call on random times

In view of the good performance of our random times LSMC primal-dual algorithm in the benchmark example, we now analyze the algorithm in a more complex example for which analytic reference values are no longer available. We consider M jump-diffusions, interpreted as underlying asset prices driving the value of potential investment projects, with dynamics governed by the following stochastic differential equations:

$$\mathrm{d}X_t^m = (r - \delta^m - \mu_J^m \lambda_J^m) X_t^m \,\mathrm{d}t + \sigma^m X_t^m \,\mathrm{d}W_t^m + \mu_J^m X_{t-}^m \,\mathrm{d}N_t^m,\ m = 1, \ldots, M,$$

where the W_t^m are independent standard Brownian motions, the N_t^m are independent homogeneous Poisson processes with arrival rates $\lambda_J^m \geq 0$ and r, δ^m, $\mu_J^m \in (-1, \infty)$ and $\sigma^m \geq 0$ are deterministic constants. The processes W_t^m and N_t^m are assumed to be mutually independent.

We consider the payoff of a max-call:

$$Z_t = \begin{cases} e^{-rt} \left(\max_{m=1,\ldots,M} X_t^m - \mathcal{K}\right)^+, & 0 < t \leq T \\ 0, & \text{otherwise} \end{cases}.$$

We suppose that the opportunities to stop are generated by a Poisson process over the interval $[0, T]$. As such, this example can be seen as (a multi-dimensional extension of) the finite horizon counterpart of the perpetual call as considered e.g., in [8].

We start by considering opportunities to stop generated by a homogeneous Poisson process with rate λ that is independent of all other processes described above. Following [2], we set $r = 5\%$, $\delta^m = 10\%$ and $\sigma^m = 0.2$, for $m = 1, \ldots, M$.

X_0	λ	Primal	(s.e.)	Dual	(s.e.)
90	1	3.0870	0.0058	3.0917	0.0064
90	2	3.7346	0.0061	3.7392	0.0062
90	5	4.1626	0.0063	4.1854	0.0073
100	1	6.0708	0.0079	6.0726	0.0080
100	2	7.1100	0.0082	7.1151	0.0082
100	5	7.7242	0.0082	7.7520	0.0088
110	1	10.6437	0.0099	10.6462	0.0099
110	2	12.1099	0.0097	12.1164	0.0098
110	5	12.9234	0.0095	12.9535	0.0100

Table 2: Value of a call option on random times ($M = 1$).

Notes. This table displays the value of a call option with stochastic exercise opportunities. The payoff of this contract is $\max(X_t - \mathcal{K}, 0)$ with $\mathcal{K} = 100$. The parameters are $r = 5\%$, $\delta = 10\%$, $\sigma = 20\%$ and $T = 3$. Exercise opportunities arrive according to a homogeneous Poisson process with rate λ that is independent of the asset price dynamics. The truncation error, numbers of simulations and specification of the regression variables are identical to those in Table 1.

An appropriate choice of the basis functions that we employ in the LSMC regressions is crucial. Denote the ordered asset prices by $X_t^{(1)}, \ldots, X_t^{(M)}$, where (1) refers to the largest and (M) to the smallest asset price at time t. For $M = 1$, we use the 24 variables $\phi_i(t)x^j$, where $i \in \{0, 1, 2, 3, 4, 5\}$ and $j \in \{0, 1, 2, 3\}$. Here, ϕ_i denotes the i-th Laguerre polynomial. For $M \geq 2$, we use the 51 variables $\phi_i(t)(x_t^{(m)})^j$ and $(x_t^{(1)})^j(x_t^{(2)})^k$, where $i \in \{0, 1, 2, 3, 4, 5\}$, $j,\ k \in \{0, 1, 2, 3\}$ and $m \in \{1, 2\}$.

Just like in the previous subsection, we use $200{,}000$ simulated paths ($100{,}000$ antithetic) to estimate the regression functions and $2{,}000{,}000$ ($1{,}000{,}000$ antithetic) simulated paths to obtain a primal estimate for the value function based on the random times LSMC policy. We then use $1{,}500$ simulated trajectories (750 antithetic) to obtain dual estimates, where we use the random times LSMC policy to determine the martingale increments. We use $10{,}000$ sub-simulations ($5{,}000$ antithetic) along each of these trajectories to estimate the conditional expectations required for the dual method.

We report the primal and dual estimates obtained using random times LSMC when $\lambda_J^m \equiv \mu_J^m \equiv 0$ and $M \equiv 1$ in Table 2. We observe that the differences between the primal and dual estimates, i.e., the duality gaps, are typically (very) small and are generally increasing in the arrival rate of the opportunities to stop, λ. The sizes of the duality gaps are typically substantially smaller than 1% of the price of the option, whereas the relative standard errors are in the range 0.1% to 0.2%.

Next, we consider the case of multiple underlying assets in Table 3 (for $M = 2$) and Table 4 (for $M = 5$), *ceteris paribus*. We find that the performance of our algorithm is robust when increasing the number of underlying assets.

The results in Tables 3–4 may be viewed as the random times counterpart of Table 2 in [2], which considers a Bermudan max-call option on the same underlying assets. More specifically, let us consider the case $M = 2$ in Table 3 and compare our results to those in Table 2 of [2]. They consider 9 equidistant exercise opportunities on the interval $[0, 3]$, whereas in our setting there are in expectation $\lambda T = \{3, 6, 15\}$ times at which the contract may be exercised. It is remarkable that, even for the case $\lambda = 5$, i.e., when there are 15 exercise opportunities in expectation, our estimated values are substantially lower than those in [2]: the difference is about $0.50\$$, which amounts to 2.5% of the asset price. For $\lambda = 2$, this difference is even larger than 1 dollar. This entails that there is a substantial premium on “market frictions”.

Finally, we analyze the performance of our algorithm when the underlying asset price dynamics includes jumps. We report in Table 5 the primal and dual estimates obtained using random times LSMC when $\lambda_J^m = 1$ and $\mu_J^m = 0.06$, for $m = 1, \ldots, M$, *ceteris paribus*. The results illustrate that the differences between the

X_0	λ	Primal	(s.e.)	Dual	(s.e.)
90	1	5.6876	0.0076	5.6950	0.0081
90	2	6.8617	0.0080	6.8754	0.0084
90	5	7.6772	0.0083	7.7168	0.0094
100	1	10.5710	0.0101	10.5769	0.0102
100	2	12.3455	0.0104	12.3607	0.0108
100	5	13.4558	0.0105	13.5195	0.0184
110	1	17.1219	0.0123	17.1269	0.0123
110	2	19.5328	0.0123	19.5600	0.0151
110	5	20.8970	0.0123	20.9724	0.0208

Table 3: Value of a max-call option on random times ($M = 2$).

Notes. This table displays the value of a max-call option with stochastic exercise opportunities on 2 independent geometric Brownian motions X_t^1, X_t^2. The payoff of this contract is $\max(\max_{m=1,2} X_t^m - \mathcal{K}, 0)$ with $\mathcal{K} = 100$, where the dynamics of the individual processes and exercise opportunities are as in Table 2. The truncation error and numbers of simulations are identical to those in Table 2. For the specification of the regression variables, consider the ordered price processes where $X_t^{(1)}$ denotes the largest price at time t and $X_t^{(M)}$ denotes the smallest price at time t. We then use the regression variables $\phi_i(t)(x_t^{(m)})^j$ and $(x_t^{(1)})^j(x_t^{(2)})^k$, where $i \in \{0,1,2,3,4,5\}, j,\ k \in \{0,1,2,3\}$ and $m \in \{1,2\}$.

X_0	λ	Primal	(s.e.)	Dual	(s.e.)
90	1	11.6716	0.0103	11.6791	0.0104
90	2	14.0658	0.0108	14.0987	0.0118
90	5	15.7416	0.0111	15.8176	0.0132
100	1	19.5441	0.0129	19.5669	0.0135
100	2	22.8863	0.0131	22.9397	0.0144
100	5	25.0735	0.0134	25.1935	0.0155
110	1	28.6483	0.0151	28.6904	0.0161
110	2	32.9037	0.0150	32.9806	0.0167
110	5	35.5506	0.0152	35.7108	0.0177

Table 4: Value of a max-call option on random times ($M = 5$).

Notes. This table displays the value of a max-call option with stochastic exercise opportunities on 5 independent geometric Brownian motions $X_t^1, \dots, X_t^5$. The payoff of this contract is $\max(\max_{m=1,\dots,5} X_t^m - \mathcal{K}, 0)$ with $\mathcal{K} = 100$, where the dynamics of the individual processes and exercise opportunities are as in Table 2. The truncation error, numbers of simulations and specification of the regression variables are identical to those in Table 3.

primal and the dual estimates remain small, even in the presence of jumps in the asset price dynamics. The standard errors are generally only slightly larger.

5.3 A max-call on asset price dependent random times

We now analyze the performance of our random times LSMC algorithm when the occurrence of exercise opportunities is dependent on the evolution of the underlying asset price, in the spirit of Example 2.11(ii). In particular, exercise opportunities are generated by a homogeneous Poisson process that is active only when the lowest underlying asset price, $X^{(M)}$, lies above a certain threshold value, which we specify to be 80. The remaining parameters are chosen as in the previous subsection. The numbers of simulated trajectories are also identical to those in the previous subsection.

For $M \geq 2$, we now use the following basis functions: $\phi_i(t)(x_t^{(m)})^j$ and $(x_t^{(m)})^j(x_t^{(n)})^k$, where $i \in \{0,1,2,3,4,5\}, j,\ k \in \{0,1,2,3\}$ and $m \neq n \in \{1,\dots,M\}$.

X_0	λ	Primal	(s.e.)	Dual	(s.e.)
90	1	6.2414	0.0083	6.2480	0.0085
90	2	7.5256	0.0087	7.5413	0.0107
90	5	8.3956	0.0090	8.4617	0.0168
100	1	11.2297	0.0108	11.2353	0.0109
100	2	13.1204	0.0111	13.1394	0.0133
100	5	14.2867	0.0113	14.4018	0.0608
110	1	17.7946	0.0130	17.8030	0.0131
110	2	20.3377	0.0131	20.3695	0.0157
110	5	21.7476	0.0131	21.8834	0.0592

Table 5: Value of a max-call option on random times with jumps in the underlying ($M = 2$).

Notes. This table displays the value of a max-call option with stochastic exercise opportunities on 2 independent geometric Brownian motions with jumps X_t^1, X_t^2. The payoff, dynamics of exercise opportunities and Brownian components of the processes X_t^1, X_t^2 are identical to those in Table 3; for the jump components of X_t^1, X_t^2, we take $\lambda_J^m = 1$, $\mu_J^m = 0.06$, for $m = 1, 2$. The truncation error, numbers of simulations and specification of the regression variables are identical to those in Table 3.

As is apparent from the left panels in Tables 6–8, the duality gaps remain small, confirming the good performance of our algorithm.

X_0	λ	Primal	(s.e.)	Dual	(s.e.)	Andersen	(s.e.)	PI	(s.e.)
90	2	3.7298	0.0061	3.7361	0.0064	3.6341	0.0059	3.7295	0.0103
90	4	4.0989	0.0063	4.1219	0.0074	3.9925	0.0059	4.0952	0.0109
100	2	7.0937	0.0082	7.1013	0.0083	6.9797	0.0078	7.0892	0.0136
100	4	7.6307	0.0082	7.6545	0.0090	7.5141	0.0078	7.6145	0.0139
110	2	12.1033	0.0097	12.1092	0.0098	11.9888	0.0095	12.0641	0.0160
110	4	12.7935	0.0095	12.8159	0.0099	12.6985	0.0092	12.8023	0.0168

Table 6: Value of a call option on asset price dependent random times ($M = 1$).

Notes. This table displays the value of a call option with stochastic exercise opportunities of which the arrival depend on the underlying asset price X_t. The payoff and dynamics of the process X_t are identical to those in Table 2. Exercise opportunities arrive according to a homogeneous Poisson process with rate $\lambda \in \{2, 4\}$ as long as $X_t \geq 80$. The left panel provides results for random times LSMC. The truncation error, numbers of simulations and specification of the regression variables are identical to those in Table 2. The right panel describes the results of the random times Andersen policy as described in Section 5.4. Subsequently, one step of policy iteration is performed with a window parameter equal to $\bar{K}$.

X_0	λ	Primal	(s.e.)	Dual	(s.e.)	Andersen	(s.e.)	PI	(s.e.)
90	2	3.9446	0.0064	3.9488	0.0065	3.8006	0.0061	3.9407	0.0115
90	4	4.5826	0.0067	4.6098	0.0101	4.3296	0.0062	4.5681	0.0138
100	2	9.4180	0.0093	9.4283	0.0096	9.0255	0.0086	9.4097	0.0198
100	4	10.5569	0.0095	10.5846	0.0105	9.9252	0.0086	10.5426	0.0222
110	2	17.2964	0.0116	17.3122	0.0121	16.5340	0.0107	17.2817	0.0277
110	4	18.8159	0.0116	18.8605	0.0134	17.7089	0.0105	18.6733	0.0301

Table 7: Value of a max-call option on asset price dependent random times ($M = 2$).

Notes. This table displays the value of a max-call option with stochastic exercise opportunities of which the arrival depend on the underlying asset prices X_t^1, X_t^2. The payoff and dynamics of the processes X_t^1, X_t^2 are identical to those in Table 3. Exercise opportunities arrive according to a homogeneous Poisson process with rate $\lambda \in \{2, 4\}$ as long as $\min\{X_t^1, X_t^2\} \geq 80$. The left panel provides results for random times LSMC. The truncation error, numbers of simulations and specification of the regression variables are identical to those in Table 3. In the right panel, the algorithms are the same as those in Table 6.

X_0	λ	Primal	(s.e.)	Dual	(s.e.)	Andersen	(s.e.)	PI	(s.e.)
90	2	2.0032	0.0043	2.0454	0.0333	1.9804	0.0043	2.0106	0.0062
90	4	2.6713	0.0048	2.6823	0.0063	2.5971	0.0046	2.6945	0.0083
100	2	9.7759	0.0086	9.7793	0.0087	9.4475	0.0082	9.7932	0.0173
100	4	11.8653	0.0088	11.8826	0.0096	11.1532	0.0081	11.8678	0.0221
110	2	21.7111	0.0122	21.7313	0.0138	20.6577	0.0113	21.7473	0.0303
110	4	24.8285	0.0118	24.8642	0.0133	22.9580	0.0107	24.7423	0.0348

Table 8: Value of a max-call option on asset price dependent random times ($M = 5$).

Notes. This table displays the value of a max-call option with stochastic exercise opportunities of which the arrival depend on the underlying asset prices $X_t^1, \dots, X_t^5$. The payoff and dynamics of the processes $X_t^1, \dots, X_t^5$ are identical to those in Table 4. Exercise opportunities arrive according to a homogeneous Poisson process with rate $\lambda \in \{2, 4\}$ as long as $\min_{m=1,\dots,5} X_t^m \geq 80$. The left panel provides results for random times LSMC. The truncation error and numbers of simulations are identical to those in Table 4. As regression variables, we use $\phi_i(t)(x_t^{(m)})^j$ and $(x_t^{(m)})^j (x_t^{(n)})^k$, where $i \in \{0, 1, 2, 3, 4, 5\}$, $j,\ k \in \{0, 1, 2, 3\}$ and $m,\ n \in \{1, \dots, 5\}$, $m \neq n$. In the right panel, the algorithms are the same as those in Table 6.

5.4 Policy Iteration

In this subsection, we analyze the performance of the policy iteration method developed in Section 4. For comparison purposes, we consider the same problem as in Section 5.3, where the occurrence of exercise opportunities is dependent on the evolution of the underlying asset price.

In Section 5.3, we saw that random times LSMC typically already leads to lower and upper bound estimates that generate quite narrow confidence intervals for this example. That is, the random times LSMC approach leaves little room for improvement in this example. To illustrate how policy iteration may improve upon a sub-optimal initial policy, we will therefore consider policy iteration using a random times version of the policy of Andersen [1] as initial policy. To determine this initial policy, we make use of the same truncation as that outlined at the start of Section 5, i.e., we truncate after $\bar{K}$ opportunities to stop. Formally, the Andersen method aims at constructing an optimal "threshold policy" (σ_k) due to thresholds $h_0, \dots, h_{\bar{K}} \geq 0$ of the form

$$\sigma_k = \inf\left\{k \leq j \leq \bar{K} : Z_{\tau_j} \geq h_j\right\}, \quad k = 0, \dots, \bar{K},$$

which is initialized by $h_{\bar{K}} = 0$ and $\sigma_{\bar{K}} = \bar{K}$. Then, given $h_{k+1}, \dots, h_{\bar{K}}$ and σ_{k+1}, one defines recursively,

$$\begin{aligned} h_k &:= \arg\max_{h \geq 0} \mathbb{E}\left[Z_{\tau_k} 1_{\{Z_{\tau_k} \geq h\}} + Z_{\tau_{\sigma_{k+1}}} 1_{\{Z_{\tau_k} < h\}}\right], \quad \text{and} \\ \sigma_k &:= k 1_{\{Z_{\tau_k} \geq h_k\}} + \sigma_{k+1} 1_{\{Z_{\tau_k} < h_k\}}. \end{aligned}$$

The Monte Carlo analogue of this method due to sample trajectories $n = 1, \dots, N$ is obtained as follows. Initialize $h_{\bar{K},N} = 0$ and $\sigma_{\bar{K},n} = \bar{K},\ n = 1, \dots, N$. Then, given $h_{k+1,N}, \dots, h_{\bar{K},N}$ and $\sigma_{k+1,n} \in \left\{k+1, \dots, \bar{K}\right\}$, $n = 1, \dots, N$, set

$$\begin{aligned} h_{k,N} &:= \arg\max_{h \geq 0} \frac{1}{N} \sum_{n=1}^{N} \left(Z_{\tau_k}^{(n)} 1_{\{Z_{\tau_k}^{(n)} \geq h\}} + Z_{\tau_{\sigma_{k+1,n}}}^{(n)} 1_{\{Z_{\tau_k}^{(n)} < h\}}\right), \quad \text{and} \\ \sigma_{k,n} &:= k 1_{\{Z_{\tau_k}^{(n)} \geq h_{k,N}\}} + \sigma_{k+1,n} 1_{\{Z_{\tau_k}^{(n)} < h_{k,N}\}} \quad \text{for} \quad n = 1, \dots, N. \end{aligned}$$

We implement the Andersen policy as follows. We generate $N = 200{,}000$ simulated paths ($100{,}000$ antithetic) to estimate the sequence of constant thresholds $h_0, \dots, h_{\bar{K}}$. We then simulate $2{,}000{,}000$ ($1{,}000{,}000$ antithetic) new paths, along which we follow this policy, to obtain a primal estimate of the value based on this random times Andersen policy; this ensures comparability with the LSMC policy.

For the policy iteration step, we then use $100{,}000$ simulated outer paths ($50{,}000$ antithetic). The random times Andersen policy gives us the initial stopping times (σ_i) required in Section 4.2. (Note that these follow from the above construction of the random times Andersen policy and thus do not follow from the definition in Eqn. 4.8, as we do not use an LSMC initial policy.) We set the policy iteration window parameter equal to $\kappa = \bar{K}$. We now perform one step of policy iteration with 500 sub-simulations to obtain $\widehat{\sigma}_0$ and subsequently the expected payoff of following the iterated policy.

Tables 6–8 show that a one step improvement, by the method developed in Section 4, of the very simple Andersen policy based on a threshold boundary, may lead to an improved policy that is qualitatively comparable with a policy constructed by random times LSMC using a large enough set of basis functions. If the dimension of the underlying process is relatively low, the LSMC method will usually be faster, since the simulation due to the iterated policy requires one degree of nesting. However, it is known that the LSMC method suffers from the curse of dimensionality, in contrast to the policy iteration method. For instance, in [4] it was demonstrated for an exotic cancelable coupon swap due to 20 underlying assets that the classical Longstaff-Schwartz algorithm with even a huge number of basis functions did not give satisfactory results (duality gaps of around 5$\%$ in relative terms). On the other hand, in [4] it was shown, for this example, that a one step improvement of an input policy, obtained via the classical Longstaff-Schwartz algorithm with a low number of basis functions, yields duality gaps of under 1.5$\%$ in relative terms, and moreover is up to 8 times faster. Undoubtedly, in the present context of randomly arriving opportunities, similar results can be obtained. That is, for very high dimensional underlying processes, a combination of random times LSMC with policy improvement may be much more efficient than running the LSMC algorithm alone. We consider this an interesting issue for further study, but beyond the scope of the present paper.

6 Concluding Remarks and Outlook

In this paper, we have studied optimal stopping with randomly arriving opportunities. As solution approach, we have recast the problem into a discrete-time infinite horizon problem. In this context, we have extended the least squares method, the dual martingale method, and the policy improvement approach to optimal stopping problems with infinite horizon. Naturally, in order to deal with *multiple* stopping with randomly arriving opportunities, one may follow a similar path. In particular, the corresponding approaches for multiple stopping with finite horizon in the literature, see for example [5] for regression methods, [23] for the dual martingale representation, and [3] for policy iteration, may be extended to infinite horizon problems in the spirit of this paper.

References

[1] ANDERSEN, L. (2000). A simple approach to the pricing of Bermudan swaptions in the multifactor LIBOR market model. *The Journal of Computational Finance* 3, 5–32.

[2] ANDERSEN, L., AND BROADIE, M. (2004). A primal-dual simulation algorithm for pricing multi-dimensional American options. *Management Science* 50, 1222–1234.

[3] BENDER, C., SCHOENMAKERS, J. (2006). An iterative method for multiple stopping: Convergence and stability. *Advances in Applied Probability* 38, 729–749.

[4] BENDER, C., KOLODKO, A., SCHOENMAKERS, J. (2008). Enhanced policy iteration for American options via scenario selection. *Quantitative Finance* 8(2), 135–146.

[5] CARMONA, R. AND TOUZI, N. (2008). Optimal multiple stopping and valuation of swing options. *Mathematical Finance* 18(2), 239–268.

[6] CLEMENT, E., LAMBERTON, D., AND PROTTER, P. (2002). An analysis of a least squares regression method for American option pricing. *Finance and Stochastics* 6, 449–471.

[7] DAVIS, M., AND KARATZAS, I. (1994). A deterministic approach to optimal stopping. In: Kelly, F. P. (ed.). *Probability, statistics and optimisation. A tribute to Peter Whittle*. Wiley, Chichester. Wiley Series in Probability and Mathematical Statistics. 455–466.

[8] DUPUIS, P., AND WANG, H. (2002). Optimal stopping with random intervention times. *Advances in Applied Probability* 34, 141–157.

[9] FISCHER, T. (2013). On simple representations of stopping times and stopping time sigma-algebras. *Statistics and Probability Letters* 83, 345–349.

[10] GLASSERMAN, P. (2004). *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, New York.

[11] HAUGH, M. B., AND KOGAN, L. (2004). Pricing American options: A duality approach. *Operations Research* 52, 258–270.

[12] HOBSON, D. (2021). The shape of the value function under Poisson optimal stopping. *Stochastic Processes and their Applications* 133, 229–246.

[13] HOBSON, D., ZENG, M. (2022). Constrained optimal stopping, liquidity and effort. *Stochastic Processes and their Applications* 150, 819–843.

[14] KARATZAS, I., SHREVE, S. E. (1998). *Brownian Motion and Stochastic Calculus*. Springer, Berlin.

[15] KOLODKO, A., SCHOENMAKERS, J. (2006). Iterative construction of the optimal Bermudan stopping time. *Finance and Stochastics* 10(1), 1177–1209.

[16] KRÄTSCHMER, V., LADKAU, M., LAEVEN, R., SCHOENMAKERS, J., STADJE, M. (2018). Optimal stopping under uncertainty in drift and jump intensity. *Mathematics of Operations Research* 43, 27–49.

[17] LANGE, R. J., RALPH, D., AND STORE, K. (2020). Real-option valuation in multiple dimensions using Poisson optional stopping times. *Journal of Financial and Quantitative Analysis* 55, 653–677.

[18] LEMPA, J. (2012). Optimal stopping with information constraint. *Applied Mathematics and Optimization* 66, 147–173.

[19] LONGSTAFF, F., AND SCHWARTZ, E. (2001). Valuing American options by simulation: A simple least-squares approach. *Review of Financial Studies* 14, 113–147.

[20] PESKIR, G., AND SHIRYAEV, A. N. (2006). *Optimal Stopping and Free-Boundary Problems*. Birkhäuser Verlag, Basel.

[21] PROTTER, P. (1990). *Stochastic Integration and Differential Equations: A New Approach*. Springer Verlag, Berlin.

[22] ROGERS, L. C. G. (2002). Monte Carlo valuation of American options. *Mathematical Finance* 12, 271–286.

[23] SCHOENMAKERS, J. (2012). A pure martingale dual for multiple stopping. *Finance and Stochastics* 16(2), 319–334.

[24] SNELL, L. J. (1952). Applications of martingale system theorems. *Trans. Amer. Math. Soc.* 73, 293–312.

[25] WALD, A. (1950). *Statistical Decision Functions*. Wiley, New York.