SIAM/ASA J. UNCERTAINTY QUANTIFICATION
Vol. xx, pp. x

Multilevel simulation based policy iteration for optimal stopping – convergence and complexity*

Denis Belomestny†, Marcel Ladkau‡, and John Schoenmakers§

Abstract. This paper presents a novel approach to reduce the complexity of simulation based policy iteration methods for solving optimal stopping problems. Typically, Monte Carlo construction of an improved policy gives rise to a nested simulation algorithm. In this respect our new approach uses the multilevel idea in the context of the nested simulations, where each level corresponds to a specific number of inner simulations. A thorough analysis of the convergence rates in the multilevel policy improvement algorithm is presented. A detailed complexity analysis shows that a significant reduction in computational effort can be achieved in comparison to the standard Monte Carlo based policy iteration. The performance of the multilevel method is illustrated in the case of pricing a multidimensional American derivative.

Key words. Optimal stopping, Policy iteration, Multilevel Monte Carlo

AMS subject classifications. 65C30, 65C20, 65C05, 60H35

1. Introduction.

Solving high-dimensional stopping problems in an efficient way has been a challenge for decades, particularly due to the need of pricing high-dimensional American derivatives in finance. For low or moderate dimensions, deterministic (PDE) based methods may be applicable, but for higher dimensions Monte Carlo based methods are practically the only way out. Besides the dimension independent convergence rates, Monte Carlo methods are also popular because of their generic applicability. In the late nineties several regression methods for constructing "good" exercise policies yielding lower bounds for the optimal value were introduced in the financial literature (see [8], [18], and [21], for an overview see also [12]). Among many other approaches we mention that [7] developed a stochastic mesh method, [2] introduced quantization methods, and [17] considered a class of policy iterations. In [5] it is demonstrated that the latter approach can be effectively combined with the Longstaff-Schwartz approach.

The methods mentioned above commonly provide a (generally suboptimal) exercise policy, hence a lower bound for the optimal value (or for the price of an American product). As a next breakthrough in Monte Carlo simulation of optimal stopping problems in financial context, a dual approach was developed by [20] and independently by [14], related to earlier ideas in [9]. Due to the dual formulation one considers "good" martingales rather than "good" stopping times. In fact, based on a "good" martingale the optimal value can be bounded from above by an expected path-wise maximum due to this martingale. Probably one of the most popular numerical methods for computing dual upper bounds is the method of [1]. However, this method has a drawback, namely a high computational complexity due to the need of nested

*The research by Denis Belomestny was made in IITP RAS and supported by Russian Scientific Foundation grant (project N 14-50-00150)
†Duisburg-Essen University and IITP RAS. (denis.belomestny@uni-due.de)
‡Weierstrass Institute of Applied Mathematics. (ladkau@wias-berlin.de)
§Weierstrass Institute of Applied Mathematics. (schoenma@wias-berlin.de)

Monte Carlo simulations. In a recent paper, [4] mend this problem by considering a multilevel version of the [1] algorithm.

In this paper we consider a new multilevel primal approach due to Monte Carlo based policy iteration. The basic concept of policy iteration goes back to [15] in fact (see also [19]). A detailed probabilistic treatment of a class of policy iterations (that includes Howard's one as a special case) as well as the description of the corresponding Monte Carlo algorithms is provided in [17]. In the spirit of [4] we here develop a multilevel estimator, where the multilevel concept is applied to the number of inner Monte Carlo simulations needed to construct a new policy, rather than the discretization step size of a particular SDE as in [11]. In this context we give a detailed analysis of the bias rates and the related variance rates that are crucial for the performance of the multilevel algorithm. In particular, as one main result, we provide conditions under which the bias of the estimator due to a simulation based policy improvement is of order $1/M$ with M being the number of inner simulations needed to construct the improved policy (Theorem 3.4). (cf. the bias analysis of nested simulation algorithms in portfolio risk measurement, see e.g. [13]). The proof of Theorem 3.4 is rather involved and has some flavor of large deviation theory. The amount of work (complexity) needed to compute, in the standard way, a policy improvement by simulation with accuracy ϵ is equal to $O(\epsilon^{-2-1/\gamma})$ with γ determining the bias convergence rate. As a result, the multilevel version of the algorithm will reduce the complexity by a factor of order $\epsilon^{1/(2\gamma)}$. In this paper, we restrict ourself to the case of Howard's policy iteration (improvement) for transparency, but, with no doubt the results carry over to the more refined policy iteration procedure in [17] as well.

The contents of the paper is as follows. In Section 2 we recap some results on iterative construction of optimal exercise policies from [17]. A description of the Monte Carlo based policy iteration algorithm, and a detailed convergence analysis is presented in Section 3. After a concise assessment of the complexity of the standard Monte Carlo approach in Section 4, we then introduce its multilevel version in Section 5 and provide a detailed analysis of the multilevel complexity and the corresponding computational gain with respect to the standard approach. In Section 6 we present a numerical example to illustrate the power of the multilevel approach. All proofs are deferred to Section 7 and an Appendix (on convergent Edgeworth expansions) concludes.

2. Policy iteration for optimal stopping.

In this section we review the (probabilistic) policy iteration (improvement) method for the optimal stopping problem in discrete time. For illustration, we formalize this in the context of pricing an American (Bermudan) derivative. We will work in a stylized setup where $(\Omega, \mathbb{F}, \mathbb{P})$ is a filtered probability space with discrete filtration $\mathbb{F} = (\mathcal{F}_j)_{j=0,\ldots,T}$ for $T \in \mathbb{N}_+$. An American derivative on a nonnegative adapted cash-flow process $(Z_j)_{j\geq 0}$ entitles the holder to exercise or receive cash Z_j at an exercise time $j \in \{0, ..., T\}$ that may be chosen once. It is assumed that Z_j is expressed in units of some specific pricing numeraire N with $N_0 := 1$ (w.l.o.g. we may take $N \equiv 1$). Then the value of the American option at time $j \in \{0, ..., T\}$ (in units of the numeraire) is given by the solution of the optimal stopping problem:

$$Y_j^* = \operatorname*{ess.sup}_{\tau \in \mathcal{T}[j,\ldots,T]} \mathbb{E}_{\mathcal{F}_j}[Z_\tau], \tag{2.1}$$

provided that the option is not exercised before j. In (2.1), $\mathcal{T}[j,\ldots,T]$ is the set of $\mathbb{F}$-stopping times taking values in $\{j,\ldots,T\}$ and the process $\left(Y_j^*\right)_{j\geq 0}$ is called the Snell envelope. It is well known that Y^* is a supermartingale satisfying the backward dynamic programming equation (Bellman principle)

$$Y_j^* = \max\left(Z_j, \mathbb{E}_{\mathcal{F}_j}[Y_{j+1}^*]\right), \quad 0 \leq j < T, \quad Y_T^* = Z_T.$$

An exercise policy is a family of stopping times $(\tau_j)_{j=0,\ldots,T}$ such that $\tau_j \in \mathcal{T}[j,\ldots,T]$.

Definition 2.1. *An exercise policy $(\tau_j)_{j=0,\ldots,T}$ is said to be **consistent** if*

$$\tau_j > j \Longrightarrow \tau_j = \tau_{j+1}, \quad 0 \leq j < T, \quad \text{and} \quad \tau_T = T.$$

Definition 2.2. ***(standard) policy iteration***
Given a consistent stopping family $(\tau_j)_{j=0,\ldots,T}$ we consider a new family $(\widehat{\tau}_j)_{j=0,\ldots,T}$ defined by

$$\widehat{\tau}_j = \inf\left\{k : j \leq k < T, Z_k > \mathbb{E}_{\mathcal{F}_k}[Z_{\tau_{k+1}}]\right\} \wedge T, \quad j = 0,...,T \tag{2.2}$$

with $\wedge$ denoting the minimum operator and $\inf\varnothing := +\infty$. The new family $(\widehat{\tau}_j)$ is termed a policy iteration of (τ_j).

Definition 2.3. *Let us introduce $\widehat{Y}_j := \mathbb{E}_{\mathcal{F}_j}[Z_{\widehat{\tau}_j}]$ and $Y_j = \mathbb{E}_{\mathcal{F}_j}[Z_{\tau_j}]$.*

The basic idea behind (2.2) goes back to [15] (see also [19]). The key issue is that (2.2) is actually a *policy improvement* due to the following theorem.

Theorem 2.4.

(i) *It holds that*

$$Y_j^* \geq \widehat{Y}_j \geq Y_j, \quad j = 0,...,T. \tag{2.3}$$

(ii) *If $\tau_j^{(0)} := \tau_j$, $\tau_j^{(m+1)} := \widehat{\tau_j^{(m)}}$ (cf. (2.2)), $Y_j^{(0)} := Y_j$, $Y_j^{(m)} := \mathbb{E}_{\mathcal{F}_j}[Z_{\tau_j^{(m)}}]$, $j = 0,...,T$, $m = 0,1,2,...,$ then*

$$Y_k^{(T-j)} = Y_k^*, \quad k = j,...,T.$$

Theorem 2.4 is in fact a corollary of Th. 3.1 and Prop. 4.3 in [17], where a detailed analysis is provided for a whole class of policy iterations of which (2.2) is a special case. See also [6] for a further analysis regarding stability issues, and extensions to policy iteration methods for multiple stopping. Due to Theorem 2.4, one may iterate any consistent policy in finitely many steps to the optimal one. Moreover, the respective (lower) approximations to the Snell envelope converge in a nondecreasing manner.

3. Simulation based policy iteration.

In order to apply the policy iteration method in practice, we henceforth assume that the cash-flow Z_j is of the form (while slightly abusing of notation) $Z_j = Z_j(X_j)$ for some underlying (possibly high-dimensional) Markovian process X. As a consequence, the Snell envelope process then has the Markovian form $Y_j^* = Y_j^*(X_j), \quad j = 0,...,T$, as well. Furthermore, it is assumed that a consistent stopping family (τ_j) depends on ω only through the path $X.$ in the following way: For each j the event $\{\tau_j = j\}$ is measurable w.r.t. X_j, and τ_j is measurable w.r.t. $(X_k)_{j\leq k\leq T}$, i.e.

$$\tau_j(\omega) = h_j(X_j(\omega),...,X_T(\omega)) \tag{3.1}$$

for some Borel measurable function h_j. A typical example of such a stopping family is

$$\tau_j = \inf\{k : j \le k \le T, \quad Z_k(X_k) \ge f_k(X_k)\}$$

for a set of real valued functions $f_k(x)$. The next issue is the estimation of the conditional expectations in (2.2). A canonical approach is the use of sub simulations. In this respect we consider an enlarged probability space $(\Omega, \mathbb{F}', \mathbb{P})$, where $\mathbb{F}' = (\mathcal{F}'_j)_{j=0,\ldots,T}$ and $\mathcal{F}_j \subset \mathcal{F}'_j$ for each j. By assumption, $\mathcal{F}'_j$ specified as

$$\mathcal{F}'_j = \mathcal{F}_j \vee \sigma\left\{X_\cdot^{i,X_i},\ i \le j,\right\} \text{ with } \mathcal{F}_j = \sigma\left\{X_i,\ i \le j\right\},$$

where for a generic $(\omega, \omega_{in}) \in \Omega$, $X_\cdot^{i,X_i} := X_k^{i,X_i(\omega)}(\omega_{in})$, $k \ge i$ denotes a sub trajectory starting at time i in the state $X_i(\omega) = X_i^{i,X_i(\omega)}$ of the outer trajectory $X(\omega)$. In particular, the random variables $X_\cdot^{i,X_i}$ and $X_\cdot^{i',X_{i'}}$ are by assumption independent, conditionally $\{X_i, X_{i'}\}$, for $i \neq i'$. On the enlarged space we consider $\mathcal{F}'_j$ measurable estimations $\mathcal{C}_{j,M}$ of $C_j := \mathbb{E}_{\mathcal{F}_j}\left[Z_{\tau_{j+1}}\right]$ as being standard Monte Carlo estimates based on M sub simulations. More precisely, for

$$C_j(X_j) := \mathbb{E}_{X_j}\left[Z_{\tau_{j+1}}\right]$$

define

$$\mathcal{C}_{j,M} := \frac{1}{M}\sum_{m=1}^{M} Z_{\tau_{j+1}^{(m)}}(X_{\tau_{j+1}^{(m)}}^{j,X_j,(m)}),$$

where the stopping times

$$\tau_{j+1}^{(m)} := h_{j+1}(X_{j+1}^{j,X_j,(m)}, ..., X_T^{j,X_j,(m)})$$

(cf. (3.1)) are evaluated on sub-trajectories $X_\cdot^{j,X_j,(m)}$, $m = 1, ..., M$, all starting at time j in X_j. Obviously, $\mathcal{C}_{j,M}$ is an unbiased estimator for C_j with respect to $\mathbb{E}_{\mathcal{F}_j}[\cdot]$. We thus end up with a simulation based version of (2.2),

$$\widehat{\tau}_{j,M} = \min\left\{k : j \le k < T,\ Z_k > \mathcal{C}_{k,M}\right\} \wedge T.$$

Now set

$$\widehat{Y}_{j,M} := \mathbb{E}_{\mathcal{F}_j}[Z_{\widehat{\tau}_{j,M}}].$$

Next we analyze the bias and the variance of the estimator $\widehat{Y}_{0,M}$.

Proposition 3.1. *Suppose that $|Z_j| < B$ for some $B > 0$. Let us further assume that there exist a constant $D > 0$ and $\alpha > 0$, such that for any $\delta > 0$ and $j = 0, \ldots, T-1$,*

$$\mathbb{P}(|C_j - Z_j| \le \delta) \le D\delta^{\alpha}. \tag{3.2}$$

It then holds,

$$\mathbb{P}\left(\widehat{\tau}_{0,M} \neq \widehat{\tau}_0\right) \le D_1 M^{-\alpha/2} \tag{3.3}$$

for some constant $D_1 > 0$.

Corollary 3.2. *Under the assumptions of Proposition 3.1, it follows immediately by (3.3) that*

$$\widehat{Y}_{0,M} - \widehat{Y}_0 = O(M^{-\alpha/2}) \quad \text{and} \quad \mathbb{E}[\left(Z_{\widehat{\tau}_{0,M}} - Z_{\widehat{\tau}_0}\right)^2] = O(M^{-\alpha/2}).$$

Proof. We have

$$\widehat{Y}_{0,M} - \widehat{Y}_0 = \mathbb{E}[\left(Z_{\widehat{\tau}_{0,M}} - Z_{\widehat{\tau}_0}\right) \mathbf{1}_{\{\widehat{\tau}_{0,M} \neq \widehat{\tau}_0\}}] \leq B\mathbb{P}\left(\widehat{\tau}_{0,M} \neq \widehat{\tau}_0\right)$$

and

$$\mathbb{E}[\left(Z_{\widehat{\tau}_{0,M}} - Z_{\widehat{\tau}_0}\right)^2] = \mathbb{E}[\left(Z_{\widehat{\tau}_{0,M}} - Z_{\widehat{\tau}_0}\right)^2 \mathbf{1}_{\{\widehat{\tau}_{0,M} \neq \widehat{\tau}_0\}}] \leq B^2\mathbb{P}\left(\widehat{\tau}_{0,M} \neq \widehat{\tau}_0\right).$$

■

Remark 3.3. *The boundedness condition $|Z_j| < B$ is made for the sake of simplicity. This condition can be replaced by a kind of moment condition as in Theorem 7 of [3].* Under somewhat more restrictive assumptions than the ones of Proposition 3.1 we can prove the following theorem.

Theorem 3.4. *Suppose that*

(i) *the transition densities $\mathfrak{p}_j\left(x;y\right)$ of the chain X_j, given $X_{j-1} = y$, are living an open domain, do not vanish, have bounded derivatives of any order in x, y, and are such that $\partial_x \mathfrak{p}_j\left(x;y\right)$ is integrable w.r.t. x ;*

(ii) *the cash-flow is bounded, i.e. there exists a constant B such that $|Z_j(x)| < B$ a.s. for all x;*

(iii) *the function*

$$\sigma_j^2(x) := \mathbb{E}\left[\left(Z_{\tau_{j+1}}(X^{j,x}_{\tau_{j+1}}) - C_j(x)\right)^2\right] = Var\left[Z_{\tau_{j+1}}(X^{j,x}_{\tau_{j+1}})\right]$$

is bounded (due to (i)) and bounded away from zero uniformly in x and j;

(iv) *the density of the random variable*

$$Z_j(X_j) - C_j(X_j)$$

conditional on $\mathcal{F}_{j-1}$, i.e. given $X_{j-1} = x_{j-1}$, is of the form $x \to h(x; x_{j-1})$, where $h(\cdot; x_{j-1})$ is at least two times differentiable for each x_{j-1}.

(v) *on the j-th exercise boundary, i.e. on the set $\{x : Z_j(x) = C_j(x)\}$, the gradient of the function*

$$\left(Z_j(x) - C_j(x)\right)/\sigma_j^2(x)$$

exist, is bounded with bounded continuous derivatives, and does not vanish.

Then it holds

$$\left|\widehat{Y}_{0,M} - \widehat{Y}_0\right| = O(M^{-1}), \quad M \to \infty.$$

Discussion. Theorem 3.4 controls the bias of the estimator $\widehat{Y}_{0,M}$ for the lower approximation $\widehat{Y}_0$ to the Snell envelope due to the improved policy $(\widehat{\tau}_j)$. Concerning the difference between $\widehat{Y}_0$ and Y_0, we infer from [17], Lemma 4.5, that

$$0 \leq \widehat{Y}_0 - Y_0 \leq \mathbb{E} \sum_{k=\tau_0}^{\widehat{\tau}_0 - 1} [\mathbb{E}_{\mathcal{F}_k} Y_{k+1} - Y_k]$$

(where automatically $\widehat{\tau}_0 \geq \tau_0$ when (τ_j) is consistent). Hence, for a bounded cash-flow process with $|Z_j| < B$ we get

$$0 \leq \widehat{Y}_0 - Y_0 \leq TB\mathbb{P}(\tau_j \neq \widehat{\tau}_j) \leq TB\mathbb{P}(\tau_j \neq \tau_j^*),$$

as $\tau_j = \tau_j^*$ implies $\tau_j = \widehat{\tau}_j = \tau_j^*$. If $\mathbb{P}(\tau_j \neq \tau_j^*) = 0$, we get $Y_0 = \widehat{Y}_0 = Y_0^*$.

4. Standard Monte Carlo approach. Within Markovian setup as introduced in Section 3, consider for some fixed natural numbers N and M, the estimator:

$$\widehat{\mathcal{Y}}_{N,M} := \frac{1}{N} \sum_{n=1}^{N} Z_{\widehat{\tau}_M}^{(n)} \tag{4.1}$$

for $\widehat{Y}_M := \widehat{Y}_{0,M}$ with $\widehat{\tau}_M := \widehat{\tau}_{0,M}$, based on n realizations $Z_{\widehat{\tau}_M}^{(n)}$, $n = 1, ..., N$, of the stopped cash-flow $Z_{\widehat{\tau}_M}$. Let us investigate the complexity, i.e. the required computational costs, in order to compute $\widehat{Y} := \widehat{Y}_0$ with a prescribed (root-mean-square) accuracy ϵ, by using the estimator (4.1). Under the assumptions of Corollary 3.2 we have with $\gamma = \alpha/2$, or $\gamma = 1$ if Theorem 3.4 applies, for the mean squared error,

$$\begin{aligned} \mathbb{E}\left[\widehat{\mathcal{Y}}_{N,M} - \widehat{Y}\right]^2 &= N^{-1} \mathrm{Var}\left[Z_{\widehat{\tau}_M}\right] + \left|\widehat{Y} - \widehat{Y}_M\right|^2 \\ &\lesssim N^{-1}\sigma_\infty^2 + \mu_\infty^2 M^{-2\gamma}, \quad M \to \infty, \end{aligned} \tag{4.2}$$

for some constants μ_∞ and $\sigma_\infty^2 := \mathrm{limsup}_{M\to\infty} \mathrm{Var}\left[Z_{\widehat{\tau}_M}\right]$, where M_0 denotes some fixed minimum number of sub trajectories used for computing the stopping time $\widehat{\tau}_M$. In order to bound (4.2) by ϵ^2, we set

$$M = \left\lceil \left(\frac{2^{1/2}\mu_\infty}{\epsilon}\right)^{1/\gamma} \right\rceil, \quad N = \left\lceil \frac{2\sigma_\infty^2}{\epsilon^2} \right\rceil$$

with $\lceil x \rceil$ denoting the smallest integer bigger or equal than x. For notational simplicity we will henceforth omit the brackets and carry out calculations with generally non-integer M, N. This will neither affect complexity rates nor the asymptotic proportionality constants. Thus the computational complexity for reaching accuracy ϵ when $\epsilon \downarrow 0$ is given by

$$\mathcal{C}_{\mathrm{stand}}^{N,M}(\epsilon) := NM = \frac{2\sigma_\infty^2 \left(2^{1/2}\mu_\infty\right)^{1/\gamma}}{\epsilon^{2+1/\gamma}}, \tag{4.3}$$

where, again for simplicity, it is assumed that both the cost of simulating one outer trajectory and one sub trajectory is equal to one unit. In typical applications we have $\gamma = 1$ and the complexity of the standard Monte Carlo method is of order $O(\epsilon^{-3})$. However, if $\gamma = 1/2$ the complexity is as high as $O(\epsilon^{-4})$.

5. Multilevel Monte Carlo approach. For a fixed natural number L and a sequence of natural numbers $\mathbf{m} := (m_0, ..., m_L)$ satisfying $1 \leq m_0 < ... < m_L$, we consider in the spirit of [11] the telescoping sum:

$$\widehat{Y}_{m_L} = \widehat{Y}_{m_0} + \sum_{l=1}^{L} \left(\widehat{Y}_{m_l} - \widehat{Y}_{m_{l-1}} \right). \tag{5.1}$$

Further we approximate the expectations $\widehat{Y}_{m_l}$ in (5.1). We take a set of natural numbers $\mathbf{n} := (n_0, ..., n_L)$ satisfying $n_0 > ... > n_L \geq 1$, and simulate the initial set of cash-flows

$$\left\{ Z^{(j)}_{\widehat{\tau}_{m_0}}, \quad j = 1, ..., n_0 \right\},$$

due to the initial set of trajectories $X^{0,x,(j)}_{\cdot}$, $j = 1, ..., n_0$. Next we simulate *independently* for each level $l = 1, ..., L$, a set of pairs

$$\left\{ (Z^{(j)}_{\widehat{\tau}_{m_l}}, Z^{(j)}_{\widehat{\tau}_{m_{l-1}}}), \quad j = 1, ..., n_l \right\}$$

due to a set of trajectories $X^{0,x,(j)}_{\cdot}$, $j = 1, ..., n_l$, to obtain a multilevel estimator

$$\widehat{\mathcal{Y}}_{\mathbf{n},\mathbf{m}} := \frac{1}{n_0} \sum_{j=1}^{n_0} Z^{(j)}_{\widehat{\tau}_{m_0}} + \sum_{l=1}^{L} \frac{1}{n_l} \sum_{j=1}^{n_l} \left(Z^{(j)}_{\widehat{\tau}_{m_l}} - Z^{(j)}_{\widehat{\tau}_{m_{l-1}}} \right) \tag{5.2}$$

as an approximation to $\widehat{Y}$ (cf. [4]). Henceforth we always take $\mathbf{m}$ to be a geometric sequence, i.e., $m_l = m_0 \kappa^l$, for some $m_0, \kappa \in \mathbb{N}$, $\kappa \geq 2$.

Complexity analysis. Let us now study the complexity of the multilevel estimator (5.2) under the assumption that the conditions of Proposition 3.1 or Theorem 3.4 are fulfilled. For the bias we have

$$\left| \mathbb{E}\left[\widehat{\mathcal{Y}}_{\mathbf{n},\mathbf{m}} \right] - \widehat{Y} \right| = \left| \mathbb{E}\left[Z_{\widehat{\tau}_{m_L}} - Z_{\widehat{\tau}} \right] \right| \leq \mu_\infty m_L^{-\gamma}, \tag{5.3}$$

and for the variance it holds

$$\mathrm{Var}\left[\widehat{\mathcal{Y}}_{\mathbf{n},\mathbf{m}} \right] = \frac{1}{n_0} \mathrm{Var}\left[Z_{\widehat{\tau}_{m_0}} \right] + \sum_{l=1}^{L} \frac{1}{n_l} \mathrm{Var}\left[Z_{\widehat{\tau}_{m_l}} - Z_{\widehat{\tau}_{m_{l-1}}} \right],$$

where due to Proposition 3.1, the terms with $l > 0$ may be estimated by

$$\begin{aligned} \mathrm{Var}\left[Z_{\widehat{\tau}_{m_l}} - Z_{\widehat{\tau}_{m_{l-1}}} \right] &\leq \mathbb{E}\left[\left(Z_{\widehat{\tau}_{m_l}} - Z_{\widehat{\tau}_{m_{l-1}}} \right)^2 \right] \\ &\leq 2\mathbb{E}\left[\left(Z_{\widehat{\tau}_{m_l}} - Z_{\widehat{\tau}} \right)^2 \right] + 2\mathbb{E}\left[\left(Z_{\widehat{\tau}_{m_{l-1}}} - Z_{\widehat{\tau}} \right)^2 \right] \\ &\leq C \left(m_l^{-\beta} + m_{l-1}^{-\beta} \right) \leq C m_l^{-\beta} \left(1 + \kappa^\beta \right) \leq \mathcal{V}_\infty m_l^{-\beta}, \end{aligned} \tag{5.4}$$

with $\beta := \alpha/2$, and suitable constants C, $\mathcal{V}_\infty$. In typical applications, we have that $C_j - Z_j$ in (3.2) has a positive but non-exploding density in zero which implies $\alpha = 1$, hence $\beta = 1/2$. This rate is confirmed by numerical experiments. Henceforth, we assume $\beta < 1$.

We are now going to analyze the optimal complexity of the multilevel algorithm. Our optimization approach is based on a separate treatment of n_0 and $n_i,\ i = 1, ..., L$. In particular, we assume that

$$n_l = n_1 \kappa^{(1+\beta)/2 - l(1+\beta)/2}, \quad 1 \le l \le L,$$

where the integers n_0 and n_1 are to be determined, and for the sub-simulations we take

$$m_l = m_0 \kappa^l, \quad 0 \le l \le L.$$

We further reuse the sub-simulations related to m_{l-1} for the computation of $\widehat{Y}_{m_l}$ so that the multilevel complexity becomes

$$\begin{aligned} \mathcal{C}_{ML}^{\mathbf{n},\mathbf{m}} &= n_0 m_0 + \sum_{l=1}^{L} n_l m_l \\ &= n_0 m_0 + n_1 m_0 \kappa \frac{\kappa^{L(1-\beta)/2} - 1}{\kappa^{(1-\beta)/2} - 1}. \end{aligned} \tag{5.5}$$

Theorem 5.1. *The asymptotic complexity of the multilevel estimator $\widehat{\mathcal{Y}}_{\mathbf{n},\mathbf{m}}$ for $0 < \beta < 1$ is given by*

$$\begin{aligned} &\mathcal{C}_{ML}^{*} := \mathcal{C}_{ML}\left(n_0^*, n_1^*, L^*, m_0, \epsilon\right) := \\ &\frac{(1-\beta)\,\mathcal{V}_\infty \mu_\infty^{(1-\beta)/\gamma}}{2\gamma\left(1-\kappa^{-(1-\beta)/2}\right)^2}\left(1 + 2\gamma/\left(1-\beta\right)\right)^{1+(1-\beta)/(2\gamma)}\left(1 + O\left(\epsilon^{(1-\beta)/(2\gamma)}\right)\right)\epsilon^{-2-(1-\beta)/\gamma}, \end{aligned} \tag{5.6}$$

where the optimal values n_0^*, n_1^*, L^* have to be chosen according to

$$\begin{aligned} &n_0^* := n_0^*\left(L^*, m_0, \epsilon\right) := \\ &\frac{\sigma_\infty \mathcal{V}_\infty^{1/2} \mu_\infty^{(1-\beta)/(2\gamma)}\left(1-\beta\right)}{2\gamma m_0^{1/2}\left(1-\kappa^{-(1-\beta)/2}\right)}\left(1 + 2\gamma/\left(1-\beta\right)\right)^{1+(1-\beta)/(4\gamma)} \times \\ &\epsilon^{-2-(1-\beta)/(2\gamma)}\left(1 + O\left(\epsilon^{(1-\beta)/(2\gamma)}\right)\right) \quad \textit{and,} \end{aligned} \tag{5.7}$$

$$\begin{aligned} &n_1^* := n_1^*\left(L^*, m_0, \epsilon\right) := \\ &\frac{\mathcal{V}_\infty \mu_\infty^{(1-\beta)/(2\gamma)}\left(1-\beta\right)}{2\gamma m_0^{(1+\beta)/2}\left(1-\kappa^{-(1-\beta)/2}\right)}\left(1 + 2\gamma/\left(1-\beta\right)\right)^{1+(1-\beta)/(4\gamma)}\kappa^{-(1+\beta)/2} \times \\ &\epsilon^{-2-(1-\beta)/(2\gamma)}\left(1 + O\left(\epsilon^{(1-\beta)/(2\gamma)}\right)\right) \quad \textit{and,} \end{aligned} \tag{5.8}$$

$$L^* := \frac{\ln \epsilon^{-1} + \ln\left[\frac{\mu_\infty}{m_0^\gamma}\left(1 + 2\gamma/\left(1-\beta\right)\right)^{1/2}\right]}{\gamma \ln \kappa} + O\left(\epsilon^{(1-\beta)/(2\gamma)}\right). \tag{5.9}$$

Note that, asymptotically, the optimal complexity $\mathcal{C}_{ML}^*$ is independent of m_0. We therefore propose to choose m_0 by experience. In typical numerical examples $m_0 = 100$ turns out to be a robust choice.

Discussion. For the standard algorithm given optimally chosen M^*, N^* we have the complexity given by (4.3), so the gain ratio of the multilevel approach over the standard Monte Carlo algorithm is asymptotically given by

$$\begin{aligned}\mathcal{R}^*(\epsilon) &:= \frac{\mathcal{C}^*_{ML}(\epsilon)}{\mathcal{C}^{N^*,M^*}_{\text{stan}}(\epsilon)} \\ &\sim \frac{(1-\beta)\left(1+2\gamma/(1-\beta)\right)^{1+(1-\beta)/(2\gamma)}\mathcal{V}_\infty}{2^{2+1/(2\gamma)}\gamma\left(1-\kappa^{-(1-\beta)/2}\right)^2\sigma^2_\infty\mu_\infty^{\beta/\gamma}}\epsilon^{\beta/\gamma}, \quad \epsilon\downarrow 0.\end{aligned} \tag{5.10}$$

For the variance and bias rate β and γ, respectively, cf. (5.4) and (5.3). Typically, we have that $\beta = 1/2$ and that $\gamma \geq 1/2$, where the value of γ depends on whether Theorem 3.4 applies or not. In any case we may conclude that the smaller γ the larger the complexity gain.

6. Numerical comparison of the two estimators. In this section we will compare both algorithms in a numerical example. The usual way would be to take both algorithms, take optimal parameters and compare the complexities given an accuracy ϵ, like we did in the previous section in general. The optimal parameters depend on knowledge of some quantities, e.g. the coefficients of the bias rates. This knowledge might be gained by pre-computation (based on relatively smaller sample sizes) for instance. Here we propose a more pragmatic and robust approach (cf. [4]).

Let us assume that a practitioner knows his standard algorithm well and provides us with his "optimal" M (inner simulations), N (outer simulations). So his computational budget amounts to MN. Given the same budget MN we are now going to configure the multilevel estimator such that $m_L = M$, i.e. the bias is the same for both algorithms. We next show that n_0, n_1, and L can be chosen in such a way that the variance of the multilevel estimator is significantly below the variance of the standard one. Although this approach will not achieve the optimal gain (5.10) for $\epsilon \downarrow 0$ (hence for $M \to \infty$), it has the advantage that we may compare the accuracy of the multilevel estimator with the standard one for any fixed M and arbitrary N. The details are spelled out below.

Taking

$$M = m_L = m_0\kappa^L \tag{6.1}$$

we have for the biases

$$\mathbb{E}\left[\widehat{\mathcal{Y}}_{\mathbf{n},\mathbf{m}} - \widehat{Y}\right] = \mathbb{E}\left[\widehat{\mathcal{Y}}_{N,M} - \widehat{Y}\right] \leq \frac{\mu_\infty}{M^\gamma}.$$

As stated above we assume the same computational budget for both algorithms leading to the following constraint (see (5.5))

$$NM = n_0m_0 + n_1m_0\kappa\frac{\kappa^{L(1-\beta)/2}-1}{\kappa^{(1-\beta)/2}-1}.$$

Let us write for $\xi \in \mathbb{R}_+$,

$$\begin{aligned}n_1 &:= \xi n_0, \\ n_0 &= \frac{NM}{m_0 + \xi m_0\kappa\frac{\kappa^{L(1-\beta)/2}-1}{\kappa^{(1-\beta)/2}-1}}.\end{aligned} \tag{6.2}$$

With (6.1) and (6.2) we have for the variance estimate (7.6)

$$
\begin{aligned}
\operatorname{Var}\left[\widehat{\mathcal{Y}}_{\mathbf{n},\mathbf{m}}\right] &\leq \frac{\sigma_\infty^2}{n_0} + \frac{\mathcal{V}_\infty \kappa^{-\beta}}{\xi n_0 M^\beta \kappa^{-\beta L}} \frac{\kappa^{L(1-\beta)/2}-1}{\kappa^{(1-\beta)/2}-1} \\
&= \frac{\sigma_\infty^2 \kappa^{-L}}{N} \left(1 + \frac{\mathcal{V}_\infty \kappa^{-\beta+\beta L}}{\xi M^\beta \sigma_\infty^2} \frac{\kappa^{L(1-\beta)/2}-1}{\kappa^{(1-\beta)/2}-1}\right) \\
&\quad \times \left(1 + \xi\kappa \frac{\kappa^{L(1-\beta)/2}-1}{\kappa^{(1-\beta)/2}-1}\right) \\
&= \frac{\sigma_\infty^2 \kappa^{-L}}{N} \left(1 + \frac{a}{\xi}\right)(1 + b\xi)
\end{aligned}
\tag{6.3}
$$

Expression (6.3) attains its minimum at

$$
\xi^\circ := \sqrt{\frac{a}{b}} = \frac{\mathcal{V}_\infty^{1/2} \kappa^{(-\beta-1+\beta L)/2}}{M^{\beta/2}\sigma_\infty},
\tag{6.4}
$$

which gives the "optimal" values n_0° and n_1° via (6.2), and

$$
\begin{aligned}
\operatorname{Var}\left[\widehat{\mathcal{Y}}_{\mathbf{n}^\circ,\mathbf{m}}\right] &\leq \frac{\sigma_\infty^2 \kappa^{-L}}{N} \left(1 + \sqrt{ab}\right)^2 \\
&= \frac{\sigma_\infty^2 \kappa^{-L}}{N} \left(1 + \frac{\kappa^{L(1-\beta)/2}-1}{\kappa^{(1-\beta)/2}-1} \frac{\mathcal{V}_\infty^{1/2} \kappa^{(1-\beta+\beta L)/2}}{M^{\beta/2}\sigma_\infty}\right)^2.
\end{aligned}
$$

The ratio of the corresponding standard deviations is thus given by

$$
\begin{aligned}
\mathcal{R}^\circ(M, L) &= \frac{\sqrt{\operatorname{Var}\left[\widehat{\mathcal{Y}}_{\mathbf{n}^\circ,\mathbf{m}}\right]}}{\sqrt{\operatorname{Var}\left[\widehat{\mathcal{Y}}_{N,M}\right]}} \\
&= \kappa^{-L/2} + \frac{\mathcal{V}_\infty^{1/2}}{M^{\beta/2}\sigma_\infty} \frac{1-\kappa^{-(1-\beta)L/2}}{1-\kappa^{-(1-\beta)/2}}.
\end{aligned}
\tag{6.5}
$$

Note that the ratio (6.5) is independent of N. By setting the derivative of (6.5) w.r.t. L equal to zero we solve,

$$
L^\circ := \frac{2}{\beta \ln \kappa} \ln\left[\frac{M^{\beta/2}\sigma_\infty}{\mathcal{V}_\infty^{1/2}(1-\beta)} \left(1 - \kappa^{-(1-\beta)/2}\right)\right].
\tag{6.6}
$$

Since $L^\circ > 0$, we require

$$
M > \left(\frac{\mathcal{V}_\infty^{1/2}(1-\beta)}{\sigma_\infty\left(1-\kappa^{-(1-\beta)/2}\right)}\right)^{2/\beta}.
\tag{6.7}
$$

It is easy to see that (6.5) attains its minimum for L° given by (6.6) and M satisfying (6.7). It then holds $\mathcal{R}^\circ(M, L^\circ) < 1$, hence the multilevel estimator outperforms the standard in terms of the variance.

Remark 6.1.*Suppose the practitioner using the standard algorithm makes up his mind and changes his choice of N to N', connected with the number of inner simulations M. He so chooses a new budget $M \times N'$ say. Then with this new budget we can adapt the parameters accordingly, yielding the same variance reduction (6.5) with the same (6.6), as the latter are independent of N.*

6.1. Numerical example: American max-call. We now proceed to a numerical study of multilevel policy iteration in the context of American max-call option based on d assets. Each asset is assumed to be governed by the following SDE

$$dS_t^i = (r - \delta) S_t^i dt + \sigma S_t^i dW_t^i, \quad i = 1, ..., d,$$

under the risk- neutral measure, where $\left(W_t^1, ..., W_t^d\right)$ is a d-dimensional standard Brownian motion. Further, $T_0, T_1, ..., T_n$ are equidistant exercise dates between $T_0 = 0$ and T_n. For notational convenience we shall write S_j instead of S_{T_j}. The discounted cash-flow process of the option is specified by

$$Z_k = e^{-rk}\left(\max_{i=1,...,d} S_k^i - K\right)^+.$$

We take the following benchmark parameter values (see [1])

$$r = 0.05, \quad \sigma = 0.2, \quad \delta = 0.1, \quad K = 100, \quad d = 5, \quad n = 9, \quad T_n = 3$$

and $S_0^i = 100$, $i = 1, ..., d$. For the input stopping family $(\tau_j)_{0 \le j \le T}$ we take

$$\tau_j = \inf\left\{k : j \le k < T : Z_k > \mathbb{E}_{\mathcal{F}_k}\left[Z_{k+1}\right]\right\} \wedge T,$$

where $\mathbb{E}_{\mathcal{F}_k}\left[Z_{k+1}\right]$ is the (discounted) value of a still-alive one period European option. The value of a European max-call option can be computed via the Johnson's formula (1987) ([16]),

$$\begin{aligned}
&\mathbb{E}\left[e^{-rT}\left(\max_{i=1,...,5} S_T^i - K\right)^+\right] \\
&= \sum_{i=1}^{5} S_0^i \frac{e^{-\delta T}}{\sqrt{2\pi}} \int_{-\infty}^{d_+^i} \exp\left[-\frac{1}{2}z^2\right] \prod_{i'=1, i' \neq i}^{5} \mathcal{N}\left(\frac{\ln\left(\frac{S_0^i}{S_0^{i'}}\right)}{\sigma\sqrt{k}} - z + \sigma\sqrt{T}\right) dz \\
&- Ke^{-rT} + Ke^{-rT} \prod_{i=1}^{5}\left(1 - \mathcal{N}\left(d_-^i\right)\right)
\end{aligned}$$

with

$$d_+^i := \frac{\ln\left(\frac{S_0^i}{K}\right) + \left(r - \delta + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}, \quad d_-^i = d_+^i - \sigma\sqrt{T}.$$

Figure 6.1. *The SD ratio function $\mathcal{R}^\circ(M,L)$ for different M, measuring the variance reduction due to the ML approach.*

For evaluating the integrals we use an adaptive Gauss-Kronrod procedure (with 31 points).

For this example we follow the approach of Section 6. We see that the final gain (6.5) due to the multilevel approach depends on κ as well. Our general experience is that an "optimal" κ for our method is typically larger than two. In this example we took $\kappa = 5$. A pre-simulation based on 10^3 trajectories yield the following estimates,

$$\begin{aligned} &\gamma = 1, \quad \beta = 0.5, \\ &\mathrm{Var}\left[Z_{\widehat{\tau}_m}\right] =: \sigma^2(m) \le \sigma_\infty^2 = 350, \\ &\sqrt{m_l}\mathrm{Var}\left[Z_{\widehat{\tau}_{m_l}} - Z_{\widehat{\tau}_{m_{l-1}}}\right] \le \mathcal{V}_\infty = 645, \end{aligned} \tag{6.8}$$

where we used antithetic sampling in (6.8). This yields Figure 6.1, where $\mathcal{R}(M,L)$ is plotted for different M as a function of L. For each particular M one may read off the optimal value of L° from this figure.

Assume, for example, that the user of the standard algorithm decides to calculate the value of the option with $M = 7500$ inner trajectories. From Figure 6.1 we see that $L = 4$ is for this M the best choice (that doesn't depend on N). For the present illustration we take $N = 1000$ and then compute n_0°, n_1° from (6.2) and (6.4), where $\mathcal{V}_\infty$ is replaced by the estimate

$$\max_{l=1,\ldots,4}\left\{\sqrt{m_l}\widehat{v}(m_l,m_{l-1})\right\}$$

with

$$\widehat{v}(m_l,m_{l-1}) := \frac{1}{n}\sum_{r=1}^{n}\left[Z_{\widehat{\tau}_{m_l}}^{(r)} - Z_{\widehat{\tau}_{m_{l-1}}}^{(r)} - \left(\overline{Z_{\widehat{\tau}_{m_l}} - Z_{\widehat{\tau}_{m_{l-1}}}}\right)\right]^2,$$

for $n = 10^3$ and the bar denoting the corresponding sample average, where antithetic variables are used in the simulation of inner trajectories. Let us further define

$$\widehat{v}(m,0) := \frac{1}{n}\sum_{r=1}^{n}\left[Z_{\widehat{\tau}_m}^{(r)} - \overline{Z_{\widehat{\tau}_m}}\right]^2$$

with $n = 10^3$ again. Table 1 shows the resulting values n_l°, the approximative level variances $\widehat{v}(m_l,m_{l-1})$, $l = 1,\ldots,4$, as well as the option prices estimates. As can be seen from the table, the variance of the multilevel estimate $\widehat{\mathcal{Y}}_{\mathbf{n}^\circ,\mathbf{m}}$ with the "optimal" choice $L^\circ = 4$ (cf. (6.6) and Figure 6.1) is significantly smaller than the variance of the standard Monte Carlo estimate $\widehat{Y}_{1000,7500}$.

Concluding remarks. One may argue that the variance reduction demonstrated in the above example looks not too spectacular. In this respect we underline that this variance reduction is obtained via a pragmatic approach (Section 6), where detailed knowledge of the optimal allocation of the standard algorithm (in particular the precise decay of the bias) is not necessary. However, in a situation where the bias decay is additionally known (from

Table 1

The performance of the ML estimator with the optimal choice of n_l°, $l = 0, \ldots, 4$, compared to standard policy iteration

l	n_l°	m_l	$\frac{1}{n_l^\circ}\sum_{n=1}^{n_l^\circ}\left[Z^{(n)}_{\widehat{\tau}_{m_l}} - Z^{(n)}_{\widehat{\tau}_{m_{l-1}}}\right]$	$\widehat{v}\,(m_l, m_{l-1})$
0	47368	12	25.5772	350
1	5223	60	0.0668629	53.4224
2	1847	300	−0.0623856	37.2088
3	653	1500	0.201612	15.8769
4	231	7500	−0.0319232	5.19074
			$\widehat{\mathcal{Y}}_{\mathbf{n}^\circ,\mathbf{m}} = 25.7513661$	$sd\left(\widehat{\mathcal{Y}}_{\mathbf{n}^\circ,\mathbf{m}}\right) =$ 0.2820887804
ST	$N =$ 1000	$M =$ 7500	$\widehat{Y}_{N,M} = 25.2373$	$sd\left(\widehat{Y}_{N,M}\right) =$ 0.5899033819

some additional pre-computation for example), one may may parameterize the multilevel algorithm following the asymptotic complexity analysis in Section 5, and thus end up with an (asymptotically) optimized complexity gain (5.10) that blows up when the required accuracy gets smaller and smaller.

7. Proofs.

7.1. Proof of Proposition 3.1. Let us write $\{\widehat{\tau}_{0,M} \neq \widehat{\tau}_0\} = \{\widehat{\tau}_{0,M} > \widehat{\tau}_0\} \cup \{\widehat{\tau}_{0,M} < \widehat{\tau}_0\}$. It then holds

$$\begin{aligned}\{\widehat{\tau}_{0,M} > \widehat{\tau}_0\} &\subset \bigcup_{j=0}^{T-1} \{C_j < Z_j \leq \mathcal{C}_{j,M}\} \cap \{\widehat{\tau}_0 = j\} \\ &=: \bigcup_{j=0}^{T-1} A_j^{M+} \cap \{\widehat{\tau}_0 = j\},\end{aligned}$$

and similarly,

$$\begin{aligned}\{\widehat{\tau}_{0,M} < \widehat{\tau}_0\} &\subset \bigcup_{j=0}^{T-1} \{C_j \geq Z_j > \mathcal{C}_{j,M}\} \cap \{\widehat{\tau}_0 = j\} \\ &=: \bigcup_{j=0}^{T-1} A_j^{M-} \cap \{\widehat{\tau}_0 = j\}.\end{aligned}$$

So we have

$$\mathbb{P}\left(\widehat{\tau}_{0,M} \neq \widehat{\tau}_0\right) \leq \sum_{j=0}^{T-1} \mathbb{P}\left(A_j^{M+} \cup A_j^{M-}\right).$$

By the conditional version of the Bernstein inequality we have,

$$\begin{aligned}\mathbb{P}_{\mathcal{F}_T}\left(A_j^{M+}\right) &= \mathbb{P}_{X_j}\left(0 < Z_j - C_j \leq \frac{1}{M}\sum_{m=1}^{M}\left(Z_{\tau_{j+1}^{(m)}}(X_{\tau_{j+1}^{(m)}}^{j,X_j(m)}) - C_j\right)\right) \\ &\leq \mathbf{1}_{\{|Z_j - C_j| \leq M^{-1/2}\}} + \sum_{k=1}^{\infty}\mathbf{1}_{\{2^{k-1}M^{-1/2} < |Z_j - C_j| \leq 2^k M^{-1/2}\}} \cdot \\ &\cdot \mathbb{P}_{X_j}\left(2^{k-1}M^{-1/2} < \frac{1}{M}\sum_{m=1}^{M}\left(Z_{\tau_{j+1}^{(m)}}(X_{\tau_{j+1}^{(m)}}^{j,X_j,(m)}) - C_j\right)\right) \\ &\leq \mathbf{1}_{\{|Z_j - C_j| \leq M^{-1/2}\}} + \sum_{k=1}^{\infty}\mathbf{1}_{\{2^{k-1}M^{-1/2} < |Z_j - C_j| \leq 2^k M^{-1/2}\}} \cdot \\ &\cdot \exp\left[-\frac{2^{2k-3}M}{MB^2 + B2^{k-1}M^{1/2}/3}\right] \\ &\leq \mathbf{1}_{\{|Z_j - C_j| \leq M^{-1/2}\}} + \sum_{k=1}^{\infty}\mathbf{1}_{\{|Z_j - C_j| \leq 2^k M^{-1/2}\}} \cdot \\ &\cdot \exp\left[-\frac{2^{2k-3}}{B^2 + B2^{k-1}/3}\right].\end{aligned}$$

So by assumption (3.2),

$$\begin{aligned}\mathbb{P}\left(A_j^{M+}\right) &\leq DM^{-\alpha/2} + D\sum_{k=1}^{\infty}2^{\alpha k}M^{-\alpha/2}\exp\left[-\frac{2^{2k-3}}{B^2 + B2^{k-1}M^{-1/2}/3}\right] \\ &\leq B_1 M^{-\alpha/2}\end{aligned}$$

for B_1 depending on B, and α. After obtaining a similar estimate $\mathbb{P}\left(A_j^{M-}\right) \leq B_2 M^{-\alpha/2}$, we finally conclude that

$$\mathbb{P}\left(\widehat{\tau}_{0,M} \neq \widehat{\tau}_0\right) \leq M^{-\alpha/2}T\max(B_1, B_2) =: D_1 M^{-\alpha/2}.$$

7.2. Proof of Theorem 3.4. Define $\widehat{\tau}_M := \widehat{\tau}_{0,M}$, $\widehat{\tau} := \widehat{\tau}_0$, and use induction to the number of exercise dates T. For $T = 0$ the statement is trivially fulfilled. Suppose it is shown that

$$\mathbb{E}\left(Z_{\widehat{\tau}_M} - Z_{\widehat{\tau}}\right) = O\Big(\frac{1}{M}\Big)$$

for T exercise dates. Now consider the cash-flow process $Z_0, ..., Z_{T+1}$. Note that the filtration $(\mathcal{F}_j)$ is generated by the outer trajectories. Note, since $T+1$ is the last exercise date, the event $\{\widehat{\tau} = T+1\} = \Omega\backslash\{\widehat{\tau} \leq T\}$ is $\mathcal{F}_T$-measurable. Further, the event $\{\widehat{\tau}_M = T+1\} = \Omega\backslash\{\widehat{\tau}_M \leq T\}$ is measurable with respect to the information generated by the inner simulated trajectories starting from an outer trajectory at time T, and so, in particular, does **not** depend on the information generated by the the outer trajectories from T until $T+1$. That is, we have

$$\mathbb{E}_{\mathcal{F}_{T+1}}\left[\left(1_{\widehat{\tau}_M = T+1} - 1_{\widehat{\tau} = T+1}\right)\right] = \mathbb{E}_{\mathcal{F}_T}\left[\left(1_{\widehat{\tau}_M = T+1} - 1_{\widehat{\tau} = T+1}\right)\right]$$

and so

$$\mathbb{E}\left[Z_{T+1}\left(1_{\widehat{\tau}_M=T+1}-1_{\widehat{\tau}=T+1}\right)\right]=\mathbb{E}\left[Z_{T+1}\mathbb{E}_{\mathcal{F}_{T+1}}\left(1_{\widehat{\tau}_M=T+1}-1_{\widehat{\tau}=T+1}\right)\right]$$
$$\mathbb{E}\left[Z_{T+1}\mathbb{E}_{\mathcal{F}_T}\left[\left(1_{\widehat{\tau}_M=T+1}-1_{\widehat{\tau}=T+1}\right)\right]\right]. \tag{7.1}$$

By (7.1) and applying the induction hypothesis to the modified cash-flow $Z_j 1_{j\le T}$, it then follows that

$$\begin{aligned}\left|\mathbb{E}\left(Z_{\widehat{\tau}_M}-Z_{\widehat{\tau}}\right)\right| &= \left|\mathbb{E}\left(Z_{\widehat{\tau}_M}1_{\widehat{\tau}_M\le T}+Z_{T+1}1_{\widehat{\tau}_M=T+1}-Z_{\widehat{\tau}}1_{\widehat{\tau}\le T}-Z_{T+1}1_{\widehat{\tau}=T+1}\right)\right| \\ &= \left|\mathbb{E}\left(Z_{\widehat{\tau}_M}1_{\widehat{\tau}_M\le T}-Z_{\widehat{\tau}}1_{\widehat{\tau}\le T}\right)+\mathbb{E}\left(Z_{T+1}\left(1_{\widehat{\tau}_M=T+1}-1_{\widehat{\tau}=T+1}\right)\right)\right| \\ &\le O(\frac{1}{M})+\left|\mathbb{E}\left(Z_{T+1}\mathbb{E}_{\mathcal{F}_T}\left(1_{\widehat{\tau}_M=T+1}-1_{\widehat{\tau}=T+1}\right)\right)\right|. \end{aligned} \tag{7.2}$$

Let us estimate the second term $\mathbb{E}\left[Z_{T+1}\mathbb{E}_{\mathcal{F}_T}\left[1_{\widehat{\tau}_M=T+1}-1_{\widehat{\tau}=T+1}\right]\right]$. Denote $\varepsilon_{M,j}=1_{Z_j\le\mathcal{C}_{j,M}}-1_{Z_j\le C_j}$ for $j=0,...,T$, and $\overline{\varepsilon}_{M,j}=\mathbb{E}_{\mathcal{F}_j}\left[1_{Z_j\le\mathcal{C}_{j,M}}-1_{Z_j\le C_j}\right]$. Then by the identity ($i_0:=+\infty$)

$$\prod_{i=1}^n a_i-\prod_{i=1}^n b_i=\sum_{l=1}^n\sum_{i_l<i_{l-1}<\cdots<i_0}\prod_{r=1}^l(a_{i_r}-b_{i_r})\cdot\prod_{j\ne i_l,j\ne i_{l-1},\ldots,j\ne i_1}b_j$$

it holds

$$\mathbb{E}_{\mathcal{F}_T}\left[1_{\widehat{\tau}_M=T+1}-1_{\widehat{\tau}=T+1}\right]=\mathbb{E}_{\mathcal{F}_T}\left[\prod_{j=0}^T 1_{Z_j\le\mathcal{C}_{j,M}}-\prod_{j=0}^T 1_{Z_j\le C_j}\right]=\mathcal{R}_1+\mathcal{R}_2,$$

where

$$\mathcal{R}_1=\mathbb{E}_{\mathcal{F}_T}\left[\sum_{j=0}^T\varepsilon_{M,j}\prod_{i\ne j}1_{Z_i\le C_i}\right]=\sum_{j=0}^T\overline{\varepsilon}_{M,j}\prod_{i\ne j}1_{Z_i\le C_i}$$

and

$$\begin{aligned}\mathcal{R}_2 &= \mathbb{E}_{\mathcal{F}_T}\left[\sum_{j_2<j_1}\varepsilon_{M,j_1}\varepsilon_{M,j_2}\prod_{i\ne j_1,j_2}1_{Z_i\le C_i}\right] \\ &\quad+\mathbb{E}_{\mathcal{F}_T}\left[\sum_{j_3<j_2<j_1}\varepsilon_{M,j_1}\varepsilon_{M,j_2}\varepsilon_{M,j_3}\prod_{i\ne j_1,j_2,j_3}1_{Z_i\le C_i}\right]+\ldots \\ &= \sum_{j_2<j_1}\overline{\varepsilon}_{M,j_1}\overline{\varepsilon}_{M,j_2}\prod_{i\ne j_1,j_2}1_{Z_i\le C_i} \\ &\quad+\sum_{j_3<j_2<j_1}\overline{\varepsilon}_{M,j_1}\overline{\varepsilon}_{M,j_2}\overline{\varepsilon}_{M,j_3}\prod_{i\ne j_1,j_2,j_3}1_{Z_i\le C_i}+\ldots\end{aligned}$$

were we note that conditional $\mathcal{F}_T$ the $\varepsilon_{M,j}$ are independent. It is easy to show that

$$\overline{\varepsilon}_{M,j}=O_P\left(\frac{1}{\sqrt{M}}\right),\quad\text{hence}\quad\mathbb{E}[Z_{T+1}\mathcal{R}_2]=O\left(\frac{1}{M}\right).$$

Let us write

$$\begin{aligned}\mathbb{E}[Z_{T+1}\mathcal{R}_1] &= \sum_{j=0}^{T}\mathbb{E}\left[Z_{T+1}\overline{\varepsilon}_{M,j}\prod_{i\neq j}1_{Z_i\le C_i}\right]\\ &= \sum_{j=0}^{T}\mathbb{E}\left[\overline{\varepsilon}_{M,j}\mathbb{E}_{\mathcal{F}_j}\left[Z_{T+1}\prod_{i\neq j}1_{Z_i\le C_i}\right]\right]\\ &=: \sum_{j=0}^{T}\mathbb{E}\left[\overline{\varepsilon}_{M,j}W_j\right].\end{aligned}$$

By assumption, $Z_j = Z_j(X_j)$, $j = 0,\dots,T$. Let us set

$$f_j(x) := Z_j(x) - \mathbb{E}[Z_{\tau_{j+1}}(X_{\tau_{j+1}})|X_j = x] = Z_j(x) - C_j(x)$$

and consider for fixed j,

$$\mathcal{C}_{j,M} - C_j = \frac{1}{M}\sum_{m=1}^{M}\left(Z_{\tau_{j+1}^{(m)}}(X^{j,x,(m)}_{\tau_{j+1}^{(m)}}) - C_j(x)\right) =: \sigma_j(x)\frac{\Delta_{j,M}(x)}{\sqrt{M}}$$

where σ_j is defined in **(iii)**, and denote by $p_{j,M}(\cdot;x)$ the conditional density of the r.v. $\Delta_{j,M}(x)$ given $X_j = x$. Then

$$\begin{aligned}\mathbb{E}[Z_{T+1}\mathcal{R}_1] &= \sum_{j=0}^{T}\mathbb{E}\left[W_j\mathbb{E}_{\mathcal{F}_j}\left[1_{Z_j\le\mathcal{C}_{j,M}} - 1_{Z_j\le C_j}\right]\right]\\ &= \sum_{j=0}^{T}\mathbb{E}\left[W_j\mathbb{E}_{\mathcal{F}_j}\left[1_{\{f_j(X_j)\le\mathcal{C}_{j,M}-C_j\}} - 1_{f_j(X_j)\le 0}\right]\right]\\ &= \sum_{j=0}^{T}\mathbb{E}\left[W_j\mathbb{E}_{\mathcal{F}_j}\left[1_{\left\{f_j(X_j)\le\sigma_j(x)\frac{\Delta_{j,M}(X_j)}{\sqrt{M}}\right\}} - 1_{f_j(X_j)\le 0}\right]\right]\\ &= \sum_{j=0}^{T}\mathbb{E}\left[W_j\int p_{j,M}(z;X_j)\left(1_{\left\{f_j(X_j)\le\sigma_j(X_j)\frac{z}{\sqrt{M}}\right\}} - 1_{f_j(X_j)\le 0}\right)dz\right]\\ &= \sum_{j=0}^{T}\mathbb{E}\left[1_{f_j(X_j)>0}W_j\int p_{j,M}(z;X_j)\left(1_{\left\{f_j(X_j)\le\sigma_j(X_j)\frac{z}{\sqrt{M}}\right\}} - 1_{f_j(X_j)\le 0}\right)dz\right]\\ &+ \sum_{j=0}^{T}\mathbb{E}\left[1_{f_j(X_j)\le 0}W_j\int p_{j,M}(z;X_j)\left(1_{\left\{f_j(X_j)\le\sigma_j(X_j)\frac{z}{\sqrt{M}}\right\}} - 1_{f_j(X_j)\le 0}\right)dz\right]\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=0}^{T} \mathbb{E}\left[W_j \int p_{j,M}(z;X_j) 1_{\left\{0<f_j(X_j)\le \sigma_j(X_j)\frac{z}{\sqrt{M}}\right\}} dz\right] \\
&- \sum_{j=0}^{T} \mathbb{E}\left[W_j \int p_{j,M}(z;X_j) 1_{\left\{\sigma_j(X_j)\frac{z}{\sqrt{M}}<f_j(X_j)\le 0\right\}} dz\right] \\
&= \sum_{j=0}^{T} (I)_j - \sum_{j=0}^{T} (II)_j
\end{aligned}
$$

Note that

$$
\begin{aligned}
W_j &= \prod_{i<j} 1_{Z_i \le C_i} \mathbb{E}_{\mathcal{F}_j}\left[Z_{T+1} \prod_{i>j} 1_{Z_i\le C_i}\right] \\
&=: \prod_{i<j} 1_{Z_i\le C_i} V_j(X_j),
\end{aligned}
$$

so

$$
(I)_j = \mathbb{E}\left[\prod_{i<j} 1_{Z_i\le C_i} \mathbb{E}_{\mathcal{F}_{j-1}} V_j(X_j) \int p_{j,M}(z;X_j) 1_{\left\{0<f_j(X_j)\le\sigma_j(X_j)\frac{z}{\sqrt{M}}\right\}} dz\right],
$$

Consider

$$
\begin{aligned}
&\mathbb{E}_{\mathcal{F}_{j-1}}\left[V_j(X_j) \int p_{j,M}(z;X_j) 1_{\left\{0<f_j(X_j)\le\sigma_j(X_j)\frac{z}{\sqrt{M}}\right\}}\right] dz \\
&= \int \mathfrak{p}_j\left(x; X_{j-1}\right) V_j(x) \int p_{j,M}(z;x) 1_{\left\{0<f_j(x)\le\sigma_j(x)\frac{z}{\sqrt{M}}\right\}} dz\, dx \\
&= \int \int p_{j,M}(z;x) \mathfrak{p}_j\left(x; X_{j-1}\right) V_j(x) 1_{\left\{0<f_j(x)\le\sigma_j(x)\frac{z}{\sqrt{M}}\right\}} dx\, dz
\end{aligned}
$$

Similarly,

$$
(II)_j = \mathbb{E}\left[\prod_{i<j} 1_{Z_i\le C_i} \mathbb{E}_{\mathcal{F}_{j-1}}\left(V_j(X_j) \int p_{j,M}(z;X_j) 1_{\left\{\sigma_j(X_j)\frac{z}{\sqrt{M}}<f_j(X_j)\le 0\right\}}\right) dz\right]
$$

where

$$
\begin{aligned}
&\mathbb{E}_{\mathcal{F}_{j-1}}\left[V_j(X_j) \int p_{j,M}(z;X_j) 1_{\left\{\sigma_j(X_j)\frac{z}{\sqrt{M}}<f_j(X_j)\le 0\right\}} dz\right] \\
&= \int dz \int p_{j,M}(z;x) \mathfrak{p}_j\left(x; X_{j-1}\right) V_j(x) 1_{\left\{\sigma_j(x)\frac{z}{\sqrt{M}}<f_j(x)\le 0\right\}} dx,
\end{aligned}
$$

yielding

$$(I)_j - (II)_j = \mathbb{E}\left[\prod_{i<j} 1_{Z_i \le C_i} \int dz \int p_{j,M}(z;x)\mathfrak{p}_j\left(x;X_{j-1}\right)V_j(x)1_{\left\{0<f_j(x)\le\sigma_j(x)\frac{z}{\sqrt{M}}\right\}}dx\right]$$
$$- \mathbb{E}\left[\prod_{i<j} 1_{Z_i \le C_i} \int dz \int p_{j,M}(z;x)\mathfrak{p}_j\left(x;X_{j-1}\right)V_j(x)1_{\left\{\sigma_j(x)\frac{z}{\sqrt{M}}<f_j(x)\le 0\right\}}dx\right]$$

$$= \int dz \int p_{j,M}(z;x)V_j(x)\psi_j(x)1_{\left\{0<f_j(x)\le\sigma_j(x)\frac{z}{\sqrt{M}}\right\}}dx$$
$$- \int dz \int p_{j,M}(z;x)V_j(x)\psi_j(x)1_{\left\{\sigma_j(x)\frac{z}{\sqrt{M}}<f_j(x)\le 0\right\}}dx$$
$$=: (*)_1 - (*)_2,$$

where

$$\psi_j(x) := \mathbb{E}\left[\prod_{i<j} 1_{Z_i \le C_i}\mathfrak{p}_j\left(x;X_{j-1}\right)\right] \tag{7.3}$$

We may assume that

$$p_{j,M}(z;x) = \phi\left(z\right)\left(1 + \frac{D_{j,M}(z;x)}{\sqrt{M}}\right)$$

with ϕ being the standard normal density and with $D_{j,M}$ satisfying for all x and M the normalization condition

$$\int \phi\left(w\right)D_{j,M}(w;x)dw = 0,$$

and the growth bound

$$D_{j,M}(w;x) = O(e^{aw^2/2}) \text{ for some } a<1 \text{ uniformly in } j, M \text{ and } x. \tag{7.4}$$

For example, (7.4) is fulfilled if the cash-flow $Z_j(x)$ is uniformly bounded in j and x (see Appendix). Since by assumption $\sigma_j(x)$ is uniformly in x and j, bounded and bounded away from zero, we thus have

$$(*)_1 = \int dz \int \phi\left(z\right)V_j(x)\psi_j(x)1_{\left\{0<f_j(x)/\sigma_j(x)\le\frac{z}{\sqrt{M}}\right\}}dx$$
$$+ \int dz \int \phi\left(z\right)\frac{D_{j,M}(z;x)}{\sqrt{M}}V_j(x)\psi_j(x)1_{\left\{0<f_j(x)/\sigma_j(x)\le\frac{z}{\sqrt{M}}\right\}}dx$$
$$=: (*)_{1a} + (*)_{1b}$$

It is easy to see that

$$\int V_j(x)\psi_j(x)dx < \infty.$$

Indeed, V_j is bounded since by assumption **(ii)** Z_j is bounded, and by (7.3) it holds,

$$\psi_j \le \mathbb{E}\left[\mathfrak{p}_j\left(x; X_{j-1}\right)\right], \quad \text{hence} \quad \int \psi_j(x)dx \le 1.$$

Let now $\xi_j(dy)$ be the image, or the push-forward, of the absolutely continuous finite measure

$$V_j(x)\psi_j(x)dx$$

under the map

$$x \to \frac{f_j(x)}{\sigma_j(x)},$$

and consider

$$\begin{aligned}(*)_{1a} &= \int dz\phi\left(z\right) 1_{z>0}\xi_j((0,\frac{z}{\sqrt{M}}]) \\ &= \sqrt{M}\int 1_{t>0}dt\,\phi\left(t\sqrt{M}\right)\xi_j((0,t]).\end{aligned}$$

It follows from assumption **(i)** that $V_j, \psi_j > 0$ and that $V_j\psi_j$ has integrable derivatives. Then by an application of the implicit function theorem due to assumption **(v)** it can be shown that ξ has a density $g(0)$ in $t = 0$, such that

$$\xi_j((-t,0]) = tg(0) + O(t^2) \quad \text{and} \quad \xi_j((0,t]) = tg(0) + O(t^2), \quad t > 0. \tag{7.5}$$

By next following the standard Laplace method for integrals (e.g. see [10]) we get

$$\begin{aligned}(*)_{1a} &= \frac{\sqrt{M}}{2\pi}\int 1_{t>0}dte^{-\ Mt^2/2}\xi_j((0,t]) \\ &= d_jM^{-1/2} + O(M^{-1}).\end{aligned}$$

Further we have for some constant C

$$\begin{aligned}|(*)_{1b}| &\le \frac{C}{\sqrt{M}}\int dz\int\phi\left(z\right)e^{az^2/2}V_j(x)\psi_j(x)1_{\left\{0<\frac{f_j(x)}{\sigma_j(x)}\le\frac{w}{\sqrt{M}}\right\}}dx \\ &= \frac{C}{\sqrt{2\pi M}}\int e^{-\frac{1}{2}(1-a)z^2}\xi_j((0,\frac{z}{\sqrt{M}}])\,dz \\ &= \frac{C}{\sqrt{2\pi}}\int e^{-\frac{1}{2}(1-a)t^2M}\xi_j((0,t])\,dt = O(M^{-1}).\end{aligned}$$

Due to (7.5) we get in the same way $(*)_2 = (*)_{2a} + (*)_{2b}$,

$$\begin{aligned}(*)_{2a} &= \sqrt{M}\int 1_{t<0}dt\phi\left(t\sqrt{M}\right)\xi_j((t,0]) = \int 1_{t>0}dt\phi\left(t\sqrt{M}\right)\xi_j((-t,0]) \\ &= d_jM^{-1/2} + O(M^{-1})\end{aligned}$$

and $(*)_{2b} = O(M^{-1})$. Gathering all together we obtain $(*)_1 - (*)_2 = O(M^{-1})$, hence $(I)_j - (II)_j = O(M^{-1})$ for all j, and we so finally arrive at

$$\mathbb{E}\left[Z_{T+1}(1_{\widehat{\tau}_M=T+1} - 1_{\widehat{\tau}=T+1})\right] = O(M^{-1}).$$

7.3. Proof of Theorem 5.1. First, we analyze the variance of the estimator $\widehat{\mathcal{Y}}_{\mathbf{n},\mathbf{m}}$, that is given by

$$\begin{aligned}\operatorname{Var}\left[\widehat{\mathcal{Y}}_{\mathbf{n},\mathbf{m}}\right] &= \frac{1}{n_0}\operatorname{Var}\left[Z_{\widehat{\tau}_{m_0}}\right] + \sum_{l=1}^{L}\frac{1}{n_l}\operatorname{Var}\left[Z_{\widehat{\tau}_{m_l}} - Z_{\widehat{\tau}_{m_{l-1}}}\right] \\ &\leq \frac{\sigma_\infty^2}{n_0} + \frac{\mathcal{V}_\infty \kappa^{-\beta}}{n_1 m_0^\beta}\frac{\kappa^{L(1-\beta)/2}-1}{\kappa^{(1-\beta)/2}-1}, \end{aligned} \tag{7.6}$$

cf. (4.2) and (5.4). Let us now minimise the complexity (5.5) over the parameters n_0 and n_1, for given L, m_0 and accuracy ϵ, that is (cf. (4.2)),

$$\left(\frac{\mu_\infty}{m_0^\gamma \kappa^{\gamma L}}\right)^2 + \frac{\sigma_\infty^2}{n_0} + \frac{\mathcal{V}_\infty \kappa^{-\beta}}{n_1 m_0^\beta}\frac{\kappa^{L(1-\beta)/2}-1}{\kappa^{(1-\beta)/2}-1} = \epsilon^2.$$

We thus have to choose L such that $\frac{\mu_\infty}{m_0^\gamma \kappa^{\gamma L}} < \epsilon$, i.e.,

$$L > \gamma^{-1}\frac{\ln \epsilon^{-1} + \ln\left(\mu_\infty / m_0^\gamma\right)}{\ln \kappa}. \tag{7.7}$$

With a Lagrangian optimization we find

$$n_0^*\left(L, m_0, \epsilon\right) = \frac{\sigma_\infty^2 + \sigma_\infty \mathcal{V}_\infty^{1/2} m_0^{-\beta/2}\frac{\kappa^{L(1-\beta)/2}-1}{1-\kappa^{-(1-\beta)/2}}}{\epsilon^2 - \left(\frac{\mu_\infty}{m_0^\gamma \kappa^{L\gamma}}\right)^2}, \tag{7.8}$$

$$n_1^*\left(L, m_0, \epsilon\right) = n_0^*\left(L, m_0, \epsilon\right)\sigma_\infty^{-1}\kappa^{-(1+\beta)/2}\mathcal{V}_\infty^{1/2} m_0^{-\beta/2}. \tag{7.9}$$

This results in a complexity (see (5.5))

$$\begin{aligned}\mathcal{C}_{ML}\left(n_0^*, n_1^*, L, m_0, \epsilon\right) &:= n_0^*\left(L, m_0, \epsilon\right) m_0 + n_1^*\left(L, m_0, \epsilon\right) m_0 \kappa \frac{\kappa^{L(1-\beta)/2}-1}{\kappa^{(1-\beta)/2}-1} \\ &= \frac{\left(\sigma_\infty m_0^{\beta/2} + \sqrt{\mathcal{V}_\infty}\frac{\kappa^{L(1-\beta)/2}-1}{\kappa^{(1-\beta)/2}-1}\kappa^{(1-\beta)/2}\right)^2 m_0^{1-\beta}}{\epsilon^2 - \left(\frac{\mu_\infty}{m_0^\gamma \kappa^{L\gamma}}\right)^2}. \end{aligned} \tag{7.10}$$

Next we are going to optimize over L. To this end we differentiate (7.10) to L and set the derivative equal to zero, which yields,

$$\begin{aligned}\epsilon^2\kappa^{2L\gamma} &= \frac{\mu_\infty^2}{m_0^{2\gamma}}\left(1 + 2\gamma/\left(1-\beta\right)\right) \\ &\quad + \frac{2\gamma}{1-\beta}\frac{\mu_\infty^2}{m_0^{2\gamma}}\left(1 + \sigma_\infty m_0^{\beta/2}\mathcal{V}_\infty^{-1/2}\left(1 - \kappa^{-(1-\beta)/2}\right)\right)\kappa^{-L(1-\beta)/2} \\ &=: p + q\kappa^{-L(1-\beta)/2}, \quad \text{with} \end{aligned} \tag{7.11}$$

$$L = \frac{\ln \epsilon^{-1}}{\gamma \ln \kappa} + \frac{\ln p}{2\gamma \ln \kappa} + \frac{\ln\left(1 + q\kappa^{-L(1-\beta)/2}/p\right)}{2\gamma \ln \kappa}. \tag{7.12}$$

From (7.11) we see that there is at most one solution in L, and since $\beta < 1$ we see from (7.12) that $L \to \infty$ as $\epsilon \downarrow 0$. So we may write

$$L = \frac{\ln \epsilon^{-1}}{\gamma \ln \kappa} + \frac{\ln p}{2\gamma \ln \kappa} + O\left(\kappa^{-L(1-\beta)/2}\right), \quad \epsilon \downarrow 0. \tag{7.13}$$

Due to (7.13) we have that

$$L = \frac{\ln \epsilon^{-1}}{\gamma \ln \kappa} + O(1), \quad \epsilon \downarrow 0, \tag{7.14}$$

hence by iterating (7.13) with (7.14) once, we obtain the asymptotic solution

$$L^* := \frac{\ln \epsilon^{-1}}{\gamma \ln \kappa} + \frac{\ln p}{2\gamma \ln \kappa} + O\left(\epsilon^{(1-\beta)/(2\gamma)}\right), \quad \epsilon \downarrow 0, \tag{7.15}$$

that obviously satisfies (7.7) for ϵ small enough. We now are ready to prove the following asymptotic complexity theorem. Due to (7.15) it holds for $a > 0$,

$$\kappa^{aL^*} = p^{a/(2\gamma)} \epsilon^{-a/\gamma} \left(1 + O\left(\epsilon^{(1-\beta)/(2\gamma)}\right)\right), \quad \text{hence}$$

$$\kappa^{L^*(1-\beta)/2} = p^{(1-\beta)/(4\gamma)} \epsilon^{-(1-\beta)/(2\gamma)} + O(1) \quad \text{and} \tag{7.16}$$

$$\kappa^{\gamma L^*} = p^{1/2} \epsilon^{-1} \left(1 + O\left(\epsilon^{(1-\beta)/(2\gamma)}\right)\right). \tag{7.17}$$

So by inserting (7.16), (7.17) with (7.11) in (7.10) we get after elementary algebraic and asymptotic manipulations (5.6). By inserting (7.16), (7.17) with (7.11) in (7.8) and (7.9) respectively we get in the same way (5.7) and (5.8), respectively. Finally, combining (7.11) and (7.15) yields (5.9).

8. Appendix.

Convergent Edgeworth type expansions. Let p_M be the density of the square-root scaled sum:

$$\frac{\Delta_1 + \ldots + \Delta_M}{\sqrt{M}},$$

where $\Delta_1, \ldots, \Delta_M$ are i.i.d. with $\mathtt{E}[\Delta_m] = 0$ and $\mathtt{Var}[\Delta_m] = 1$, $m = 1, ..., M$. The density p_M has a formal representation:

$$p_M(z) = \phi(z) \left[\sum_{j=0}^{\infty} \frac{h_j(z)\Gamma_{j,M}}{j!}\right]$$

with

$$h_j(z) = (-1)^j \left[\frac{d^j}{dz^j} \exp\left(-z^2/2\right)\right] \exp\left(z^2/2\right).$$

The coefficients $\Gamma_{j,M}$ are found from

$$\exp\left(\sum_{j=1}^{\infty}(\kappa_{j,M}-\alpha_j)\beta^j/j!\right)=\sum_{j=1}^{\infty}\Gamma_{j,M}\beta^j/j!,$$

where $\kappa_{j,M}$ are the cumulants of the distribution due to p_M and α_j are the cumulants of the standard normal distribution. It is clear that

$$\Gamma_{0,M}=1$$

and that

$$\Gamma_{n,M}=\sum_{k=1}^{n}\frac{k}{n}\Gamma_{n-k,M}(\kappa_{k,M}-\alpha_k)$$

for $n>0$. Note that $\alpha_1=\kappa_{1,M}$ and $\alpha_k=0$ for $k>1$. Hence $\Gamma_{1,M}=\Gamma_{2,M}=0$ and

$$\Gamma_{n,M}=\sum_{k=3}^{n}\frac{k}{n}\Gamma_{n-k,M}\kappa_{k,M}$$

for $n>2$.

Lemma 8.1. *Let the random variable Δ_1 be bounded, i.e., $|\Delta_1|<A$ a.s., then*

$$|\Gamma_{n,M}|\leq\frac{C^n}{\sqrt{M}}$$

for some constant C depending on A.

Proof. First note that $\Gamma_{3,M}=\kappa_{3,M}$ and

$$|\kappa_{k,M}|\leq C^kM^{1-k/2},\quad k\in\mathbb{N},$$

for some constant C depending on A. Assume that the statement is proved for all $n\leq n_0$. Then

$$\begin{aligned}|\Gamma_{n_0+1,M}|&\leq\frac{C^{n_0+1}}{\sqrt{M}}\sum_{k=3}^{n_0+1}\frac{k}{n_0+1}M^{1-k/2}\\&\leq\frac{C^{n_0+1}}{\sqrt{M}}\sum_{k=3}^{n_0+1}\frac{k}{n_0+1}M^{-k/6}\leq\frac{C^{n_0+1}}{\sqrt{M}}\end{aligned}$$

for M large enough. ■

Since

$$|h_j(z)|\leq B^j|z|^j$$

for some $B>0$, it holds

$$p_M(z)=\phi(z)\left[1+\frac{D_M(z)}{\sqrt{M}}\right],$$

where

$$|D_M(z)| \leq \left| \sum_{j=3}^{\infty} \frac{h_j(z)\sqrt{M}\Gamma_{j,M}}{j!} \right| \leq \left| \sum_{j=3}^{\infty} \frac{|z|^j (BC)^j}{j!} \right|$$
$$\leq \exp(BC|z|),$$

which implies (7.4).

REFERENCES

[1] Leif Andersen and Mark Broadie, *Primal-dual simulation algorithm for pricing multidimensional american options*, Management Science, 50 (2004), pp. 1222–1234.

[2] Vlad Bally and Gilles Pagès, *A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems*, Bernoulli, 9 (2003), pp. 1003–1049.

[3] Denis Belomestny, Fabian Dickmann, and Tigran Nagapetyan, *Pricing american options via multi-level approximation methods*, arXiv preprint arXiv:1303.1334, (2013).

[4] Denis Belomestny, John Schoenmakers, and Fabian Dickmann, *Multilevel dual approach for pricing american style derivatives*, Finance and Stochastics, 17 (2013), pp. 717–742.

[5] Christian Bender, Anastasia Kolodko, and John Schoenmakers, *Enhanced policy iteration for American options via scenario selection*, Quant. Finance, 8 (2008), pp. 135–146.

[6] Christian Bender and John Schoenmakers, *An iterative method for multiple stopping: convergence and stability*, Adv. in Appl. Probab., 38 (2006), pp. 729–749.

[7] Mark Broadie and Paul Glasserman, *A stochastic mesh method for pricing high-dimensional american options*, Journal of Computational Finance, 7 (2004), pp. 35–72.

[8] Jacques F Carriere, *Valuation of the early-exercise price for options using simulations and nonparametric regression*, Insurance: mathematics and Economics, 19 (1996), pp. 19–30.

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

[10] N. G. de Bruijn, *Asymptotic methods in analysis*, Dover Publications, Inc., New York, third ed., 1981.

[11] Michael B Giles, *Multilevel monte carlo path simulation*, Operations Research, 56 (2008), pp. 607–617.

[12] Paul Glasserman, *Monte Carlo methods in financial engineering*, vol. 53, Springer Science & Business Media, 2003.

[13] Michael B Gordy and Sandeep Juneja, *Nested simulation in portfolio risk measurement*, Management Science, 56 (2010), pp. 1833–1848.

[14] Martin B. Haugh and Leonid Kogan, *Pricing American options: a duality approach*, Oper. Res., 52 (2004), pp. 258–270.

[15] Ronald A. Howard, *Dynamic programming and Markov processes*, The Technology Press of M.I.T., Cambridge, Mass.; John Wiley & Sons, Inc., New York-London, 1960.

[16] Herb Johnson, *Options on the maximum or the minimum of several assets*, Journal of Financial and Quantitative Analysis, 22 (1987), pp. 277–283.

[17] Anastasia Kolodko and John Schoenmakers, *Iterative construction of the optimal Bermudan stopping time*, Finance Stoch., 10 (2006), pp. 27–49.

[18] Francis A Longstaff and Eduardo S Schwartz, *Valuing american options by simulation: A simple least-squares approach*, Review of Financial studies, 14 (2001), pp. 113–147.

[19] Martin L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons, Inc., New York, 1994. A Wiley-Interscience Publication.

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

[21] John N Tsitsiklis and Benjamin Van Roy, *Regression methods for pricing complex american-style options*, Neural Networks, IEEE Transactions on, 12 (2001), pp. 694–703.