Improved Lower and Upper Bound Algorithms for Pricing American Options by Simulation *

Mark Broadie[†] and Menghui Cao[‡]

December 2007

Abstract

This paper introduces new variance reduction techniques and computational improvements to Monte Carlo methods for pricing American-style options. For simulation algorithms that compute lower bounds of American option values, we apply martingale control variates and introduce the *local policy enhancement*, which adopts a local simulation to improve the exercise policy. For duality-based upper bound methods, specifically the primal-dual simulation algorithm (Andersen and Broadie 2004), we have developed two improvements. One is *sub-optimality checking*, which saves unnecessary computation when it is sub-optimal to exercise the option along the sample path; the second is *boundary distance grouping*, which reduces computational time by skipping computation on selected sample paths based on the distance to the exercise boundary. Numerical results are given for single asset Bermudan options, moving window Asian options and Bermudan max options. In some examples the computational time is reduced by a factor of several hundred, while the confidence interval of the true option value is considerably tighter than before the improvements.

Key words: American option, Bermudan option, moving window Asian option, Bermudan max option, Monte Carlo simulation, primal-dual simulation algorithm, variance reduction, option pricing.

1 Introduction

1.1 Background

Pricing American-style options is challenging, especially under multi-dimensional and path-dependent settings, for which lattice and finite difference methods are often impractical due to the curse of dimensionality. In recent years many simulation-based algorithms have been

*An earlier draft of this paper was titled *Improvements in Pricing American Options Using Monte Carlo Simulation*. We thank Paul Glasserman, Mark Joshi and seminar participants at Columbia University and CIRANO conference for their valuable comments. We thank an anonymous referee for suggestions that improved the paper. This work was supported by NSF Grant DMS-0410234.

†Graduate School of Business, Columbia University, New York, NY 10027. Email: mnb2@columbia.edu.

‡DFG Investment Advisers, 135 E 57th Street, New York, NY 10022. Email: david.cao@dfgia.com.

proposed for pricing American options, most using a hybrid approach of simulation and dynamic programming to determine an exercise policy. Because these algorithms produce an exercise policy which is inferior to the optimal policy, they provide low-biased estimators of the true option values. We use the term *lower bound algorithm* to refer to any method that produces a low-biased estimate of an American option value with a sub-optimal exercise strategy.[1]

Regression-based methods for pricing American options are proposed by Carriere (1996), Tsitsiklis and Van Roy (1999) and Longstaff and Schwartz (2001). The least-squares method by Longstaff and Schwartz projects the conditional discounted payoffs onto basis functions of the state variables. The projected value is then used as the approximate continuation value, which is compared with the intrinsic value for determining the exercise strategy. Low-biased estimates of the option values can be obtained by generating a new, i.e., independent, set of simulation paths, and exercising according to the sub-optimal exercise strategy. Clément, Lamberton and Protter (2002) analyze the convergence of the least-squares method. Glasserman and Yu (2004) study the tradeoff between the number of basis functions and the number of paths. Broadie, Glasserman and Ha (2000) propose a weighted Monte Carlo method in which the continuation value of the American option is expressed as a weighted sum of future values and the weights are selected to optimize a convex objective function subject to known conditional expectations. Glasserman and Yu (2002) analyze this 'regression later' approach, compared to the 'regression now' approach implied in other regression-based methods.

One difficulty associated with lower bound algorithms is that of determining how well they estimate the true option value. If a high-biased estimator is obtained in addition to the low-biased estimator, a confidence interval can be constructed for the true option value, and the width of the confidence interval may be used as an accuracy measure for the algorithms. Broadie and Glasserman (1997, 2004) propose two convergent methods that generate both lower and upper bounds of the true option values, one based on simulated trees and the other a stochastic mesh method. Haugh and Kogan (2004) and Rogers (2002) independently develop dual formu-

[1]The terms exercise strategy, stopping time and exercise policy will be used interchangeably in this paper.

lations of the American option pricing problem, which can be used to construct upper bounds of the option values. Andersen and Broadie (2004) show how duality-based upper bounds can be computed directly from any given exercise policy through a simulation algorithm, leading to significant improvements in their practical implementation. We call any algorithm that produces a high-biased estimate of an American option value an *upper bound algorithm*.

The duality-based upper bound estimator can often be represented as a lower bound estimator plus a penalty term. The penalty term, which may be viewed as the value of a non-standard lookback option, is a non-negative quantity that penalizes potentially incorrect exercise decisions made by the sub-optimal policy. Estimation of this penalty term requires nested simulations which is computationally demanding. Our paper addresses this major shortcoming of the duality-based upper bound algorithm by introducing improvements that may significantly reduce its computational time and variance. We also propose enhancements to lower bound algorithms which improve exercise policies and reduce the variance.

1.2 Brief results

The improvements developed and tested in this paper include martingale control variates and local policy enhancement for lower bound algorithms, and sub-optimality checking and boundary distance grouping enhancements for upper bound algorithms. The least-squares Monte Carlo method introduced by Longstaff and Schwartz (2001) is used as the lower bound algorithm and the primal-dual simulation algorithm by Andersen and Broadie (2004) is used as the upper bound method, although the improvements can be applied to other lower bound and duality-based upper bound algorithms.

Many lower bound algorithms approximate the option's continuation value and compare it with the option's intrinsic value to form a sub-optimal exercise policy. If the approximation of the continuation value is inaccurate, it often leads to a poor exercise policy. To improve the exercise policy, we propose a *local policy enhancement* which employs sub-simulation to gain a better estimate of the continuation value in circumstances where the sub-optimal policy is likely to generate incorrect decisions. Then the sub-simulation estimate is compared with the intrinsic value to potentially override the original policy's decision to exercise or continue.

In many upper bound algorithms, a time-consuming sub-simulation is carried out to estimate the option's continuation value at every exercise time. We show in Section 4 that sub-simulation is not needed when the option is sub-optimal to exercise, that is, when the intrinsic value is lower than the continuation value. Based on this idea, *sub-optimality checking* is a simple technique to save computational work and improve the upper bound estimator. It states that we can skip the sub-simulations when the option's intrinsic value is lower than an easily derived lower bound of the continuation value along the sample path. Despite being simple, this approach often leads to dramatic computational improvements in the upper bound algorithms, especially for out-of-the-money (OTM) options.

Boundary distance grouping is another method to enhance the efficiency of duality-based upper bound algorithms. For many simulation paths, the penalty term that contributes to the upper bound estimator is zero. Thus it would be more efficient if we could identify in advance the paths with non-zero penalties. The goal of boundary distance grouping is to separate the sample paths into two groups, one group deemed more likely to produce zero penalties, the 'zero' group, and its complement, the 'non-zero' group. A sampling method is used to derive the upper bound estimator with much less computational effort, through the saving of sub-simulation, on the sample paths in the 'zero' group. The fewer paths there are in the 'non-zero' group, the greater will be the computational saving achieved by this method. While the saving is most significant for deep OTM options, the technique is useful for in-the-money (ITM) and at-the-money (ATM) options as well.

Bermudan options are American-style options that can be exercised at discrete time prior to the maturity. Most computer-based algorithms effectively price Bermudan options, instead of continuously-exercisable American options, due to the finite nature of the computer algorithms. This paper provides numerical results on single asset Bermudan options, moving window Asian options and Bermudan basket options, the latter two of which are difficult to price using lattice or finite difference methods. The techniques introduced in this paper are general enough to be used for other types of Bermudan options, such as Bermudan interest rate swaptions.

The rest of this paper is organized as follows. In Section 2, the Bermudan option pricing problem is formulated. Section 3 addresses the martingale control variates and local policy enhancement for the lower bound algorithms. Section 4 introduces sub-optimality checking and boundary distance grouping for the upper bound algorithms. Numerical results are shown in Section 5. In Section 6, we conclude and suggest directions for further research. Some numerical details, including a comparison between 'regression now' and 'regression later,' the choice of basis functions, the proofs of propositions, and the variance estimation for boundary distance grouping, are given in the appendices.

2 Problem formulation

We consider a complete financial market where the assets are driven by Markov processes in a standard filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let B_t denote the value at time t of \$1 invested in a risk-free money market account at time 0, $B_t = e^{\int_0^t r_s ds}$, where r_s denotes the instantaneous risk-free interest rate at time s. Let S_t be an $\mathbb{R}^d$-valued Markov process with the initial state S_0, which denotes the process of underlying asset prices or state variables of the model. There exists an equivalent probability measure $\mathbb{Q}$, also known as the risk-neutral measure, under which discounted asset prices are martingales. Pricing of any contingent claim on the assets can be obtained by taking the expectation of discounted cash flows with respect to the $\mathbb{Q}$ measure. Let $E_t[\cdot]$ denote the conditional expectation under the $\mathbb{Q}$ measure given the information up to time t, i.e., $E_t[\cdot] = E^{\mathbb{Q}}[\cdot|\mathcal{F}_t]$. We consider here discretely-exercisable American options, also known as Bermudan options, which may be exercised only at a finite number of time steps $\Gamma = \{t_0, t_1, ..., t_n\}$ where $0 = t_0 < t_1 < t_2 < ... < t_n \leq T$. τ is the stopping time which can take values in Γ. The intrinsic value h_t is the option payoff upon exercise at time t, for example $h_t = (S_t - K)^+$ for a single asset call option with strike price K, where $x^+ := \max(x, 0)$.

The pricing of Bermudan options can be formulated as a primal-dual problem. The primal problem is to maximize the expected discounted option payoff over all possible stopping times,

$$\text{Primal: } V_0 = \sup_{\tau \in \Gamma} E_0 \left[\frac{h_\tau}{B_\tau} \right]. \tag{1}$$

More generally, the discounted Bermudan option value at time $t_i < T$ is

$$\frac{V_{t_i}}{B_{t_i}} = \sup_{\tau \geq t_i} E_{t_i} \left[\frac{h_\tau}{B_\tau} \right] = \max \left(\frac{h_{t_i}}{B_{t_i}}, E_{t_i} \left[\frac{V_{t_{i+1}}}{B_{t_{i+1}}} \right] \right), \tag{2}$$

where V_t/B_t is the discounted value process and the smallest super-martingale that dominates h_t/B_t on $t \in \Gamma$ (see Lamberton and Lapeyre 1996). The stopping time which achieves the largest option value is denoted τ^*.[2]

Haugh and Kogan (2004) and Rogers (2002) independently propose the dual formulation of the problem. For an arbitrary adapted super-martingale process π_t we have

$$\begin{aligned} V_0 &= \sup_{\tau \in \Gamma} E_0 \left[\frac{h_\tau}{B_\tau} \right] = \sup_{\tau \in \Gamma} E_0 \left[\frac{h_\tau}{B_\tau} + \pi_\tau - \pi_\tau \right] \\ &\leq \pi_0 + \sup_{\tau \in \Gamma} E_0 \left[\frac{h_\tau}{B_\tau} - \pi_\tau \right] \\ &\leq \pi_0 + E_0 \left[\max_{t \in \Gamma} \left(\frac{h_t}{B_t} - \pi_t \right) \right], \end{aligned} \tag{3}$$

which gives an upper bound of V_0. Based on this, the dual problem is to minimize the upper bound with respect to all adapted super-martingale processes,

$$\text{Dual: } U_0 = \inf_{\pi \in \Pi} \left\{ \pi_0 + E_0 \left[\max_{t \in \Gamma} \left(\frac{h_t}{B_t} - \pi_t \right) \right] \right\}, \tag{4}$$

where Π is the set of all adapted super-martingale processes. Haugh and Kogan (2004) show that the optimal values of the primal and the dual problems are equal, i.e., $V_0 = U_0$, and the optimal solution of the dual problem is achieved with π_t^* being the discounted optimal value process.

3 Improvements to lower bound algorithms

3.1 A brief review of the lower bound algorithm

Most algorithms for pricing American options are lower bound algorithms, which produce low-biased estimates of American option values. They usually involve generating an exercise strategy

[2] In general we use '*' to indicate a variable or process associated with the optimal stopping time.

and then valuing the option by following the exercise strategy. Let L_t be the lower bound value process associated with τ, which is defined as

$$\frac{L_t}{B_t} = E_t\left[\frac{h_{\tau_t}}{B_{\tau_t}}\right], \tag{5}$$

where $\tau_t = \inf\{u \in \Gamma \cap [t,T] : \mathbf{1}_u = 1\}$ and $\mathbf{1}_t$ is the adapted exercise indicator process, which equals 1 if the sub-optimal strategy indicates exercise and 0 otherwise. Clearly, the sub-optimal exercise strategy is always dominated by the optimal strategy,

$$L_0 = E_0\left[\frac{h_\tau}{B_\tau}\right] \leq V_0,$$

in other words, L_0 is a lower bound of the Bermudan option value V_0.

We denote Q_t, or Q_t^τ, as the option's continuation value at time t under the sub-optimal strategy τ,

$$Q_{t_i} = E_{t_i}\left[\frac{B_{t_i}}{B_{t_{i+1}}} L_{t_{i+1}}\right], \tag{6}$$

and $\tilde{Q}_t$ as the approximation of the continuation value. In regression-based algorithms, $\tilde{Q}_t$ is the projected continuation value through a linear combination of the basis functions

$$\tilde{Q}_{t_i} = \sum_{k=1}^{b} \hat{\beta}_k f_k(S_{1,t_i}, ..., S_{d,t_i}), \tag{7}$$

where $\hat{\beta}_k$ is the regression coefficient and $f_k(\cdot)$ is the corresponding basis function.

Low bias of sub-optimal policy is introduced when the decision from the sub-optimal policy differs from the optimal decision. Broadie and Glasserman (2004) propose *policy fixing* to prevent some of these incorrect decisions: the option is considered for exercise only if the exercise payoff exceeds a lower limit of the continuation value, $\underline{Q_t}$. A straightforward choice for this exercise lower limit is the value of the corresponding European option if it can be valued analytically. More generally it can be the value of any option dominated by the Bermudan option or the maximum among the values of all dominated options (e.g., the maximum among the values of European options that mature at each exercise time of the Bermudan option). We apply policy fixing for all lower bound computations in this paper. Note that we only use the values of single European options and not the maximum among multiple option values, because the latter invalidates the condition for Proposition 1 (refer to the proof of Proposition 1 in Appendix B for more detail). Denote the adjusted approximate continuation value as $\underline{\tilde{Q}_t} := \max(\tilde{Q}_t, \underline{Q_t})$. The sub-optimal strategy with policy fixing can be defined as

$$\tau_t = \inf\{u \in \Gamma \cap [t, T] : h_u > \underline{\tilde{Q}_u}\}. \tag{8}$$

If it is optimal to exercise the option and yet the sub-optimal exercise strategy indicates otherwise, i.e.,

$$Q_t^* < h_t \leq \underline{\tilde{Q}_t}, \quad \mathbf{1}_t^* = 1 \text{ and } \mathbf{1}_t = 0, \tag{9}$$

it is an *incorrect continuation.* Likewise when it is optimal to continue but the sub-optimal exercise strategy indicates exercise, i.e.,

$$Q_t^* \geq h_t > \underline{\tilde{Q}_t}, \quad \mathbf{1}_t^* = 0 \text{ and } \mathbf{1}_t = 1, \tag{10}$$

it is an *incorrect exercise.*

3.2 Distance to the exercise boundary

In this section we discuss an approach to quantify the distance of an option to the exercise boundary. The *exercise boundary* is the surface in the state space where the option holder, based on the exercise policy, is indifferent between holding and exercising the option. Accordingly the sub-optimal exercise boundary can be defined as the set of states at which the adjusted approximate continuation value equals the exercise payoff, i.e., $\{\omega_t : \underline{\tilde{Q}_t} = h_t\}$. The exercise region is where $\underline{\tilde{Q}_t} < h_t$ and the sub-optimal policy indicates exercise, and vice versa for the continuation region.

Incorrect decisions are more likely to occur when the option is 'close' to the exercise boundary. To determine how 'close' the option is from the sub-optimal exercise boundary we introduce a boundary distance measure

$$d_t := |\underline{\tilde{Q}_t} - h_t|. \tag{11}$$

This function is measured in units of the payoff as opposed to the underlying state variables. It does not strictly satisfy the axioms of a distance function, but it does have similar characteristics. In particular, d_t is zero only when the option is on the sub-optimal exercise boundary and it increases as $\underline{\tilde{Q}_t}$ deviates from h_t. We can use it as a measure of closeness between the sample path and the sub-optimal exercise boundary. Alternative boundary distance measures include $|\underline{\tilde{Q}_t} - h_t|/h_t$ and $|\underline{\tilde{Q}_t} - h_t|/S_t$.

3.3 Local policy enhancement

The idea of local policy enhancement is to employ a sub-simulation to estimate the continuation value $\hat{Q}_t$ and use that, instead of the approximate continuation value $\tilde{Q}_t$, to make the exercise decision. Since the sub-simulation estimate is generally more accurate than the approximate continuation value, this may improve the exercise policy, at the expense of additional computational effort.

It is computationally demanding, however, to perform a sub-simulation at every time step. To achieve a good tradeoff between accuracy and computational cost, we would like to launch a sub-simulation only when an incorrect decision is considered more likely to be made. Specifically, we launch a sub-simulation at time t if the sample path is sufficiently close to the exercise boundary.

The simulation procedure for the lower bound algorithm with local policy enhancement is as follows:

(i) Simulate the path of state variables until either the sub-optimal policy indicates exercise or the option matures.

(ii) At each exercise time, compute h_t, $\underline{Q_t}$, $\tilde{Q}_t$, and d_t. Continue if $h_t \le \underline{Q_t}$, otherwise

  a. If $d_t > \epsilon$, follow the original sub-optimal strategy, exercise if $h_t > \tilde{Q}_t$, continue otherwise.

  b. If $d_t \le \epsilon$, launch a sub-simulation with N_ϵ paths to estimate $\hat{Q}_t$, exercise if $h_t > \hat{Q}_t$, continue otherwise.

(iii) Repeat steps (i)-(ii) for N_L sample paths, obtain the lower bound estimator $\hat{L}_0$ by averaging the discounted payoffs.

Due to the computational cost for sub-simulations, local policy enhancement may prove to be too expensive to apply for some Bermudan options.

3.4 Use of control variates

The fast and accurate estimation of an option's continuation values is essential to the pricing of American options in both the lower and upper bound computations. We use the control variate technique to improve the efficiency of continuation value estimates. The control variate method is a broadly used technique for variance reduction (see, for example, Boyle, Broadie and Glasserman 1997), which adjusts the simulation estimates by quantities with known expectations. Assume we know the expectation of X and want to estimate $E[Y]$. The control variate adjusted estimator is $\bar{Y} - \beta(\bar{X} - E[X])$, where β is the adjustment coefficient. The variance-minimizing adjustment coefficient is $\beta^* = \rho_{XY}\frac{\sigma_Y}{\sigma_X}$, which can be estimated from the X and Y samples. Broadie and Glasserman (2004) use European option values as controls for pricing Bermudan options and apply them in two levels: inner controls are used for estimating continuation values and outer controls are used for the mesh estimates. Control variates contribute to tighter price bounds in two ways, by reducing both the standard errors of the lower bound estimators and the bias of the upper bound estimators.

Typically control variates are valued at a fixed time, such as the European option's maturity. Rasmussen (2005) and Broadie and Glasserman (2004) use control variates that are valued at the exercise time of the Bermudan option rather than at maturity, which leads to larger variance reduction because the control is sampled at an exercise time and so has a higher correlation with the Bermudan option value. This approach requires the control variate to have the martingale property and thus can be called a *martingale control variate*. We apply this technique in our examples, specifically by taking single asset European option values at the exercise time as controls for single asset Bermudan options, and the geometric Asian option values at the exercise time as controls for moving window Asian options. For Bermudan max options, since there is no simple analytic formula for European max options on more than two assets, we use the average of single asset European option values as the martingale control.

As discussed in Glasserman (2003), bias may be introduced if the same samples are used to estimate the adjustment coefficient β and the control variate adjusted value. In order to avoid bias which may sometimes be significant, we can fix the adjustment coefficient at a constant value. In our examples we fix the coefficient at one when estimating the single asset Bermudan option's continuation value with European option value as the control, and find it to be generally effective.

4 Improvements to upper bound algorithms

The improvements shown in this section can be applied to duality-based upper bound algorithms. In particular we use the primal-dual simulation algorithm of Andersen and Broadie (2004).

4.1 Duality-based upper bound algorithms

The dual problem of pricing Bermudan options is

$$U_0 \quad = \quad \inf_{\pi \in \Pi} \left\{ \pi_0 + E_0 \left[\max_{t \in \Gamma} \left(\frac{h_t}{B_t} - \pi_t \right) \right] \right\}.$$

Since the discounted value process V_t/B_t is a super-martingale, we can use Doob-Meyer decomposition to decompose it into a martingale process π_t^* and an adapted increasing process A_t^*

$$\frac{V_t}{B_t} = \pi_t^* - A_t^*. \tag{12}$$

This gives

$$\frac{h_t}{B_t} - \pi_t^* = \frac{h_t}{B_t} - \frac{V_t}{B_t} - A_t^* \leq 0, \quad \forall t \in \Gamma,$$

since $\frac{h_t}{B_t} \leq \frac{V_t}{B_t}$ and $A_t^* \geq 0$. Hence,

$$\max_{t \in \Gamma} \left(\frac{h_t}{B_t} - \pi_t^* \right) \leq 0$$

and using the definition of U_0 above we get $U_0 \leq \pi_0^* = V_0$. But also $V_0 \leq U_0$, so $U_0 = V_0$, i.e., there is no duality gap. For martingales other than π^* there will be a gap between the resulting upper and lower bounds, so the question is how to construct a martingale process that leads to a tight upper bound when the optimal policy is not available.

4.2 Primal-dual simulation algorithm

The primal-dual simulation algorithm is a duality-based upper bound algorithm that builds upon simulation and can be used together with any lower-bound algorithm to generate an upper bound of Bermudan option values. We can decompose L_t/B_t as

$$\frac{L_t}{B_t} = \pi_t - A_t, \tag{13}$$

where π_t is an adapted martingale process defined as,

$$\begin{aligned} \pi_0 &:= L_0, \quad \pi_{t_1} := L_{t_1}/B_{t_1}, \\ \pi_{t_{i+1}} &:= \pi_{t_i} + \frac{L_{t_{i+1}}}{B_{t_{i+1}}} - \frac{L_{t_i}}{B_{t_i}} - \mathbf{1}_{t_i} E_{t_i}\left[\frac{L_{t_{i+1}}}{B_{t_{i+1}}} - \frac{L_{t_i}}{B_{t_i}}\right] \quad \text{for } 1 \leq i \leq n-1. \end{aligned} \tag{14}$$

Since $\frac{Q_{t_i}}{B_{t_i}} = \frac{L_{t_i}}{B_{t_i}}$ when $\mathbf{1}_{t_i} = 0$ and $\frac{Q_{t_i}}{B_{t_i}} = E_{t_i}\left[\frac{L_{t_{i+1}}}{B_{t_{i+1}}}\right]$ when $\mathbf{1}_{t_i} = 1$, we have

$$\pi_{t_{i+1}} = \pi_{t_i} + \frac{L_{t_{i+1}}}{B_{t_{i+1}}} - \frac{Q_{t_i}}{B_{t_i}}. \tag{15}$$

Define the upper bound increment D as

$$D := \max_{t \in \Gamma}\left(\frac{h_t}{B_t} - \pi_t\right), \tag{16}$$

which can be viewed as the payoff from a non-standard lookback call option, with the discounted Bermudan option payoff being the state variable and the adapted martingale process being the floating strike. The duality gap $D_0 := E_0[D]$ can be estimated by $\bar{D} := \frac{1}{N_H}\sum_{i=1}^{N_H} D_i$, and $\hat{H}_0 = \hat{L}_0 + \hat{D}_0$ will be the upper bound estimator from the primal-dual simulation algorithm. The sample variance of the upper bound estimator can be approximated as the sum of sample variances from the lower bound estimator and the duality gap estimator,

$$\frac{\hat{s}_H^2}{N_H} \approx \frac{\hat{s}_L^2}{N_L} + \frac{\hat{s}_D^2}{N_H}, \tag{17}$$

because the two estimators are uncorrelated when estimated independently. The simulation procedure for the primal-dual simulation algorithm is as follows:

(i) Simulate the path of state variables until the option matures.

(ii) At each exercise time, launch a sub-simulation with N_S paths to estimate Q_t/B_t and update π_t using equation (15).

(iii) Calculate the upper bound increment D for the current path.

(iv) Repeat steps (i)–(iii) for N_H sample paths, estimate the duality gap $\hat{D}_0$ and combine it with $\hat{L}_0$ to obtain the upper bound estimator $\hat{H}_0$.

Implementation details are given in Anderson and Broadie (2004). Note that A_t is not necessarily an increasing process since L_t/B_t is not a super-martingale. In fact,

$$A_{t_{i+1}} - A_{t_i} = -\mathbf{1}_{t_i} E_{t_i}\left[\frac{L_{t_{i+1}}}{B_{t_{i+1}}} - \frac{L_{t_i}}{B_{t_i}}\right] = \begin{cases} 0, & \mathbf{1}_{t_i} = 0, \\ \frac{h_{t_i}}{B_{t_i}} - \frac{Q_{t_i}}{B_{t_i}}, & \mathbf{1}_{t_i} = 1, \end{cases}$$

which decreases when an incorrect exercise decision is made.

The ensuing Propositions 1 and 2 illustrate some properties of the primal-dual simulation algorithm. Proofs are provided in Appendix B.

Proposition 1

(i) If $h_{t_i} \le \underline{Q_{t_i}}$ *for* $1 \le i \le k$*, then* $\pi_{t_k} = \frac{L_{t_k}}{B_{t_k}}$ *and* $\frac{h_{t_k}}{B_{t_k}} - \pi_{t_k} \le 0.$

(ii) If $h_{t_i} \le \underline{Q_{t_i}}$ *for* $l \le i \le k$*, then* $\pi_{t_k} = \pi_{t_{l-1}} - \frac{Q_{t_{l-1}}}{B_{t_{l-1}}} + \frac{L_{t_k}}{B_{t_k}}$ *and* $\frac{h_{t_k}}{B_{t_k}} - \pi_{t_k} \le \frac{Q_{t_{l-1}}}{B_{t_{l-1}}} - \pi_{t_{l-1}}.$

Proposition 1(i) states that the martingale process π_t is equal to the discounted lower bound value process and there is no contribution to the upper bound increment before the option enters the exercise region, and 1(ii) means that the computation of π_t does not depend on the path during the period that option stays in the continuation region. It follows from 1(ii) that the sub-simulation is not needed when the option is sub-optimal to exercise. In independent work, Joshi

(2007) derives results very akin to those shown in Proposition 1 by using a hedging portfolio argument. He shows that the upper bound increment is zero in the continuation region, simply by changing the payoff function to negative infinity when the option is sub-optimal to exercise.

If an option stays in the continuation region throughout its life, the upper bound increment D for the path is zero. The result holds even if the option stays in the continuation region until the final step. Furthermore, if it is sub-optimal to exercise the option except in the last two exercise dates and the optimal exercise policy is available at the last step before maturity (for example, if the corresponding European option can be valued analytically), the upper bound increment D is also zero, because $\pi_{t_{n-1}} = \frac{L_{t_{n-1}}}{B_{t_{n-1}}} = \frac{V_{t_{n-1}}}{B_{t_{n-1}}} \geq \frac{h_{t_{n-1}}}{B_{t_{n-1}}}$.

Proposition 2 *For a given sample path,*

(i) If $\exists \delta > 0$ such that $|\tilde{Q}_t - Q_t| < \delta$, and $d_t \geq \delta$ or $h_t \leq \underline{Q_t}$ holds $\forall t \in \Gamma$, then A_t is an increasing process and $D = 0$ for the path.

(ii) If $\exists \delta > 0$ such that $|\tilde{Q}_t - Q_t^*| < \delta$, and $d_t \geq \delta$ or $h_t \leq \underline{Q_t}$ holds $\forall t \in \Gamma$, then $\mathbf{1}_t \equiv \mathbf{1}_t^*$.

The implication of Proposition 2 is that, given a uniformly good approximation of the sub-optimal continuation value ($|\tilde{Q}_t - Q_t|$ is bounded above by a constant δ), the upper bound increment will be zero for a sample path if it never gets close to the sub-optimal exercise boundary. And if the approximation is uniformly good relative to the optimal continuation value ($|\tilde{Q}_t - Q_t^*|$ is bounded above by a constant δ), the sub-optimal exercise strategy will always coincide with the optimal strategy for the path never close to the sub-optimal boundary.

4.3 Sub-optimality checking

The primal-dual simulation algorithm launches a sub-simulation to estimate continuation values at every exercise time along the sample path. The continuation values are then used to determine the martingale process and eventually an upper bound increment. These sub-simulations are computationally demanding, however, many of them are not necessary.

Sub-optimality checking is an effective way to address this issue. It is based on the idea of Proposition 1, and can be easily implemented by comparing the option exercise payoff with the exercise lower limit $\underline{Q_t}$. The sub-simulations will be skipped when the exercise payoff is lower than the exercise lower limit, in other words, when it is sub-optimal to exercise the option.

Despite being simple, sub-optimality checking may bring dramatic computational improvement, especially for deep OTM options. Efficiency of the simulation may be measured by the product of sample variance and simulation time, and we can define an *effective saving factor* (ESF) as the ratio of the efficiency before and after improvement. Since the sub-optimality checking reduces computational time without affecting variance, its ESF is simply the ratio of computational time before and after the improvement.

The simulation procedure for the primal-dual algorithm with sub-optimality checking is as follows:

(i) Simulate the path of underlying variables until the option matures.

(ii) At each exercise time, if $h_t > \underline{Q_t}$, launch a sub-simulation with N_S paths to estimate Q_t/B_t and update π_t using Proposition 1; otherwise skip the sub-simulation and proceed to next time step.

(iii) Calculate the upper bound increment D for the current path.

(iv) Repeat (i)-(iii) for N_H sample paths, estimate the duality gap D_0 and combine it with $\hat{L}_0$ to obtain the upper bound estimator $\hat{H}_0$.

4.4 Boundary distance grouping

By Proposition 2, when the sub-optimal strategy is close to optimal, many of the simulation paths will have zero upper bound increments D. The algorithm, however, may spend a substantial amount of time to compute these zero values. We can eliminate much of this work by characterizing the paths that are more likely to produce non-zero upper bound increments than others. We do so by identifying paths that, for at least once during their life, are 'close' to the sub-optimal exercise boundary.

In boundary distance grouping, we separate the sample paths into two groups according to the distance of each path to the sub-optimal exercise boundary. Paths that are ever within a certain distance to the boundary during the option's life are placed into the 'non-zero' group, because it is suspected that the upper bound increment is non-zero. All other paths, the ones that never get close to the sub-optimal exercise boundary, are placed into the 'zero' group. A sampling method is used to eliminate part of the simulation work for the 'zero' group when estimating upper bound increments. If the fraction of paths in the 'non-zero' group is small, the computational saving from doing this can be substantial. The two groups are defined as follows:

$$Z := \{\omega : \forall t \in \Gamma, d_t(\omega) \geq \delta \text{ or } h_t \leq \underline{Q_t}\}, \tag{18}$$

$$\bar{Z} := \{\omega : \exists t \in \Gamma, d_t(\omega) < \delta \text{ and } h_t > \underline{Q_t}\}. \tag{19}$$

If there exists a small constant $\delta_0 > 0$ such that $P(\{\omega : \max_{t\in\Gamma} |\tilde{Q}_t(\omega) - Q_t(\omega)| < \delta_0\}) = 1$, the distance threshold δ could simply be chosen as δ_0 so that by Proposition 2, D is zero for all the sample paths that belong to Z. In general δ_0 is not known, and the appropriate choice of δ still remains, as we will address below.

Assume the D estimator has mean μ_D and variance σ_D^2. Without loss of generality, we assume $n_{\bar{Z}}$ out of the N_H paths belong to group $\bar{Z}$ and are numbered from 1 to $n_{\bar{Z}}$, i.e., $\omega_1, ..., \omega_{n_{\bar{Z}}} \in \bar{Z}$, and $\omega_{n_{\bar{Z}}+1}, ..., \omega_{N_H} \in Z$. Let $p_{\bar{Z}}$ be the probability that a sample path belongs to group $\bar{Z}$,

$$p_{\bar{Z}} = P(\omega \in \bar{Z}) = P(\{\omega : \exists t \in \Gamma, d_t(\omega) < \delta \text{ and } h_t > \underline{Q_t}\}). \tag{20}$$

The conditional means and variances for upper bound increments in the two groups are $\mu_{\bar{Z}}$, $\sigma_{\bar{Z}}^2$, μ_Z and σ_Z^2. In addition to the standard estimator which is the simple average, an alternative estimator of the duality gap can be constructed by estimating D_is from a selected set of paths, more specifically the $n_{\bar{Z}}$ paths in group $\bar{Z}$ and l_Z paths randomly chosen from group Z ($l_Z \ll N_H - n_{\bar{Z}}$). For simplicity, we pick the first l_Z paths from group Z. The new estimator is

$$\tilde{D} := \frac{1}{N_H}\left(\sum_{i=1}^{n_{\bar{Z}}} D_i + \frac{N_H - n_{\bar{Z}}}{l_Z}\sum_{i=n_{\bar{Z}}+1}^{n_{\bar{Z}}+l_Z} D_i\right), \tag{21}$$

which may be easily shown to be unbiased. Although the variance of $\tilde{D}$ is higher than the variance of $\bar{D}$, the difference is usually small (see Appendix C).

As shown in Appendix C, under certain conditions the effective saving factor of boundary distance grouping is simply the saving of computational time by only estimating D_is from group $\bar{Z}$ paths instead of all paths, i.e.,

$$ESF = \frac{\text{Var}[\bar{D}] \cdot T_{\bar{D}}}{\text{Var}[\tilde{D}] \cdot T_{\tilde{D}}} \approx 1 + \frac{T_{D_Z}}{p_{\bar{Z}} T_{D_{\bar{Z}}}}, \tag{22}$$

which goes to infinity as $p_{\bar{Z}} \to 0$, $T_{\bar{D}}$ and $T_{\tilde{D}}$ are the expected time to obtain the standard estimator and the alternative estimator, $T_{D_{\bar{Z}}}$ and T_{D_Z} are respectively the expected time to estimate upper bound increment D from a group $\bar{Z}$ path and from a group Z path.

Notice that after the grouping, we cannot directly estimate $\text{Var}[\tilde{D}]$ by calculating the sample variance from D_is because they are no longer identically distributed. Appendix C gives two indirect methods for estimating the sample variance. The simulation procedure for primal-dual algorithm with boundary distance grouping is as follows:

(i) Generate n_p pilot paths as in the standard primal-dual algorithm. For each δ among a set of values, estimate the parameters $p_{\bar{Z}}$, $\mu_{\bar{Z}}$, $\sigma_{\bar{Z}}$, T_P, T_I, etc., and calculate l'_Z, then choose the δ' that optimizes the efficient measure.

(ii) Simulate the path of underlying variables until the option matures.

(iii) Estimate the boundary distance d_t along the path, if $\exists t \in \Gamma$ such that $d_t < \delta'$ and $h_t > \underline{Q_t}$, assign the path to group $\bar{Z}$, otherwise assign it to Z.

(iv) If the current path belongs to group $\bar{Z}$ or is among the first l'_Z paths in group Z, estimate the upper bound increment D as in the regular primal-dual algorithm, otherwise skip it.

(v) Repeat steps (ii)-(iv) for N_H sample paths, estimate the duality gap using the alternative estimator $\tilde{D}$ and combine it with $\hat{L}_0$ to obtain the upper bound estimator $\hat{H}_0$.

5 Numerical results

Numerical results for single asset Bermudan options, moving window Asian options and Bermudan max options are presented in this section. The underlying assets are assumed to follow the standard single and multi-asset Black-Scholes model. In the results below, $\hat{L}_0$ is the lower bound estimator obtained through the least-squares method (Longstaff and Schwartz 2001), t_L is the computational time associated with it, $\hat{H}_0$ is the upper bound estimator obtained through the primal-dual simulation algorithm (Andersen and Broadie 2004), t_H is the associated computational time, and $t_T = t_L + t_H$ is the total computational time. The point estimator is obtained by taking the average of lower bound and upper bound estimators. All computations are done on a Pentium 4 2.0GHz computer and computation time is measured in minutes.

In the four summary tables below (Tables 1-4), we show the improvements from methods introduced in this paper, through measures including the low and high estimators, the standard errors and the computational time. Each table is split into three panels: the top panel contains results before improvement, the middle panel shows the reduction of upper bound computational time through sub-optimality checking and boundary distance grouping, and the bottom panel shows the additional variance reduction and estimator improvement through local policy enhancement and the martingale control variate. Note that the local policy enhancement is only used for moving window Asian options (Table 2), for which we find the method effective without significantly increasing the computational cost.

For all regression-based algorithms, which basis functions to use is often critical but not obvious. We summarize the choice of basis functions for our numerical examples, as well as the comparison between 'regression later' and 'regression now,' in Appendix A.

5.1 Single asset Bermudan options

The single asset Bermudan option is the most standard and simplest Bermudan-type option. We assume the asset price follows the geometric Brownian motion process

$$\frac{dS_t}{S_t} = (r - q)dt + \sigma dW_t, \tag{23}$$

where W_t is standard Brownian motion. The payoff upon exercise for a single asset Bermudan call option at time t is $(S_t - K)^+$. The option and model parameters are defined as follows: σ is the annualized volatility, r is the continuously compounded risk-free interest rate, q is the continuously compounded dividend rate, K is the strike price, T is the maturity in years, and there are $n+1$ exercise opportunities, equally spaced at time $t_i = iT/n, i = 0, 1, ..., n$.

In the implementation of lower bound algorithms, paths are usually simulated from the initial state for which the option value is desired, to determine the sub-optimal exercise policy. However, the optimal exercise policy is independent of this initial state. To approximate the optimal policy more efficiently, we disperse the initial state for regression, an idea independently proposed in Rasmussen (2005). The paths of state variables are generated from a distribution of initial states, more specifically by simulating the state variables from strike K at time $-T/2$ instead of from S_0 at time 0. This dispersion method can be particularly helpful when pricing deep OTM and deep ITM options, given that simulating paths from the initial states of these options is likely to contribute little to finding the optimal exercise strategy, since most of the paths will be distant from the exercise boundary. The regression only needs to be performed once for pricing options with same strike and different initial states, in which case the total computational time is significantly reduced. In terms of regression basis functions, using powers of European option values proves to be more efficient than using powers of the underlying asset prices.

Table 1 shows the improvements in pricing single asset Bermudan call options using techniques introduced in this paper. It demonstrates that the simulation algorithm may work remarkably well, even compared to the binomial method. In each of the seven cases, a tight confidence interval containing the true value can be produced in a time comparable to, or less than, the binomial method. The widths of 95% confidence intervals are all within 0.4% of the true option values.

Table 1: Summary results for single asset Bermudan call options

S_0	$\hat{L}_0$(s.e.)	t_L	$\hat{H}_0$(s.e.)	t_H	95% C.I.	t_T	Point est.	True
70	0.1261(.0036)	0.03	0.1267(.0036)	4.26	0.1190, 0.1338	4.29	0.1264	0.1252
80	0.7075(.0090)	0.04	0.7130(.0091)	4.73	0.6898, 0.7309	4.77	0.7103	0.6934
90	2.3916(.0170)	0.05	2.4148(.0172)	6.12	2.3584, 2.4484	6.16	2.4032	2.3828
100	5.9078(.0253)	0.07	5.9728(.0257)	7.95	5.8583, 6.0231	8.02	5.9403	5.9152
110	11.7143(.0296)	0.08	11.8529(.0303)	8.33	11.6562, 11.9123	8.41	11.7836	11.7478
120	20.0000(.0000)	0.03	20.1899(.0076)	5.83	20.0000, 20.2049	5.87	20.0950	20.0063
130	30.0000(.0000)	0.00	30.0523(.0043)	3.55	30.0000, 30.0608	3.55	30.0261	30.0000
70	0.1281(.0036)	0.03	0.1288(.0037)	0.00	0.1210, 0.1360	0.03	0.1285	0.1252
80	0.7075(.0090)	0.04	0.7113(.0091)	0.01	0.6898, 0.7291	0.05	0.7094	0.6934
90	2.3916(.0170)	0.05	2.4185(.0172)	0.12	2.3584, 2.4523	0.16	2.4051	2.3828
100	5.9078(.0253)	0.07	5.9839(.0258)	0.61	5.8583, 6.0344	0.68	5.9459	5.9152
110	11.7143(.0296)	0.08	11.8624(.0304)	1.98	11.6562, 11.9219	2.06	11.7883	11.7478
120	20.0000(.0000)	0.03	20.2012(.0075)	2.09	20.0000, 20.2159	2.12	20.1006	20.0063
130	30.0000(.0000)	0.00	30.0494(.0040)	1.75	30.0000, 30.0572	1.75	30.0247	30.0000
70	0.1251(.0001)	0.03	0.1251(.0001)	0.00	0.1249, 0.1254	0.04	0.1251	0.1252
80	0.6931(.0003)	0.04	0.6932(.0003)	0.01	0.6925, 0.6939	0.05	0.6932	0.6934
90	2.3836(.0007)	0.05	2.3838(.0007)	0.12	2.3821, 2.3852	0.16	2.3837	2.3828
100	5.9167(.0013)	0.07	5.9172(.0013)	0.61	5.9141, 5.9198	0.68	5.9170	5.9152
110	11.7477(.0019)	0.08	11.7488(.0019)	1.98	11.7441, 11.7524	2.07	11.7482	11.7478
120	20.0032(.0015)	0.04	20.0105(.0016)	2.19	20.0003, 20.0136	2.22	20.0069	20.0063
130	30.0000(.0000)	0.00	30.0007(.0004)	0.71	30.0000, 30.0015	0.71	30.0004	30.0000

Note: Option parameters are $\sigma = 20\%$, $r = 5\%$, $q = 10\%$, $K = 100$, $T = 1$, $n = 50$, $b = 3$, $N_R = 100,000$, $N_L = 100,000$, $N_H = 1000$, $N_S = 500$. The three panels respectively contain results before improvement (top), after the improvement of sub-optimality checking and boundary distance grouping (middle), and additionally with martingale control variate (bottom)–European call option value sampled at the exercise time in this case. The true value is obtained through a binomial lattice with 36,000 time steps, which takes approximately two minutes per option.

5.2 Moving window Asian options

A moving window Asian option is a Bermudan-type option that can be exercised at any time t_i before T ($i \geq m$), with the payoff dependent on the average of the asset prices during a period of fixed length. Consider the asset price S_t following the geometric Brownian motion process defined in equation (23), and let A_{t_i} be the arithmetic average of S_t over the m periods up to time t_i, i.e.,

$$A_{t_i} = \frac{1}{m} \sum_{k=i-m+1}^{i} S_{t_k}. \tag{24}$$

The moving window Asian option can be exercised at any time t_i with payoff $(A_{t_i} - K)^+$ for a call and $(K - A_{t_i})^+$ for a put. Notice that it becomes a standard Asian option when $m = n$, and a single asset Bermudan option when $m = 1$. The European version of this option is a forward starting Asian option or Asian tail option.

The early exercise feature, along with the payoff's dependence on the historic average, makes the moving window Asian option difficult to value by lattice or finite difference methods. Monte Carlo simulation appears to be a good alternative to price these options. Polynomials of underlying asset price and arithmetic average are used as the regression basis functions.

As shown in Table 2, the moving window Asian call options can be priced with high precision using Monte Carlo methods along with the improvements in this paper. For all seven cases, the 95% confidence interval widths lie within 1% of the true option values, compared to 2-7 times that amount before improvements. The lower bound computing time is longer after the improvements due to the sub-simulations in local policy enhancement, but the total computational time is reduced in every case.

5.3 Symmetric Bermudan max options

A Bermudan max option is a discretely-exercisable option on multiple underlying assets whose payoff depends on the maximum among all asset prices. We assume the asset prices follow

Table 2: Summary results for moving window Asian call options

S_0	$\hat{L}_0$(s.e.)	t_L	$\hat{H}_0$(s.e.)	t_H	95% C.I.	t_T	Point est.
70	0.345(.007)	0.04	0.345(.007)	3.08	0.331, 0.359	3.12	0.345
80	1.715(.017)	0.05	1.721(.017)	3.67	1.682, 1.754	3.72	1.718
90	5.203(.030)	0.05	5.226(.030)	5.13	5.144, 5.285	5.18	5.214
100	11.378(.043)	0.08	11.427(.044)	6.97	11.293, 11.512	7.05	11.403
110	19.918(.053)	0.10	19.992(.053)	8.24	19.814, 20.097	8.34	19.955
120	29.899(.059)	0.10	29.992(.060)	8.58	29.782, 30.109	8.68	29.945
130	40.389(.064)	0.10	40.490(.064)	8.61	40.264, 40.616	8.71	40.440
70	0.345(.007)	0.04	0.345(.007)	0.00	0.331, 0.358	0.04	0.345
80	1.715(.017)	0.05	1.721(.017)	0.01	1.682, 1.754	0.06	1.718
90	5.203(.030)	0.05	5.227(.030)	0.10	5.144, 5.286	0.15	5.215
100	11.378(.043)	0.08	11.419(.044)	0.25	11.294, 11.504	0.33	11.399
110	19.918(.053)	0.10	19.990(.054)	0.55	19.814, 20.095	0.65	19.954
120	29.899(.059)	0.10	29.995(.060)	1.26	29.782, 30.112	1.36	29.947
130	40.389(.064)	0.11	40.478(.064)	1.67	40.264, 40.604	1.78	40.433
70	0.338(.001)	0.08	0.338(.001)	0.00	0.336, 0.340	0.08	0.338
80	1.699(.003)	0.30	1.702(.003)	0.01	1.694, 1.708	0.31	1.701
90	5.199(.005)	0.91	5.206(.006)	0.11	5.189, 5.217	1.02	5.203
100	11.406(.007)	2.01	11.417(.008)	0.25	11.391, 11.433	2.26	11.411
110	19.967(.009)	3.36	19.987(.010)	0.55	19.949, 20.007	3.92	19.977
120	29.961(.010)	4.24	29.972(.011)	1.26	29.942, 30.993	5.50	29.967
130	40.443(.010)	4.22	40.453(.011)	1.68	40.423, 40.475	5.89	40.448

Note: Option parameters are $\sigma = 20\%$, $r = 5\%$, $q = 0\%$, $K = 100$, $T = 1$, $n = 50$, $m = 10$, $b = 6$, $N_R = 100,000$, $N_L = 100,000$, $N_H = 1000$, $N_S = 500$. The three panels respectively contain results before improvement (top), after the improvement of sub-optimality checking and boundary distance grouping (middle), and additionally with local policy enhancement and martingale control variate (bottom)–geometric Asian option value sampled at the exercise time in this case. For the local policy enhancement $\epsilon = 0.5$ and $N_\epsilon = 100$.

correlated geometric Brownian motion processes, i.e.,

$$\frac{dS_{j,t}}{S_{j,t}} = (r - q_j)dt + \sigma_j dW_{j,t}, \tag{25}$$

where $W_{j,t}, j = 1, ..., d$, are standard Brownian motions and the instantaneous correlation between $W_{j,t}$ and $W_{k,t}$ is ρ_{jk}. The payoff of a 5-asset Bermudan max call option is $(\max_{1\leq j\leq 5} S_{j,t} - K)^+$.

For simplicity, we assume $q_j = q$, $\sigma_j = \sigma$ and $\rho_{jk} = \rho$, for all $j, k = 1, ..., d$ and $j \neq k$. We call this the symmetric case because the common parameter values mean the future asset returns do not depend on the index of specific asset. Under these assumptions the assets are numerically indistinguishable, which facilitates simplification in the choice of regression basis functions. In particular, the polynomials of sorted asset prices can be used as the (non-distinguishing) basis functions, without referencing to a specific asset index.

Table 3 provides pricing results for 5-asset Bermudan max call options before and after the improvements in this paper. Considerably tighter price bounds and reduced computational time are obtained, in magnitudes similar to that observed for the single asset Bermudan option and moving window Asian option.

Next we consider the more general case, in which the assets have asymmetric parameters and are thus distinguishable.

5.4 Asymmetric Bermudan max options

We use the 5-asset max call option with asymmetric volatilities (ranging from 8% to 40%) as an example. Table 4 shows that the magnitude of the improvements from the techniques in this paper are comparable to their symmetric counterpart. The lower bound estimator in the asymmetric case may be significantly improved by including basis functions that distinguish the assets (see Table 5). Nonetheless, for a reasonably symmetric or a large basket of assets, it is often more efficient to use the non-distinguishing basis functions, because of the impracticality to include the large number of asset-specific basis functions.

Table 6 illustrates that the local policy enhancement can effectively improve the lower bound

Table 3: Summary results for 5-asset symmetric Bermudan max call options

S_0	$\hat{L}_0$(s.e.)	t_L	$\hat{H}_0$(s.e.)	t_H	95% C.I.	t_T	Point est.
70	3.892(.006)	0.72	3.904(.006)	2.74	3.880, 3.916	3.46	3.898
80	9.002(.009)	0.81	9.015(.009)	3.01	8.984, 9.033	3.82	9.009
90	16.622(.012)	0.97	16.655(.012)	3.39	16.599, 16.679	4.36	16.638
100	26.120(.014)	1.12	26.176(.015)	3.72	26.093, 26.205	4.83	26.148
110	36.711(.016)	1.18	36.805(.017)	3.90	36.681, 36.838	5.08	36.758
120	47.849(.017)	1.18	47.985(.019)	3.92	47.816, 48.023	5.10	47.917
130	59.235(.018)	1.16	59.403(.021)	3.86	59.199, 59.445	5.02	59.319
70	3.892(.006)	0.72	3.901(.006)	0.05	3.880, 3.913	0.77	3.897
80	9.002(.009)	0.80	9.015(.009)	0.10	8.984, 9.033	0.90	9.008
90	16.622(.012)	0.97	16.662(.012)	0.45	16.599, 16.686	1.42	16.642
100	26.120(.014)	1.12	26.165(.015)	0.91	26.093, 26.194	2.03	26.142
110	36.711(.016)	1.19	36.786(.017)	1.33	36.681, 36.819	2.52	36.749
120	47.849(.017)	1.19	47.994(.020)	1.62	47.816, 48.033	2.81	47.921
130	59.235(.018)	1.16	59.395(.021)	2.14	59.199, 59.437	3.30	59.315
70	3.898(.001)	0.70	3.903(.001)	0.06	3.896, 3.906	0.76	3.901
80	9.008(.002)	0.79	9.014(.002)	0.10	9.004, 9.019	0.90	9.011
90	16.627(.004)	0.95	16.644(.004)	0.46	16.620, 16.653	1.41	16.636
100	26.125(.005)	1.09	26.152(.006)	0.91	26.115, 26.164	2.00	26.139
110	36.722(.006)	1.15	36.781(.009)	1.34	36.710, 36.798	2.49	36.752
120	47.862(.008)	1.16	47.988(.012)	1.62	47.847, 48.011	2.78	47.925
130	59.250(.009)	1.45	59.396(.013)	2.14	59.233, 59.423	3.59	59.323

Note: Option parameters are $\sigma = 20\%$, $q = 10\%$, $r = 5\%$, $K = 100$, $T = 3$, $\rho = 0$, $n = 9$, $b = 18$, $N_R = 200,000$, $N_L = 2,000,000$, $N_H = 1500$ and $N_S = 1000$. The three panels respectively contain results before improvement (top), after the improvement of sub-optimality checking and boundary distance grouping (middle), and additionally with martingale control variate (bottom)–average of European option values sampled at the exercise time in this case.

Table 4: Summary results for 5-asset asymmetric Bermudan max call options

S_0	$\hat{L}_0$(s.e.)	t_L	$\hat{H}_0$(s.e.)	t_H	95% C.I.	t_T	Point est.
70	11.756(.016)	0.74	11.850(.019)	2.75	11.723, 11.888	3.49	11.803
80	18.721(.020)	0.96	18.875(.024)	3.43	18.680, 18.921	4.39	18.798
90	27.455(.024)	1.25	27.664(.028)	4.26	27.407, 27.719	5.52	27.559
100	37.730(.028)	1.57	38.042(.033)	5.13	37.676, 38.107	6.70	37.886
110	49.162(.031)	1.75	49.555(.037)	5.73	49.101, 49.627	7.48	49.358
120	61.277(.034)	1.82	61.768(.040)	5.99	61.211, 61.848	7.81	61.523
130	73.709(.037)	1.83	74.263(.044)	6.07	73.638, 74.349	7.89	73.986
70	11.756(.016)	0.74	11.850(.019)	0.19	11.723, 11.883	0.93	11.801
80	18.721(.020)	0.96	18.875(.024)	0.38	18.680, 18.933	1.34	18.803
90	27.455(.024)	1.25	27.664(.028)	0.62	27.407, 27.741	1.87	27.570
100	37.730(.028)	1.57	38.042(.033)	1.47	37.676, 38.106	3.04	37.886
110	49.162(.031)	1.75	49.555(.037)	2.58	49.101, 49.626	4.33	49.357
120	61.277(.034)	1.82	61.768(.040)	3.17	61.211, 61.830	4.99	61.514
130	73.709(.037)	1.83	74.263(.044)	4.11	73.638, 74.351	5.94	73.986
70	11.778(.003)	0.75	11.842(.007)	0.19	11.772, 11.856	0.95	11.810
80	18.744(.004)	0.98	18.866(.011)	0.39	18.736, 18.887	1.38	18.805
90	27.480(.006)	1.29	27.659(.014)	0.62	27.468, 27.686	1.90	27.570
100	37.746(.008)	1.62	37.988(.016)	1.48	37.730, 38.020	3.10	37.867
110	49.175(.010)	1.79	49.492(.020)	2.58	49.155, 49.531	4.37	49.334
120	61.294(.015)	1.86	61.686(.023)	3.17	61.269, 61.730	5.04	61.490
130	73.723(.015)	1.88	74.184(.026)	4.12	73.694, 74.234	6.00	73.953

Note: Option parameters are the same as in Table 3, except that $\sigma_i = 8\%, 16\%, 24\%, 32\%$ and 40% respectively for $i = 1, 2, 3, 4, 5$. The three panels respectively contain results before improvement (top), after the improvement of sub-optimality checking and boundary distance grouping (middle), and additionally with martingale control variate (bottom)–average of European option values sampled at the exercise time in this case.

Table 5: Impact of basis functions on 5-asset asymmetric Bermudan max call options

	$S_0 = 90$			$S_0 = 100$			$S_0 = 110$		
b	$\hat{L}_0^S$	$\hat{L}_0^A$	$\Delta\hat{L}_0$	$\hat{L}_0^S$	$\hat{L}_0^A$	$\Delta\hat{L}_0$	$\hat{L}_0^S$	$\hat{L}_0^A$	$\Delta\hat{L}_0$
18	27.049	27.517	+0.468	37.089	37.807	+0.718	48.408	49.254	+0.846
12	27.325	27.480	+0.155	37.529	37.746	+0.217	48.910	49.175	+0.265

Note: Option parameters are the same as in Table 4. $\hat{L}_0^S$ represents the lower bound estimator using symmetric (non-distinguishing) basis functions, $\hat{L}_0^A$ is the estimator using asymmetric (distinguishing) basis functions. The 95% upper bounds for $S_0 = 90, 100, 110$ are respectively 27.686, 38.020, and 49.531.

Table 6: Lower bound improvements by local policy enhancement (5-asset Bermudan max call)

	$S_0 = 90$			$S_0 = 100$			$S_0 = 110$		
b	$\hat{L}_0$	$\hat{L}_0^E$	$\Delta\hat{L}_0$	$\hat{L}_0$	$\hat{L}_0^E$	$\Delta\hat{L}_0$	$\hat{L}_0$	$\hat{L}_0^E$	$\Delta\hat{L}_0$
6	16.563	16.613	+0.050	26.040	26.108	+0.068	36.632	36.713	+0.081
12	16.606	16.629	+0.023	26.106	26.138	+0.032	36.715	36.756	+0.041
18	16.627	16.634	+0.007	26.125	26.139	+0.014	36.722	36.750	+0.028
19	16.618	16.630	+0.012	26.113	26.134	+0.021	36.705	36.739	+0.034

Note: Option parameters are the same as in Table 3 except that the local policy enhancement is used here with $\epsilon = 0.05$ and $N_\epsilon = 500$. $\hat{L}_0$ is the regular lower bound estimator, $\hat{L}_0^E$ is the estimator with local policy enhancement. The 95% upper bounds for $S_0 = 90, 100, 110$ are, respectively, 16.652, 26.170, and 36.804.

estimator, especially when the original exercise policy is far from optimal. The lower bound estimator using 12 basis functions with local policy enhancement consistently outperforms the estimator using 18 basis functions without local policy enhancement. This indicates that the local policy enhancement can help reduce the number of basis functions needed for regression in order to achieve the same level of accuracy.

Table 7 shows the effective saving factor by sub-optimality checking and boundary distance grouping, which is calculated as the ratio of the product of computational time and variance of estimator, with and without improvements. Both methods are most effective on deep OTM options (for example, the $S_0 = 70$ case for a moving window Asian call shows an effective saving factor of more than 1000), and show considerable improvements for ATM and ITM options. The sub-optimality checking shows greater improvements in most cases, while the boundary distance grouping works better for deep ITM single asset and max Bermudan options, for which many sample paths are far above the exercise boundary thus will be placed in the 'zero' group and save the computation. Also notice that the improvements of two methods are not orthogonal, especially for OTM options, since the time saving for both methods comes mainly from the sample paths that never go beyond the exercise lower limit. More specifically, the expected time for estimating upper bound increment from a group Z path is shorter than that from a group $\bar{Z}$ path after applying the sub-optimality checking, which limits the additional saving through boundary distance grouping.

Table 7: Effective saving factor by sub-optimality checking and boundary distance grouping

Option type	S_0	t_H	t_H^S	ESF^{S*}	t_H^B	ESF^B	$t_H^{S,B}$	$ESF^{S,B}$
Single	70	4.22	0.01	739.7	0.03	196.8	0.00	$N.A.^{**}$
asset	90	5.87	0.24	32.3	1.13	6.9	0.10	92.3
Bermudan	110	7.90	3.26	2.3	3.71	2.6	1.95	6.0
call	130	3.58	4.21	1.2	1.18	4.5	1.08	5.4
Moving	70	3.05	0.01	1163.1	0.06	69.4	0.00	1172.1
window	90	5.00	0.15	43.3	2.30	2.7	0.11	62.0
Asian	110	7.97	1.04	9.1	4.88	1.9	0.55	16.5
call	130	8.37	2.33	4.4	6.33	1.3	1.68	5.7
Bermudan	70	2.72	0.10	26.5	0.32	8.2	0.06	47.2
max	90	3.49	1.01	3.4	1.16	2.4	0.46	7.4
call	110	4.04	3.93	1.0	2.70	1.6	1.34	1.6
	130	4.08	4.30	1.0	2.30	1.6	2.14	1.6

$*$: Effective saving factor (ESF) is defined in Section 4.

$**$: Due to zero upper bound increment after improvements being the denominator.

Note: Option parameters are the same as in Table 1, 2 and 3 respectively for three types of options. t_H, t_H^S, t_H^B, $t_H^{S,B}$ are the computational time for the primal-dual simulation estimates, respectively with no improvement, with sub-optimal checking, boundary distance grouping and two methods combined.

Table 8 demonstrates how the computational time increases with the number of exercise opportunities n, using the moving window Asian call option as an example. For the least-squares lower bound estimator, the computational time t_L increases linearly with n. After applying the local policy enhancement, dependence becomes between linear and quadratic, because sub-simulations are performed, but only when the path is close to the sub-optimal exercise boundary.

The computational time for primal-dual simulation algorithm, t_H, has a quadratic dependence on the number of exercise steps, as a sub-simulation is needed at every step of the sample path. By combining the sub-optimality checking and boundary distance grouping, the dependence becomes more than quadratic, while the computational time is actually reduced–that is because the boundary distance grouping is more effective for options with fewer exercise opportunities, in which case there are fewer sample paths in the 'non-zero' group.

Table 8: Computational time vs. number of exercise opportunities (moving window Asian call)

S_0	70				100				130			
n	t_L	t_L^E	t_H	$t_H^{S,B}$	t_L	t_L^E	t_H	$t_H^{S,B}$	t_L	t_L^E	t_H	$t_H^{S,B}$
10	0.01	0.01	0.20	0.00	0.02	0.14	0.35	0.00	0.02	0.55	0.46	0.05
25	0.02	0.03	1.13	0.00	0.04	0.73	2.08	0.05	0.05	1.76	2.54	0.31
50	0.04	0.08	4.31	0.00	0.08	2.01	8.00	0.25	0.10	4.22	9.59	1.68
100	0.07	0.21	16.83	0.01	0.15	5.51	30.89	1.30	0.21	11.36	37.23	8.64

Note: Option parameters are the same as in Table 2 except the number of exercise opportunities n. The window size is set as $m = n/5$ to ensure the consistent window length.

6 Conclusion

In this paper we introduce new variance reduction techniques and computational improvements for both lower and upper bound Monte Carlo methods for pricing American-style options.

Local policy enhancement may significantly improve the lower bound estimator of Bermudan option values, especially when the original exercise policy is far from optimal. Sub-optimality checking and boundary distance grouping are two methods that may reduce the computational time in duality-based upper bound algorithms by up to several hundred times. They both work best on out-of-the-money options. Sub-optimality checking is easy to implement and more effective in general, while boundary distance grouping performs better for options that are deep in-the-money. They can be combined to achieve a more significant reduction.

Tight lower and upper bounds for high-dimensional and path-dependent Bermudan options can be computed in the matter of seconds or a few minutes using the methods proposed here. Together they produce narrower confidence intervals using less computational time, by improving the exercise policy, reducing the variance of the estimators and saving unnecessary computations. For all the numerical examples tested, widths of 95% confidence intervals are within 1% of the option values, compared to 5% $\sim$ 10% before the improvements. And it takes up to 6 minutes to price each option, instead of several hours before the improvements. These improvements greatly enhance the practical use of Monte Carlo methods for pricing complicated American options.

References

[1] Andersen, L., and M. Broadie 2004, "A Primal-Dual Simulation Algorithm for Pricing Multi-Dimensional American Options," *Management Science* **50**, No. 9, 1222-1234.

[2] Boyle, P., M. Broadie, and P. Glasserman, 1997, "Monte Carlo Methods for Security Pricing," *Journal of Economic Dynamics and Control* **21**, 1267-1321.

[3] Broadie, M., and P. Glasserman, 1997, "Pricing American-Style Securities Using Simulation," *Journal of Economic Dynamics and Control* **21**, 8-9, 1323-1352.

[4] Broadie, M., and P. Glasserman, 2004, "A Stochastic Mesh Method for Pricing High-Dimensional American Option," *Journal of Computational Finance* **7**, No.4, 35-72.

[5] Broadie, M., P. Glasserman, and Z. Ha, 2000, "Pricing American Options by Simulation Using a Stochastic Mesh with Optimized Weights," *Probabilistic Constrained Optimization: Methodology and Applications*, 32-50, S. Uryasev, ed., Kluwer, Norwell, Mass.

[6] Carriere, J. F., 1996, "Valuation of the Early-Exercise Price for Derivative Securities using Simulations and Splines," *Insurance: Mathematics and Economics* **19**, 19-30.

[7] Clément, E., D. Lamberton, and P. Protter, 2002, "An Analysis of a Least Squares Regression Algorithm for American Option Pricing," *Finance and Stochastics* **17**, 448-471.

[8] Glasserman, P., 2003, "Monte Carlo Methods in Financial Engineering," *Springer.*

[9] Glasserman, P., and B. Yu, 2002, "Simulation for American Options: Regression Now or Regression Later?" *Monte Carlo and Quasi-Monte Carlo Methods 2002* (H. Niederreiter, ed.), 213-226, Springer, Berlin.

[10] Glasserman, P., and B. Yu, 2004, "Number of Paths Versus Number of Basis Functions in American Option Pricing," *Annals of Applied Probability* **14**, No.4, 2090-2119.

[11] Haugh, M., and L. Kogan, 2004, "Pricing American Options: A Duality Approach," *Operations Research* **52**, No.2, 258-270.

[12] Joshi, M. S., 2007, "A Simple Derivation of and Improvements to Jamshidian's and Rogers' Upper Bound Methods for Bermudan Options," *Applied Mathematical Finance,* **14**, No.3, 197-205.

[13] Lamberton, D., and B. Lapeyre, 1996, *Introduction to Stochastic Calculus Applied to Finance*, Chapman & Hall/CRC, Boca Raton.

[14] Longstaff, F., and E. Schwartz, 2001, "Valuing American Options By Simulation: A Simple Least-Squares Approach," *Review of Financial Studies* **14**, 113-148.

[15] Rasmussen, N., 2005, "Control Variates for Monte Carlo Valuation of American Options," *Journal of Computational Finance* **9**, No.1.

[16] Rogers, C., 2002, "Monte Carlo Valuation of American Options," *Mathematical Finance* **12**, No.3, 271-286.

[17] Tsitsiklis, J., and B. Van Roy, 1999, "Optimal Stopping of Markov Processes: Hilbert Space Theory, Approximation Algorithms, and an Application to Pricing High-Dimensional Financial Derivatives," *IEEE Transactions on Automatic Control* **44**, 1840-1851.

A Regression related issues

A.1 'Regression now' vs. 'regression later'

As a comparison with the standard least-squares estimator, we implement the 'regression later' technique (Broadie, Glasserman and Ha 2000, Glasserman and Yu 2002) on single asset Bermudan options and Bermudan max options. The 'regression later' approach requires the use of martingale basis functions, thus we choose S_t, S_t^2 and E_t (European option value at time t) as the basis functions for single asset Bermudan options and the first two powers of each asset price $\{S_{i,t},\ S_{i,t}^2\}$ for Bermudan max options. The corresponding martingale basis functions for S_t, S_t^2 and E_t are $e^{-(r-q)t}S_t$, $e^{-2(r-q)t-\sigma^2 t}S_t^2$ and $e^{-rt}E_t$.

As shown in Tables 9 and 10, 'regression later' approach generates more accurate estimates than 'regression now' in most cases, especially when fewer paths are used for regression and the exercise policy is far from optimal. However, when the number of regression paths are sufficiently large, 'regression later' does not lead to an observable improvement because the regression estimates approach their convergence limit and there is no room for additional improvement by 'regression later'. Note also that, the requirement of martingale basis functions limits the use of this algorithm.

A.2 Choice of basis functions

The choice of basis functions is critical for the regression-based algorithms. We list here the basis functions we use for our examples. Note that the constant c is counted as one of the basis functions.

For the single asset Bermudan options, we use 3 basis functions $\{c, E_t, E_t^2\}$ where E_t is the value of the European option at time t.

The 6 basis functions used for the moving window Asian options include the polynomials of S_t and A_t up to the second order $\{c,\ S_t,\ S_t^2,\ A_t,\ A_t^2,\ S_tA_t\}$.

Polynomials of sorted asset prices are used for the symmetric Bermudan max options. Let $S'_{i,t}$ be the i-th highest asset price at t. The 18 basis functions include the polynomials up to

Table 9: Regression now vs. regression later (single asset Bermudan call)

	$S_0 = 90$			$S_0 = 100$			$S_0 = 110$		
N_R	$\hat{L}_0^N$	$\hat{L}_0^L$	$\Delta\hat{L}_0$	$\hat{L}_0^N$	$\hat{L}_0^L$	$\Delta\hat{L}_0$	$\hat{L}_0^N$	$\hat{L}_0^L$	$\Delta\hat{L}_0$
10000	2.3759	2.3786	0.0027	5.8938	5.9017	0.0079	11.7141	11.7263	0.0122
100000	2.3832	2.3835	0.0003	5.9149	5.9154	0.0005	11.7442	11.7447	0.0005

Note: Option parameters are the same as in Table 1 except the number of basis functions $b = 4$ here.

Table 10: Regression now vs. regression later (5-asset asymmetric Bermudan max call)

	$S_0 = 90$			$S_0 = 100$			$S_0 = 110$		
N_R	$\hat{L}_0^N$	$\hat{L}_0^L$	$\Delta\hat{L}_0$	$\hat{L}_0^N$	$\hat{L}_0^L$	$\Delta\hat{L}_0$	$\hat{L}_0^N$	$\hat{L}_0^L$	$\Delta\hat{L}_0$
2000	27.268	27.310	0.042	37.473	37.489	0.016	48.817	48.858	0.041
200000	27.315	27.318	0.003	37.494	37.494	0.000	48.846	48.833	-0.013

Note: Option parameters are the same as in Table 3 except the number of basis functions $b = 11$ here.

the fifth order: $\{c,\ S_1',\ S_1'^2,\ S_2',\ S_2'^2,\ S_1'S_2',\ S_1'^3,\ S_2'^3,\ S_3',\ S_3'^2,\ S_1'S_3',\ S_2'S_3',\ S_1'^4,\ S_1'^5,\ S_1'^2S_2',\ S_1'S_2'^2,\ S_1'S_2'S_3',\ S_4'S_5'\}$; 12-case includes the polynomials of three highest asset prices, i.e., the first twelve in the list above; 6-case includes polynomials of two highest asset prices, i.e., the first six in the list above; 19-case is the same as in Longstaff and Schwartz (2001), which are $\{c,\ S_1',\ S_1'^2,\ S_1'^3,\ S_1'^4,\ S_1'^5,\ S_2',\ S_2'^2,\ S_3',\ S_3'^2,\ S_4',\ S_4'^2,\ S_5',\ S_5'^2,\ S_1'S_2',\ S_2'S_3',\ S_3'S_4',\ S_4'S_5',\ S_1'S_2'S_3'S_4'S_5'\}$.

The 18 asset-distinguishing basis functions for the asymmetric Bermudan max options are $\{c,\ S_1',\ S_1'^2,\ S_2',\ S_2'^2,\ S_1'S_2',\ S_1,\ S_2,\ S_3,\ S_4,\ S_5,\ S_1'^3,\ S_1^2,\ S_2^2,\ S_3^2,\ S_4^2,\ S_5^2,\ S_2'^3\}$; 12-case includes the first twelve in the list above.

For the comparisons between 'regression now' and 'regression later,' 4 basis functions $\{c, S_t, S_t^2, E_t\}$ are used for the single asset Bermudan options, 11 basis functions $\{c,\ S_{1,t},\ S_{2,t},\ S_{3,t},\ S_{4,t},\ S_{5,t},\ S_{1,t}^2,\ S_{2,t}^2,\ S_{3,t}^2,\ S_{4,t}^2,\ S_{5,t}^2\}$ are used for the asymmetric Bermudan max options.

B Proof of Proposition 1 and 2

In this section we give proofs of Propositions 1 and 2 in Section 4.

Proposition 1

(i) If $h_{t_i} \le \underline{Q_{t_i}}$ for $1 \le i \le k$, then $\pi_{t_k} = \frac{L_{t_k}}{B_{t_k}}$ and $\frac{h_{t_k}}{B_{t_k}} - \pi_{t_k} \le 0$.

(ii) If $h_{t_i} \le \underline{Q_{t_i}}$ for $l \le i \le k$, then $\pi_{t_k} = \pi_{t_{l-1}} - \frac{Q_{t_{l-1}}}{B_{t_{l-1}}} + \frac{L_{t_k}}{B_{t_k}}$ and $\frac{h_{t_k}}{B_{t_k}} - \pi_{t_k} \le \frac{Q_{t_{l-1}}}{B_{t_{l-1}}} - \pi_{t_{l-1}}$.

Proof.

(i) If $h_{t_i} \le \underline{Q_{t_i}}$ for $1 \le i \le k$, then $\mathbf{1}_{t_i} = 0$, and

$$\begin{aligned}\pi_{t_k} &= \pi_{t_{k-1}} + \frac{L_{t_k}}{B_{t_k}} - \frac{L_{t_{k-1}}}{B_{t_{k-1}}} \\ &= \pi_{t_{k-2}} + \frac{L_{t_{k-1}}}{B_{t_{k-1}}} - \frac{L_{t_{k-2}}}{B_{t_{k-2}}} + \frac{L_{t_k}}{B_{t_k}} - \frac{L_{t_{k-1}}}{B_{t_{k-1}}} \\ &= \pi_0 + \sum_{j=1}^{k} \left(\frac{L_{t_j}}{B_{t_j}} - \frac{L_{t_{j-1}}}{B_{t_{j-1}}} \right) = \frac{L_{t_k}}{B_{t_k}},\end{aligned}$$

therefore,

$$\frac{h_{t_k}}{B_{t_k}} - \pi_{t_k} \le \frac{\underline{Q_{t_k}}}{B_{t_k}} - \frac{L_{t_k}}{B_{t_k}} \le 0.$$

Notice that the last inequality holds for instance if $\underline{Q_t}$ is a $\mathbb{Q}$-sub-martingale, which is valid with the choice of one European option value but invalid with the choice of maximum among multiple option values.

(ii) If $h_{t_i} \le \underline{Q_{t_i}}$ for $l \le i \le k$

$$\begin{aligned}\pi_{t_k} &= \pi_{t_{k-1}} + \frac{L_{t_k}}{B_{t_k}} - \frac{L_{t_{k-1}}}{B_{t_{k-1}}} \\ &= \pi_{t_{k-2}} + \frac{L_{t_{k-1}}}{B_{t_{k-1}}} - \frac{L_{t_{k-2}}}{B_{t_{k-2}}} + \frac{L_{t_k}}{B_{t_k}} - \frac{L_{t_{k-1}}}{B_{t_{k-1}}} \\ &= \pi_{t_{l-1}} - \frac{Q_{t_{l-1}}}{B_{t_{l-1}}} + \frac{L_{t_k}}{B_{t_k}}.\end{aligned}$$

and

$$\frac{h_{t_k}}{B_{t_k}} - \pi_{t_k} = \frac{h_{t_k}}{B_{t_k}} - \pi_{t_{l-1}} + \frac{Q_{t_{l-1}}}{B_{t_{l-1}}} - \frac{L_{t_k}}{B_{t_k}} \le \frac{Q_{t_{l-1}}}{B_{t_{l-1}}} - \pi_{t_{l-1}}.\square$$

Proposition 2 *For a given sample path,*

(i) If $\exists \delta > 0$ such that $|\tilde{Q}_t - Q_t| < \delta$, and $d_t \ge \delta$ or $h_t \le \underline{Q_t}$ holds for $\forall t \in \Gamma$, then A_t is an increasing process and $D = 0$ for the path.

(ii) If $\exists\delta > 0$ such that $|\tilde{Q}_t - Q_t^*| < \delta$, and $d_t \geq \delta$ or $h_t \leq \underline{Q_t}$ holds for $\forall t \in \Gamma$, then $\mathbf{1}_t \equiv \mathbf{1}_t^*$.

Proof.

(i) For $\forall t_i \in \Gamma$, if $h_{t_i} \leq \underline{Q_{t_i}}$, $\mathbf{1}_{t_i} = 0$ and $A_{t_{i+1}} - A_{t_i} = 0$; otherwise if $d_{t_i} \geq \delta$,

(a) $\mathbf{1}_{t_i} = 0$, $A_{t_{i+1}} - A_{t_i} = 0$ and $h_{t_i} - \tilde{Q}_{t_i} < -\delta$,

$$h_{t_i} - L_{t_i} = h_{t_i} - Q_{t_i} = (h_{t_i} - \tilde{Q}_{t_i}) + (\tilde{Q}_{t_i} - Q_{t_i}) < -\delta + \delta = 0;$$

(b) $\mathbf{1}_{t_i} = 1$, $h_{t_i} = L_{t_i}$ and $h_{t_i} - \tilde{Q}_{t_i} > \delta$,

$$\begin{aligned} A_{t_{i+1}} - A_{t_i} &= \frac{h_{t_i}}{B_{t_i}} - E_{t_i}\left[\frac{L_{t_{i+1}}}{B_{t_{i+1}}}\right] \\ &= \frac{1}{B_{t_i}}(h_{t_i} - Q_{t_i}) \\ &= \frac{1}{B_{t_i}}[(h_{t_i} - \tilde{Q}_{t_i}) + (\tilde{Q}_{t_i} - Q_{t_i})] \\ &> \frac{1}{B_{t_i}}(\delta - \delta) = 0. \end{aligned}$$

Finally,

$$D = \max_t \left(\frac{h_t}{B_t} - \pi_t\right) = \max_t \left(\frac{h_t}{B_t} - \frac{L_t}{B_t} - A_t\right) = 0.$$

(ii) For $\forall t \in \Gamma$, if $h_t \leq \underline{Q_t}$, $\mathbf{1}_t = 0 = \mathbf{1}_t^*$; otherwise if $d_t \geq \delta$, since $|\tilde{Q}_t - Q_t^*| < \delta$,

(a) $\mathbf{1}_t = 0$, $Q_t^* > \tilde{Q}_t - \delta \geq h_t + \delta - \delta = h_t$ and thus $\mathbf{1}_t^* = 0$;

(b) $\mathbf{1}_t = 1$, $Q_t^* < \tilde{Q}_t + \delta \leq h_t - \delta + \delta = h_t$ and thus $\mathbf{1}_t^* = 1$.□

As we noted in the introduction, D may be interpreted as a penalty term for incorrect decisions. An incorrect continuation decision at t will be penalized and cause $\frac{h_t}{B_t} - \pi_t$ to be positive if the path never enters the exercise region before t. To see this, assume an incorrect continuation decision is made at t_i for a sample path that has not been in the exercise region before t_i. By Proposition 1, $\pi_{t_i} = \frac{L_{t_i}}{B_{t_i}}$ and $L_{t_i} \leq Q_{t_i}^* < h_{t_i} < \tilde{Q}_{t_i}$,

$$\frac{h_{t_i}}{B_{t_i}} - \pi_{t_i} = \frac{h_{t_i}}{B_{t_i}} - \frac{L_{t_i}}{B_{t_i}} \geq \frac{h_{t_i}}{B_{t_i}} - \frac{Q_{t_i}^*}{B_{t_i}} > 0.$$

On the other hand, an incorrect exercise decision at t will only get penalized (i.e., $D > 0$ for the path) either when the option leaves and re-enters the exercise region or the option never comes back into the exercise region and matures OTM. For example, suppose an incorrect exercise decision is made at t_k, and t_k is the first time that the sample path enters the exercise region. Assume the next time that the sample path enters the exercise region is at time t_l, i.e., $l = \inf_{i>k}\{i : \mathbf{1}_{t_i} = 1\} \wedge n$. By Proposition 1,

$$\pi_{t_l} = \pi_{t_k} - \frac{Q_{t_k}}{B_{t_k}} + \frac{h_{t_l}}{B_{t_l}} \quad = \quad \frac{L_{t_k}}{B_{t_k}} - \frac{Q_{t_k}}{B_{t_k}} + \frac{h_{t_l}}{B_{t_l}},$$

thus

$$\frac{h_{t_l}}{B_{t_l}} - \pi_{t_l} = \frac{Q_{t_k}}{B_{t_k}} - \frac{L_{t_k}}{B_{t_k}} \approx \frac{Q^*_{t_k}}{B_{t_k}} - \frac{L_{t_k}}{B_{t_k}} > \frac{h_{t_k}}{B_{t_k}} - \frac{L_{t_k}}{B_{t_k}} = 0,$$

assuming the sub-optimal continuation value is a good approximate of the true continuation value.

C Numerical details in boundary distance grouping

C.1 Estimators of the duality gap

The standard estimator of the duality gap D_0 is $\bar{D} = \frac{\sum_{i=1}^{N_H} D_i}{N_H}$, and the alternative estimator using boundary distance grouping is

$$\tilde{D} = \frac{\sum_{i=1}^{n_{\bar{Z}}} D_i + \frac{N_H - n_{\bar{Z}}}{l_Z} \sum_{i=n_{\bar{Z}}+1}^{n_{\bar{Z}}+l_Z} D_i}{N_H}.$$

Assume

$$E[D_i|\bar{Z}] = \mu_{\bar{Z}}, \mathrm{Var}[D_i|\bar{Z}] = \sigma^2_{\bar{Z}},$$

$$E[D_i|Z] = \mu_Z, \mathrm{Var}[D_i|Z] = \sigma^2_Z.$$

Since each path belongs to group $\bar{Z}$ with probability $p_{\bar{Z}}$ and group Z with probability $1-p_{\bar{Z}}$, $n_{\bar{Z}}$ is a binomial random variable with parameters $(N_H, p_{\bar{Z}})$. We have $E[n_{\bar{Z}}] = N_H p_{\bar{Z}}$, $\mathrm{Var}[n_{\bar{Z}}] = N_H p_{\bar{Z}}(1 - p_{\bar{Z}})$. By Proposition 2, if the $\tilde{Q}_t$ almost surely lies within δ_0 from the sub-optimal continuation value, we can choose $\delta = \delta_0$ so that $D = 0$ for almost all paths in group Z, thus μ_Z and σ_Z are both close to zero. The variance of the standard estimator is,

$$\mathrm{Var}[\bar{D}] \quad = \quad E[\mathrm{Var}(\bar{D}|n_{\bar{Z}})] + \mathrm{Var}(E[\bar{D}|n_{\bar{Z}}])$$

$$
\begin{aligned}
&= E\left[\frac{n_{\bar{Z}}\sigma_{\bar{Z}}^2 + (N_H - n_{\bar{Z}})\sigma_Z^2}{N_H^2}\right] + Var\left(\frac{n_{\bar{Z}}\mu_{\bar{Z}} + (N_H - n_{\bar{Z}})\mu_Z}{N_H}\right) \\
&= \frac{1}{N_H}\sigma_Z^2 + \frac{1}{N_H}p_{\bar{Z}}(\sigma_{\bar{Z}}^2 - \sigma_Z^2) + \frac{(\mu_{\bar{Z}} - \mu_Z)^2 Var(n_{\bar{Z}})}{N_H^2} \\
&= \frac{1}{N_H}[\sigma_Z^2 + p_{\bar{Z}}(\sigma_{\bar{Z}}^2 - \sigma_Z^2) + p_{\bar{Z}}(1 - p_{\bar{Z}})(\mu_{\bar{Z}} - \mu_Z)^2] \\
&= \frac{1}{N_H}[(1 - p_{\bar{Z}})\sigma_Z^2 + p_{\bar{Z}}\sigma_{\bar{Z}}^2 + p_{\bar{Z}}(1 - p_{\bar{Z}})(\mu_{\bar{Z}} - \mu_Z)^2] \\
&= \frac{\sigma_D^2}{N_H}.
\end{aligned}
$$

It is easy to verify the alternative estimator is unbiased,

$$
E[\tilde{D}] = E[E[\tilde{D}|n_{\bar{Z}}]] = E\left[\frac{n_{\bar{Z}}}{N_H}\mu_{\bar{Z}} + \frac{N_H - n_{\bar{Z}}}{N_H}\mu_Z\right] = p_{\bar{Z}}\mu_{\bar{Z}} + (1 - p_{\bar{Z}})\mu_Z = \mu,
$$

and the variance is

$$
\begin{aligned}
\mathrm{Var}[\tilde{D}] &= E[\mathrm{Var}(\tilde{D}|n_{\bar{Z}})] + \mathrm{Var}(E[\tilde{D}|n_{\bar{Z}}]) \\
&= E\left[\frac{1}{N_H^2}\sum_{i=1}^{n_{\bar{Z}}}\mathrm{Var}(D_i) + \frac{(N_H - n_{\bar{Z}})^2}{l_Z^2 N_H^2} Var\left(\sum_{i=n_{\bar{Z}}+1}^{n_{\bar{Z}}+l_Z} D_i\right)\right] + Var\left(\frac{n_{\bar{Z}}}{N_H}\mu_{\bar{Z}} + \frac{N_H - n_{\bar{Z}}}{N_H}\mu_Z\right) \\
&= E\left[\frac{1}{N_H^2}n_{\bar{Z}}\sigma_{\bar{Z}}^2 + \frac{(N_H - n_{\bar{Z}})^2}{l_Z^2 N_H^2} l_Z\sigma_Z^2\right] + \frac{(\mu_{\bar{Z}} - \mu_Z)^2}{N_H^2}\mathrm{Var}(n_{\bar{Z}}) \\
&= \frac{p_{\bar{Z}}\sigma_{\bar{Z}}^2}{N_H} + \frac{\sigma_Z^2}{l_Z N_H^2}E[N_H^2 - 2n_{\bar{Z}}N_H + n_{\bar{Z}}^2] + \frac{(\mu_{\bar{Z}} - \mu_Z)^2}{N_H}p_{\bar{Z}}(1 - p_{\bar{Z}}) \\
&= \frac{1}{l_Z}(1 - p_{\bar{Z}})^2\sigma_Z^2 + \frac{1}{l_Z N_H}p_{\bar{Z}}(1 - p_{\bar{Z}})\sigma_Z^2 + \frac{1}{N_H}p_{\bar{Z}}\sigma_{\bar{Z}}^2 + \frac{1}{N_H}p_{\bar{Z}}(1 - p_{\bar{Z}})(\mu_{\bar{Z}} - \mu_Z)^2.
\end{aligned}
$$

For any choice of $l_Z \leq N_H(1 - p_{\bar{Z}})$, where $N_H(1 - p_{\bar{Z}})$ is the expected number of paths in Z,

$$
\begin{aligned}
\mathrm{Var}[\tilde{D}] &> \frac{1}{l_Z}(1 - p_{\bar{Z}})^2\sigma_Z^2 + \frac{1}{N_H}p_{\bar{Z}}\sigma_{\bar{Z}}^2 + \frac{1}{N_H}p_{\bar{Z}}(1 - p_{\bar{Z}})(\mu_{\bar{Z}} - \mu_Z)^2 \\
&\geq \frac{1}{N_H}[(1 - p_{\bar{Z}})\sigma_Z^2 + p_{\bar{Z}}\sigma_{\bar{Z}}^2 + p_{\bar{Z}}(1 - p_{\bar{Z}})(\mu_{\bar{Z}} - \mu_Z)^2] \\
&= \frac{\sigma_D^2}{N_H} = \mathrm{Var}[\bar{D}],
\end{aligned}
$$

which means the variance of $\tilde{D}$ is always greater than that of $\bar{D}$. The '=' sign in the second inequality holds when $l_Z = N_H(1 - p_{\bar{Z}})$, in which case $\mathrm{Var}[\bar{D}]$ and $\mathrm{Var}[\tilde{D}]$ are only different by a small term $\frac{1}{N_H^2}p_{\bar{Z}}\sigma_Z^2$. The difference is due to the randomness in the number of paths in group $\bar{Z}$ and Z. Generally,

$$\begin{aligned}\mathrm{Var}[\tilde{D}] &< \frac{\sigma_Z^2}{l_Z} + \frac{1}{N_H}[(1-p_{\bar{Z}})\sigma_Z^2 + p_{\bar{Z}}\sigma_{\bar{Z}}^2 + p_{\bar{Z}}(1-p_{\bar{Z}})(\mu_{\bar{Z}} - \mu_Z)^2] \\ &= \frac{\sigma_D^2}{N_H} + \frac{\sigma_Z^2}{l_Z},\end{aligned}$$

i.e., $\mathrm{Var}[\tilde{D}]$ falls into $[\sigma_D^2/N_H, \sigma_D^2/N_H + \sigma_Z^2/l_Z]$, i.e., $[\mathrm{Var}[\bar{D}], \mathrm{Var}[\bar{D}] + \sigma_Z^2/l_Z]$, which is a tight interval if $\sigma_Z^2 \ll l_Z\sigma_D^2/N_H$.

C.2 Effective saving factor

The effective saving factor for boundary distance grouping can be calculated as the ratio of the efficiency before and after improvement, where the efficiency of simulation is measured by the product of sample variance and computational time. As denoted in Section 4, T_P is the expected time spent for generating one sample path, T_I is the expected time to identify which group the path belongs to, $T_{D_{\bar{Z}}}$ and T_{D_Z} are the expected time to estimate upper bound increment D from a group $\bar{Z}$ path and from a group Z path respectively, typically $T_P, T_I \ll T_{D_{\bar{Z}}}, T_{D_Z}$.

The total expected time for estimating $\bar{D}$ is,

$$T_{\bar{D}} \approx N_H T_P + N_H p_{\bar{Z}} T_{D_{\bar{Z}}} + N_H(1-p_{\bar{Z}})T_{D_Z} = N_H[T_P + p_{\bar{Z}}T_{D_{\bar{Z}}} + (1-p_{\bar{Z}})T_{D_Z}],$$

and for $\tilde{D}$,

$$T_{\tilde{D}} \approx N_H T_P + N_H T_I + p_{\bar{Z}} N_H T_{D_{\bar{Z}}} + l_Z T_{D_Z} = N_H(T_P + T_I + p_{\bar{Z}}T_{D_{\bar{Z}}}) + l_Z T_{D_Z}.$$

For a fixed boundary distance threshold $\delta > 0$, parameters $p_{\bar{Z}}, \mu_{\bar{Z}}, \mu_Z, \sigma_{\bar{Z}}^2, \sigma_Z^2$ can be estimated from simulation. We may maximize the effective saving factor with respect to l_Z (the number of paths selected from group Z to estimate D) for a fixed δ and find the optimal δ' from a pre-selected set of δ choices,

$$\delta' := \arg(\min_{\delta} \mathrm{Var}[\tilde{D}] \cdot T_{\tilde{D}}). \tag{26}$$

The variance and efficiency measure for $\tilde{D}$ are

$$\mathrm{Var}[\tilde{D}] = \frac{1}{l_Z}(1-p_{\bar{Z}})\sigma_Z^2[(1-p_{\bar{Z}}) + \frac{1}{N_H}p_{\bar{Z}}] + \frac{1}{N_H}p_{\bar{Z}}\sigma_{\bar{Z}}^2 + \frac{1}{N_H}p_{\bar{Z}}(1-p_{\bar{Z}})(\mu_{\bar{Z}} - \mu_Z)^2$$

$$\begin{aligned}
&\approx \frac{1}{l_Z}(1-p_{\bar{Z}})^2\sigma_Z^2 + \frac{1}{N_H}p_{\bar{Z}}\sigma_{\bar{Z}}^2 + \frac{1}{N_H}p_{\bar{Z}}(1-p_{\bar{Z}})(\mu_{\bar{Z}}-\mu_Z)^2, \\
\mathrm{Var}[\tilde{D}]\cdot T_{\tilde{D}} &= \frac{1}{l_Z}N_H(1-p_{\bar{Z}})^2\sigma_Z^2(T_P+T_I+p_{\bar{Z}}T_{D_{\bar{Z}}}) \\
&\quad + l_Z\frac{p_{\bar{Z}}\sigma_{\bar{Z}}^2 T_{D_Z}}{N_H} + l_Z\frac{p_{\bar{Z}}(1-p_{\bar{Z}})(\mu_{\bar{Z}}-\mu_Z)^2 T_{D_Z}}{N_H} + \text{constant}.
\end{aligned}$$

Take the partial derivative with respect to l_Z,

$$\begin{aligned}
&\frac{\partial(\mathrm{Var}[\tilde{D}]\cdot T_{\tilde{D}})}{\partial l_Z}|_{l_Z=l'_Z} \\
= &-\frac{1}{l'^2_Z}N_H(1-p_{\bar{Z}})^2\sigma_Z^2(T_P+T_I+p_{\bar{Z}}T_{D_{\bar{Z}}}) + \frac{p_{\bar{Z}}\sigma_{\bar{Z}}^2 T_{D_Z}}{N_H} + \frac{p_{\bar{Z}}(1-p_{\bar{Z}})(\mu_{\bar{Z}}-\mu_Z)^2 T_{D_Z}}{N_H},
\end{aligned}$$

thus the product achieves its minimum at

$$l'_Z = \sqrt{\frac{(1-p_{\bar{Z}})^2\sigma_Z^2}{p_{\bar{Z}}T_{D_Z}} \cdot \frac{T_P+T_I+p_{\bar{Z}}T_{D_{\bar{Z}}}}{\sigma_{\bar{Z}}^2+(1-p_{\bar{Z}})(\mu_{\bar{Z}}-\mu_Z)^2}}N_H = \gamma N_H,$$

where γ denotes the portion of the sample paths that we should choose to estimate the group Z average. As

$$\frac{\partial^2(\mathrm{Var}[\tilde{D}]\cdot T_{\tilde{D}})}{\partial l_Z^2} = \frac{2}{l_Z^3}N_H(1-p_{\bar{Z}})^2\sigma_Z^2(T_P+T_I+p_{\bar{Z}}T_{D_{\bar{Z}}}) > 0,$$

the function is strictly convex and l'_Z is the unique minimum.

The effective saving factor of boundary distance grouping can be calculated as,

$$\begin{aligned}
&ESF = \frac{\mathrm{Var}[\bar{D}]\cdot T_{\bar{D}}}{\mathrm{Var}[\tilde{D}]\cdot T_{\tilde{D}}}|_{l_Z=l'_Z} \\
= &\frac{[(1-p_{\bar{Z}})\sigma_Z^2 + p_{\bar{Z}}\sigma_{\bar{Z}}^2 + p_{\bar{Z}}(1-p_{\bar{Z}})(\mu_{\bar{Z}}-\mu_Z)^2][T_P + p_{\bar{Z}}T_{D_{\bar{Z}}} + (1-p_{\bar{Z}})T_{D_Z}]}{[\frac{(1-p_{\bar{Z}})^2}{\gamma}\sigma_Z^2 + p_{\bar{Z}}\sigma_{\bar{Z}}^2 + p_{\bar{Z}}(1-p_{\bar{Z}})(\mu_{\bar{Z}}-\mu_Z)^2](T_P+T_I+p_{\bar{Z}}T_{D_{\bar{Z}}}+\gamma T_{D_Z})}.
\end{aligned}$$

Assume $p_{\bar{Z}} \approx 0$, $\mu_Z \approx 0$ and $p_{\bar{Z}}T_{D_{\bar{Z}}} \gg T_P + T_I$, then

$$\gamma \approx \sqrt{\frac{\sigma_Z^2 T_{D_{\bar{Z}}}}{(\sigma_{\bar{Z}}^2+\mu_{\bar{Z}}^2)T_{D_Z}}}.$$

If in addition we have $\sigma_Z^2 \ll \frac{l'_Z\sigma_D^2}{N_H}$ which leads to $\mathrm{Var}[\bar{D}] \approx \mathrm{Var}[\tilde{D}]$, the effective saving is essentially the saving of time spent for estimating D,

$$ESF \approx \frac{p_{\bar{Z}}T_{D_{\bar{Z}}} + (1-p_{\bar{Z}})T_{D_Z}}{p_{\bar{Z}}T_{D_{\bar{Z}}} + \gamma T_{D_Z}}.$$

If the boundary distance can effectively identify the paths with zero upper bound increment, group Z will have approximately zero mean and variance, thus $\gamma \approx 0$ and

$$ESF \approx 1 + \frac{T_{D_Z}}{p_{\bar{Z}} T_{D_{\bar{Z}}}}.$$

C.3 Two ways to estimate Var$[\tilde{D}]$

We can not directly estimate the variance of the alternative estimator $\tilde{D}$ from D_is because they are not i.i.d. random variables after grouping, in this section we show two indirect ways to estimate it.

One is through the batching procedure, in which we estimate the batch mean $\tilde{D}^{(j)}$ with boundary distance grouping from k independent batches ($j = 1, ..., k$), each using $\lfloor \frac{N_H}{k} \rfloor$ simulation paths. $\tilde{D}^{(j)}$ are i.i.d. estimates of $\tilde{D}$ and the sample variance can be calculated by

$$\hat{s}_D^2 = \frac{\sum_{j=1}^{k} (\tilde{D}^{(j)} - \frac{\sum_{j=1}^{k} \tilde{D}^{(j)}}{k})^2}{k-1},$$

which is an unbiased estimator of $\mathrm{Var}[\tilde{D}]$.

The other alternative is to use a modified sample variance to approximate it. Similar to the regular sample variance, a modified sample variance estimator $\hat{s}$ can be constructed as below,

$$\hat{\theta} = \frac{1}{N_H(N_H-1)} \left[\sum_{i=1}^{n_{\bar{Z}}} (D_i - \tilde{D})^2 + \frac{N_H - n_{\bar{Z}}}{l_Z} \sum_{i=n_{\bar{Z}}+1}^{n_{\bar{Z}}+l_Z} (D_i - \tilde{D})^2 \right],$$

or equivalently,

$$\hat{\theta} = \frac{1}{N_H(N_H-1)} \left[\sum_{i=1}^{n_{\bar{Z}}} D_i^2 + \frac{N_H - n_{\bar{Z}}}{l_Z} \sum_{i=n_{\bar{Z}}+1}^{n_{\bar{Z}}+l_Z} D_i^2 - N_H \tilde{D}^2 \right].$$

The expectation of $\hat{\theta}$ is

$$\begin{aligned}
E[\hat{\theta}] &= E\left[\frac{\sum_{i=1}^{n_{\bar{Z}}} (D_i^2 - 2D_i\tilde{D} + \tilde{D}^2) + \frac{N_H - n_{\bar{Z}}}{l_Z} \sum_{i=n_{\bar{Z}}+1}^{n_{\bar{Z}}+l_Z} (D_i^2 - 2D_i\tilde{D} + \tilde{D}^2)}{N_H(N_H-1)} \right] \\
&= E\left[\frac{n_{\bar{Z}}(\mu_{\bar{Z}}^2 + \sigma_{\bar{Z}}^2) - 2N_H\tilde{D}^2 + (N_H - n_{\bar{Z}})(\mu_Z^2 + \sigma_Z^2) + N_H\tilde{D}^2}{N_H(N_H-1)} \right] \\
&= E\left[\frac{n_{\bar{Z}}(\mu_{\bar{Z}}^2 + \sigma_{\bar{Z}}^2) + (N_H - n_{\bar{Z}})(\mu_Z^2 + \sigma_Z^2) - N_H\bar{D}^2}{N_H(N_H-1)} \right] - E\left[\frac{\tilde{D}^2 - \bar{D}^2}{N_H - 1} \right]
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{Var}[\bar{D}] - E\left[\frac{\tilde{D}^2 - \bar{D}^2}{N_H - 1}\right] \\
&= \mathrm{Var}[\tilde{D}] - E\left[\frac{N_H(\tilde{D}^2 - \bar{D}^2)}{N_H - 1}\right] \\
&= \mathrm{Var}[\tilde{D}] - \frac{\sigma_Z^2}{l_Z}\frac{(N_H + p_{\bar{Z}} - l_Z)(1 - p_{\bar{Z}})}{N_H - 1} \\
&\geq \mathrm{Var}[\tilde{D}] - \frac{\sigma_Z^2}{l_Z}(1 - p_{\bar{Z}}) \\
&\geq \mathrm{Var}[\tilde{D}] - \frac{\sigma_Z^2}{l_Z}.
\end{aligned}
$$

Thus our new estimator can be constructed as

$$
\begin{aligned}
\hat{s}^2_{\tilde{D}} &= \hat{\theta} + \frac{\hat{\sigma}_Z^2}{l_Z} \\
&= \frac{\sum_{i=1}^{n_{\bar{Z}}}(D_i - \tilde{D})^2 + \frac{N_H - n_{\bar{Z}}}{l_Z}\sum_{i=n_{\bar{Z}}+1}^{n_{\bar{Z}}+l_Z}(D_i - \tilde{D})^2}{N_H(N_H - 1)} + \frac{\sum_{i=n_{\bar{Z}}+1}^{n_{\bar{Z}}+l_Z}(D_i - \frac{\sum_{i=n_{\bar{Z}}+1}^{n_{\bar{Z}}+l_Z} D_i}{l_Z})^2}{l_Z(l_Z - 1)},
\end{aligned}
$$

whose expectation is

$$
E[\hat{s}^2_{\tilde{D}}] = E[\hat{\theta}] + \frac{\sigma_Z^2}{l_Z} \in \left[\mathrm{Var}[\tilde{D}], \mathrm{Var}[\tilde{D}] + \frac{\sigma_Z^2}{l_Z}\right].
$$

Although the modified sample variance is not an unbiased estimator of the true variance, it closely bounds the true variance from above. Since $\mathrm{Var}[\tilde{D}] \in \left[\frac{\sigma_D^2}{N_H}, \frac{\sigma_D^2}{N_H} + \frac{\sigma_Z^2}{l_Z}\right]$, we have $E[\hat{s}^2_{\tilde{D}}] \in \left[\frac{\sigma_D^2}{N_H}, \frac{\sigma_D^2}{N_H} + 2\frac{\sigma_Z^2}{l_Z}\right]$, which is a tight interval if $\sigma_Z^2 \ll \frac{l_Z\sigma_D^2}{N_H}$.