Working Paper #  2013- 04

Efficient Monte Carlo Counterparty

Credit Risk Pricing and Measurement

Samin Ghamami, University of California at Berkeley
Bo Zhang, IBM Thomas J. Watson Research Center, NY

September 17, 2013

University of California
Berkeley
Efficient Monte Carlo Counterparty Credit Risk Pricing and Measurement

Samim Ghamami† and Bo Zhang‡

September 17, 2013

Abstract

Counterparty credit risk (CCR), a key driver of the 2007-08 credit crisis, has become one of the main focuses of the major global and U.S. regulatory standards. Financial institutions invest large amounts of resources employing Monte Carlo simulation to measure and price their counterparty credit risk. We develop efficient Monte Carlo CCR estimation frameworks by focusing on the most widely used and regulatory-driven CCR measures: expected positive exposure (EPE), credit value adjustment (CVA), and effective expected positive exposure (eEPE). Our numerical examples illustrate that our proposed efficient Monte Carlo estimators outperform the existing crude estimators of these CCR measures substantially in terms of mean square error (MSE). We also demonstrate that the two widely used sampling methods, the so-called Path Dependent Simulation (PDS) and Direct Jump to Simulation date (DJS), are not equivalent in that they lead to Monte Carlo CCR estimators which are drastically different in terms of their MSE.

1 Introduction and a Summary of Important CCR Measures

Counterparty credit risk (CCR) is the risk that a party to a derivative contract may default prior to the expiration of the contract and fail to make the required contractual payments, (see [5] for the basic CCR definitions). Counterparty credit risk has been widely considered as one of the key drivers of the 2007-08 credit crisis, and it has become one of main focuses of the major global and U.S. regulatory frameworks (Basel III† and the Dodd-Frank Act of 2009-10; see, for instance, [3]). It is well known that pricing and measuring counterparty credit risk is computationally extremely intensive; financial institutions (derivative dealers) invest large amounts of resources

---

*Samim Ghamami is grateful to Robert Anderson, Lisa Goldberg, and Travis Nesmith for helpful discussions.
†Federal Reserve Board, Washington, D.C., email: samim.ghamami@frb.gov, and Center for Risk Management Research, University of California, Berkeley email: samim.ghamami@berkeley.edu.
‡Business Analytics and Mathematical Sciences Department, IBM Thomas J. Watson Research Center, Yorktown Heights, NY, 10598, email:bozhang@gatech.edu.
†Basel III is a global regulatory standard on bank capital adequacy, stress testing and market liquidity risk agreed upon by the members of the Basel Committee on Banking Supervision in 2010-11, and scheduled to be introduced from 2013 until 2018.
developing and maintaining Monte Carlo simulation “engines” to manage their counterparty risk, (see [20], [15], and [5]). While various aspects of counterparty credit risk have been subject of extensive research post 2007-08 financial crisis, statistical efficiency of the CCR estimators has received no attention in the literature.

In this paper we develop efficient Monte Carlo frameworks for pricing and measuring counterparty risk. More specifically, we focus on efficient Monte Carlo estimation of the most widely used and regulatory-driven CCR measures, expected positive exposure (EPE), credit value adjustment (CVA), and effective EPE (eEPE), as defined below. Efficiency criteria under consideration are variance, bias, and computing time of the Monte Carlo estimators. Our proposed Monte Carlo estimators of EPE, CVA, and eEPE outperform the existing crude estimators of these CCR measures substantially in terms of mean square error (MSE). To the best of our knowledge, this paper is the first to consider efficiency improvement for Monte Carlo CCR estimation. Currently, CVA is a CCR measure that is only applied to bilateral derivatives transactions. However, EPE and eEPE are CCR measures applicable to both bilateral derivatives transactions and centrally-cleared derivatives transactions. Specifically, Basel Committee on Banking Supervision (BCBS) has devised regulatory capital charges on clearing member banks against their central counterparty credit risk; EPE and eEPE are components of these capital charges to be estimated via Monte Carlo simulation by large dealer banks.

Counterparty credit exposure [5], denoted by $V$, of a financial institution against one of its counterparties, is the larger of zero and the market value of the portfolio of derivatives contracts the financial institution holds with this counterparty. To effectively introduce our efficient Monte Carlo procedures, we consider credit exposures in the absence of the commonly used risk mitigants, collateral and netting agreements. This simple setting facilitates the communication of our main results.

EPE is a widely used counterparty credit risk measure for regulatory and economic capital calculations, (see Chapters 2 and 11 of [15]). It is defined as follows,

$$EPE \equiv \int_0^T E[V_t] dt,$$

where $E[V_t]$ is the expected value of the (credit) exposure at time $t \geq 0$, and $T > 0$ denotes the time to maturity of the longest transaction in the derivatives portfolio.

Effective EPE (eEPE), another widely used regulatory and economic capital-related counterparty risk measure [15] is defined as follows in the CCR literature:

$$eEPE_{dst} \equiv \sum_{i=1}^{n} \max_{1 \leq j \leq i} E[V_j] \Delta_i,$$

This definition is based on a discrete time grid, $0 \equiv t_0 < t_1 < \ldots < t_n \equiv T$ with $\Delta_i = t_i - t_{i-1}$, $i = 1, \ldots, n$. We prefer and propose the following continuous version of eEPE:

$$eEPE \equiv \int_0^T \max_{0 \leq u \leq t} E[V_u] dt,$$
which is consistent with the definition of EPE and has the advantage of not requiring an a priori specification of a discrete time grid. Our results in Section 5 apply to eEPE as well as eEPE_{dst}.

eEPE is the “conservative” version of EPE that accounts for roll-over risk. Roll-over risk refers to the following scenario. Expiration of some of the short-term trades in the derivatives portfolio before \( T \) would decrease some of the \( E[V_t] \) and so EPE. However, it is likely that these short-term trades are replaced by new ones. When these replacements are not captured by the Monte Carlo CCR “engine”, EPE is underestimated, (see [20]).

CVA, which is the difference between the risk free portfolio value and the true counterparty default risky portfolio value, (see [19]), has become one of the main focuses of the Basel III; derivative dealers are required to calculate CVA charges for each of their counterparties on a frequent basis.

Let \( \tau \), a positive random variable, denote the default time of the counterparty. It can be shown that CVA, the price of the counterparty credit risk, is equal to the risk neutral expected discounted loss, i.e.,

\[
\text{CVA} \equiv E[(1 - R)D_{\tau}V_{\tau}1\{\tau \leq T\}],
\]

where \( 1\{A\} \) is the indicator of the event \( A \), \( D_t = B_0/B_t \) is the stochastic discount factor at time \( t \), \( B_t \) is the value of the money market account at time \( t \), and \( R \) is the financial institution’s recovery rate, (see, for instance, Chapter 7 of [15] for a derivation of this formula). Hereafter we suppress the dependence of the CVA on the recovery rate, \( R \). When \( V \) and \( \tau \) are assumed to be independent, we refer to CVA as independent CVA. Let \( F \) denote the cumulative distribution function of \( \tau \). Independent CVA can be written as follows,

\[
\text{CVA}_I \equiv E[D_{\tau}V_{\tau}1\{\tau \leq T\}] = \int_0^T E[D_tV_t]dF_t,
\]

where the last equality follows from conditioning on \( \tau \), the independence of \( V \) and \( \tau \), and the independence of \( D \) and \( \tau \). We focus on efficient Monte Carlo estimation of independent CVA in this paper.  \(^2\)

EPE, effective EPE, and independent CVA are estimated based on the Reimann sum approximation of the integrals in (1), (3), and (5) and Monte Carlo estimation of expected exposures, \( E[V_t] \), and expected discounted exposures, \( E[D_tV_t] \).

Section 2 summarizes the common features of the Monte Carlo CCR framework widely used by financial institutions and introduces the notion of Marginal Matching, which enables us to define and differentiate the two widely used CCR sampling methods, Path Dependent Simulation (PDS) and Direct Jump to Simulation date (DJS). These two terms were first introduced by Pykhtin and Zhu in 2006 [20]. Practitioners choose either of the sampling methods arbitrarily.  \(^3\)

A recurring theme of Sections 3 through 5 of this paper is to illustrate that PDS and DJS-based CCR estimators have drastically different MSE. Section 3 introduces an efficient Monte Carlo

\(^2\)Wrong (right) way risk are referred to as cases where credit exposures are negatively (positively) correlated with the credit quality of the counterparty, (see [10], [5], and [16]).

\(^3\)One of the authors’ former employer is a large investment bank.
framework for estimating EPE, which also directly applies to efficient CVA estimation. Using our results in Section 3, we summarize our proposed Monte Carlo framework for efficient estimation of CVA\(_I\) in Section 4. Using our results in Section 3, Section 5 considers efficient Monte Carlo estimation of eEPE. Our numerical examples indicate that employing our Monte Carlo CCR schemes leads to substantial MSE reduction. We would like to emphasize that Sections 4 and 5 should not be read independently. The main components of the proposed efficient CCR framework are developed in Section 3. As will be seen in the sequel, this is because EPE and CVA\(_I\) are both weighted sums of expected exposures, and the ideas developed for efficient EPE and CVA\(_I\) estimation have implications for efficient eEPE estimation.

2 Monte Carlo Counterparty Credit Risk Estimation

Contract level credit exposure at time \(t > 0\) is the maximum of the contract’s market value and zero, \(\max\{C_t, 0\}\), where \(C_t\) denotes the time-\(t\) value of the derivative contract. Consider a financial institution that holds a portfolio of \(k\) derivative contracts with its counterparty. Counterparty level credit exposure is

\[
V_t = \sum_{i=1}^{k} \max\{C^i_t, 0\}, \quad (6)
\]

where \(C^i_t\) denotes the time-\(t\) value of the \(i\)'th derivative contract in the derivatives portfolio. When risk mitigants are employed, \(V_t\) is defined differently. For instance, in the presence of netting agreements, credit exposure becomes, (see [19]),

\[
V_t = \max\{\sum_{i=1}^{k} C^i_t, 0\}. \quad (7)
\]

A typical Monte Carlo counterparty risk engine of a derivatives dealer estimates various types of CCR measures based on sampling from the credit exposure process on a time grid, \(0 < t_1 < \ldots < t_n = T\), where \(T\) denotes the maturity of the longest transaction in a portfolio of derivatives and \(t_1, \ldots, t_n\) are sometimes referred to as valuation points. Set \(V_t = V_{t_i}\).

Some of the CCR measures are static in the sense that they are defined based on a given fixed time point. Expected exposure (EE) at time \(t_i\), is simply \(E[V_t]\). Also, Value at Risk (VaR) type of measures for a given valuation point \(t_i\) is referred to as potential future exposure. Derivatives dealers use Monte Carlo simulation to estimate EE and PFE for all the given valuation points \(t_1, \ldots, t_n\) on a frequent basis, (see [15] and [19] for more details). Note that the CCR measures considered in this paper, EPE, CVA, and eEPE, are dynamic in the sense that they depend on the time evolution of the credit exposure process.

In what follows we first summarize the simulation of the credit exposure process. Then, we introduce the notion of Marginal Matching in sampling from the time evolution of the credit exposure process.
2.1 Simulating the Credit Exposure Process

Suppose that credit exposure is a stochastic process \( \{V_t ; t \geq 0\} \) defined on a given filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq \infty}, \mathbb{P})\). Given (6) and (7), \( V_t \) can be viewed as a function of the stochastic processes that drive the values of the derivative contracts, \( C_1, \ldots, C_k \). In risk management, these underlying stochastic processes are usually referred to as risk factors, e.g., interest rates, commodity prices, and equity prices. To generate a Monte Carlo realization of \( V_t \), for a fixed \( t > 0 \), first, the underlying risk factors should be sampled from up to time \( t > 0 \). Next, given the Monte Carlo realization of the risk factors up to time \( t > 0 \), the derivative contracts should be valued. This two-step procedure leads to a single Monte Carlo realization of \( V_t \). It is a risk management common practice to use the physical probability measure in the first step and the risk-neutral measure in the second. This applies to Monte Carlo estimation of EPE and eEPE. However, since CVA is usually viewed as the market price of counterparty credit risk, risk-neutral measure is usually used in both steps. Depending on the complexity of the payoff function of the derivative contracts, the valuation step could take straightforward Black-Scholes-type analytical calculations, or it could demand approximations that depending on the desired level of accuracy might be computationally intensive. These approximations could also involve Monte Carlo simulation: Nested Monte Carlo refers to the use of a second layer of Monte Carlo simulation in the valuation step of the above procedure, (see [14]), and regression-based Monte Carlo (see [4]) uses ideas from regression-based Monte Carlo American option pricing, (see Chapter 8 of [11]).

2.2 Marginal Matching

Let \( X = (X_1, \ldots, X_n) \) denote a random vector with distribution function \( F_X \). Let \( \omega_X \equiv (E[h_1(X_1)], \ldots, E[h_n(X_n)]) \) for some functions \( h_1, \ldots, h_n \). And let \( \theta_X \equiv g(\omega_X) \) for a function \( g \) that maps \( \omega_X \) from \( R^n \) to \( R \). Two simple examples of \( \theta_X \) are as follows,

\[
\sum_{i=1}^n E[h(X_i)] \quad \text{and} \quad \max\{E[h(X_1)], \ldots, E[h(X_n)]\},
\]

that is \( \theta_X \) is defined based on the marginal distribution of (functions of) \( X_1, \ldots, X_n \). Let \( Y = (Y_1, \ldots, Y_n) \) denote another random vector with distribution function \( F_Y \) such that,

\[
X \not\equiv^d Y, \quad X_i =^d Y_i \text{ for all } i = 1, \ldots, n,
\]

where \( =^d \) denotes “being equal in distribution”. Simply note that since the marginal distributions of \( X \) and \( Y \) match, \( \theta_X = \theta_Y \). Now, suppose that \( \theta_X \) is to be estimated with Monte Carlo simulation. Given (6), samples can be drawn from \( F_X \) or \( F_Y \). Let \( \hat{\theta}_{X,m} \) and \( \hat{\theta}_{Y,m} \) denote Monte Carlo estimators of \( \theta_X \) based on \( m \) simulation runs when samples are drawn from \( F_X \) and \( F_Y \), respectively. Obviously,

\[
\hat{\theta}_{X,m} \not\equiv^d \hat{\theta}_{Y,m}.
\]
and so between $\hat{\theta}_{X,m}$ and $\hat{\theta}_{Y,m}$, i.e., when deciding on whether to sample from $F_X$ or $F_Y$, the estimator with a lower mean square error (MSE) should be chosen.

**Example: Finite-Dimensional Distributions of Brownian Motion** Let $\{X_t : t \geq 0\}$ denote a Brownian motion with drift $\mu$ and volatility parameter $\sigma$. Consider the random vector $X = (X_1, \ldots, X_n) \equiv (X_{t_1}, \ldots, X_{t_n})$ on the time grid, $0 < t_1 < t_2 < \ldots < t_n$. That is, following the basic definition of a Brownian motion, $X$ is a multivariate normal random vector with $E[X_{t_i}] = \mu t_i$ and $\text{Var}(X_{t_i}) = \sigma^2 t_i$, and $\text{cov}(X_{t_i}, X_{t_j}) = t_i > 0$ for $t_i < t_j$. Now, let $Y = (Y_1, \ldots, Y_n)$ denote a multivariate normal random vector whose marginal distributions match that of $X$ but with $\text{cov}(Y_i, Y_j) = 0$, i.e., components of $Y$ are independent normals.

**Stochastic Models of the Risk Factors** Let $\{R_t : t \geq 0\}$, representing the dynamics of a risk factor, denote a stochastic process defined on a given filtered probability space, $(\Omega, F, (F_t)_{0 \leq t \leq \infty}, P)$. In this paper we assume that $\{R_t : t \geq 0\}$ is in the following class: a Gauss-Markov process (see Chapter 5 of [17] or Chapter 3 of [11]) specified by

$$dR_t = (g_t + h_t R_t)dt + \sigma_t dB_t,$$

with $g, h,$ and $\sigma$ all deterministic functions of time and $B$ a standard one-dimensional Brownian motion. Many of the widely used continuous time stochastic processes in finance and economics are in this class. Consider the finite dimensional distribution of $R$ on a time grid, $t_1, \ldots, t_n$ and set $R_i \equiv R_{t_i}$. Suppose that $R = (R_1, \ldots, R_n)$ can be sampled from *exactly* in the sense that the distribution of the simulated $R$ is precisely that of the $R$ process at times $t_1, \ldots, t_n$; examples are Brownian motion, Ornstein-Uhlenbeck processes, GBM, and the square-root diffusion specified above whose simulations involve generating positively correlated normal random variables. Let $\tilde{R} = (\tilde{R}_1, \ldots, \tilde{R}_n)$ denote a random vector for which $\tilde{R} \neq^d R$ but $\tilde{R}_i =^d R_i$ for all $i = 1, \ldots, n$ and $\text{cov}(\tilde{R}_i, \tilde{R}_j) = 0$ for all $i \neq j$. That is, simulation of $\tilde{R}_1, \ldots, \tilde{R}_n$ can be done by generating $n$ uncorrelated or simply independent normal random variables.

**PDS Sampling versus DJS Sampling** In the CCR literature when counterparty risk measures are estimated based on sampling from the finite-dimensional distributions of the underlying risk factors, the sampling is referred to as *Path Dependent Simulation* (PDS sampling). When the notion of marginal matching is used, the sampling is referred to as *Direct Jump to Simulation date* (DJS). For instance, in the Brownian motion example above, sampling from $X$ and $Y$ when estimating $\theta_X$-type estimands are referred to as PDS and DJS sampling, respectively. In Monte Carlo estimation of CCR measures, PDS and DJS sampling have been widely considered.

---

4This assumption is only used in the proof of Proposition 1 and Proposition 2.
equivalent (see [20]). We have also observed that practitioners often choose either of the sampling methods arbitrarily. One of the main contributions of this paper is to differentiate DJS and PDS in terms of the mean square error of the estimators of EPE, CVA, and eEPE.

3 Efficient Monte Carlo Estimation of EPE

In this section we consider efficient Monte Carlo estimation of EPE,

\[ \text{EPE} = \int_{0}^{T} E[V_t]dt, \]

where \( V \) denotes the credit exposure process, and \( T > 0 \) represents the expiration time of the longest maturity derivative contract in an OTC derivatives portfolio. Consider a time grid, \( 0 \equiv t_0 < t_1 < ... < t_n \equiv T \), with a fixed \( n \). Set \( \Delta_i \equiv t_i - t_{i-1} \) and \( V_i \equiv V_{t_i}, \ i = 1, ..., n \). Let \( \hat{\theta}_{b,m,n,k} \) denote a class of Monte Carlo estimators of EPE defined as follows,

\[ \hat{\theta}_{b,m,n,k} \equiv \sum_{i=1}^{n} \bar{V}_i \Delta_i, \]

where \( \bar{V}_i \equiv \frac{\sum_{j=1}^{m} V_{ij}}{m} \) and \( V_{i1}, ..., V_{im} \) represent the \( m \) simulation samples at valuation point \( t_i \). The subscript \( b \) refers to the biased nature of the estimators, and the subscript \( k \) could take \( p \) and \( d \), referring to PDS and DJS based simulation of the credit exposure process, respectively. As mentioned in Section 2.1, simulating the credit exposure process involves sampling from the underlying risk factors. Hereafter, PDS and DJS-based simulations of the credit exposure process refer to the cases where the underlying risk factors are sampled from based on their finite dimensional distributions (PDS sampling) and based on the notion of marginal matching (DJS sampling), respectively. Note that,

\[ \text{MSE}(\hat{\theta}_{b,m,n,k}) = \text{Var} \left( \sum_{i=1}^{n} \bar{V}_i \Delta_i \right) + \left( \sum_{i=1}^{n} E[\bar{V}_i] \Delta_i - \int_{0}^{T} E[V_t]dt \right)^2. \]

We assume that Monte Carlo realizations of \( V_i \) are unbiased estimates of \( E[V_i], i = 1, ..., n \). This implies that the bias part of the MSE of \( \hat{\theta}_{b,m,n,k} \) is not affected by the choice of the sampling method (PDS or DJS). In Section 3.1, we assume that \( n \), the number of valuation points, is fixed, and we compare the efficiency of \( \hat{\theta}_{b,m,n,p} \) and \( \hat{\theta}_{b,m,n,d} \) in terms of variance and computing time both for path independent and path dependent derivatives. Next, we introduce our efficient biased, yet consistent Monte Carlo estimators of EPE. In Section 3.3 we introduce efficient unbiased estimators of EPE. Numerical examples in Section 3.4 indicate that our proposed estimators substantively outperform the crude estimators of EPE in terms of the mean square error.
3.1 Comparing PDS and DJS-based Estimation of EPE

Suppose that the credit exposure process, $V_t$, is defined on a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq \infty}, P)$, where $(\mathcal{F}_t)_{0 \leq t \leq \infty}$ denote the filtration generated by the underlying risk factors. Consider the setting where $V$ denotes the contract level exposure and a financial institution takes a position in a maturity-$T$ derivative contract with its counterparty. Let $\Pi_T$ denote the payoff function of the derivative contract. It is well known from martingale pricing that

$$C_t = n_t E \left[ \frac{\Pi_T}{n_T} | \mathcal{F}_t \right],$$

(9)

where $n$ is a numeraire. Transactions between the financial institution and its counterparty for which $V_t = \max \{ C_t, 0 \} = C_t$ for all $0 < t \leq T$ are referred to as unilateral transactions, e.g. the financial institution takes a long position in a call option with its counterparty. Transactions for which $V_t = \max \{ C_t, 0 \} \neq C_t$ for some $0 < t \leq T$ are referred to as bilateral transactions, e.g. an interest rate swap between the financial institution and its counterparty.

The following simple example reviews simulation of the exposure process under PDS and DJS.

Suppose that the credit exposure process, $V$, is a Brownian motion with drift $\mu$ and volatility $\sigma$. Consider the setting where $V$ denotes the contract level exposure and a financial institution takes a position in a maturity-$T$ GBM-driven vanilla call option. Let $\Pi_T = \max \{ V_T - K, 0 \}$, whose components are uncorrelated but marginal distributions match those of $V$. This is the so-called PDS sampling method. An alternative sampling method, using the notion of marginal matching, is to sample from the multivariate normal random vector, $X = (X_1, ..., X_n)$. This is the so-called DJS method. To be more specific, in DJS sampling, $S_i$ is generated from time zero. That is, generate $Y_i$, a normal random variable with mean $\mu t_i$ and variance $\sigma^2 t_i$, and set $S_i = S_0 e^{Y_i}$. In PDS sampling, $V_i$’s are sampled based on generating the sample path of the

\[ \theta \equiv \sum_{i=1}^{n} E[V_i] \Delta_i. \]

Recall that,

\[ \hat{\theta}_{b,m,n,k} = \sum_{i=1}^{n} \bar{V}_i \Delta_i, \]

where $\bar{V}_i$ is the $m$-simulation-run average of $V_{i1}, ..., V_{im}$. With $V_i = f(S_i)$ and $S_i = S_0 e^{Y_i}$, Monte Carlo estimation of $\theta$ requires sampling from the multivariate normal random vector, $X = (X_1, ..., X_n)$. Assuming zero short rate, $C_t = E[(S_T - K)^+ | S_t] = E[(S_t S_{T-t} - K)^+ | S_t]$. Note that the function $f$ in $f(S_t) \equiv E[(S_t S_{T-t} - K)^+ | S_t]$, which is well-defined for all values of $t \geq 0$ given the payoff function $\Pi_T$ with a fix maturity $T$, is in fact a function of $t$ and $S_t$. In Section 3, for notational simplicity, we suppress the dependence of $f$ on $t$ in the definition $C_t = n_t E[\Pi_T/n_T | S_t] \equiv f(S_t)$. 

\[8\]
GBM sequentially at \( i = 1, \ldots, n \). That is, to generate a realization of \( V_i \), \( S_i \) is generated given the previously generated value of \( S_{i-1} \). \(^6\) Note that since for any given \( t > 0 \), \( V_i \) is a function of \( S_i = S_0 e^{X_t} \), DJS-based simulation of the exposure process implies that \( \text{cov}(V_i, V_j) = 0 \) for any \( i \neq j \), \( i, j = 1, \ldots, n \).

In what follows we compare the efficiency of \( \hat{\theta}_{b,m,n,p} \) and \( \hat{\theta}_{b,m,n,d} \) in terms of variance and computing time for path independent and path dependent derivatives. We consider unilateral and bilateral transactions in both single risk-factor and multi-risk factor settings. That is, we consider two cases: a stylized setting where \( (\mathcal{F}_t)_{0 \leq t \leq \infty} \) is the filtration generated by a single risk factor; we also consider the more general multi-risk factor settings.

### 3.1.1 Path Independent Case

The above mentioned example shows that under DJS, \( \text{cov}(V_u, V_t) = 0 \) for any \( 0 < u < t < T \). Proposition 1 and Proposition 2 consider this covariance function of the contract level credit exposure process under the PDS method for unilateral and bilateral transactions, respectively, and identify conditions under which \( \text{cov}(V_u, V_t) > 0 \) for any \( 0 < u < t < T \). Condition 2 of Proposition 1 below uses the well known changes of numeraire techniques of Geman-El Karoui-Rochet \([9]\) for option type contracts with at most three distinct sources of randomness: stochastic short rate and a maximum of two risky assets. Well known examples of these contracts are options written on stocks or bonds, e.g. European options and exchange options.

**Proposition 1.** Consider the credit exposure process, \( \{V_t; t \geq 0\} \), defined on a given filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq \infty}, P)\), and a \( T \)-maturity transaction between the financial institution and its counterparty that is unilateral, i.e. the credit exposure process is the price process, \( V_t = C_t > 0 \) for all \( 0 \leq t \leq T \), where \( C_t \) denotes the time-\( t \) value of the derivative contract with payoff \( \Pi_T \). Then,

\[
\text{cov}(V_u, V_t) > 0,
\]

for any \( 0 < u < t < T \) under any of the following conditions:

**Condition 1:** Numeraire is the money market account, \( B \), with deterministic short rate, \( r \), and \( \Pi_T \) is a function of \( N \geq 1 \) exogenously given risky assets.

**Condition 2:** Short rate is stochastic and the \( T \)-payoff function is a function of at most two risky assets as follows \( \Pi_T = (\alpha_1 S_1(T) + \alpha_2 S_2(T))^+ \), where \( \alpha_1 \) and \( \alpha_2 \) are any real numbers, and \( S_1 \) and/or \( S_2 \) are risky assets.

\(^6\)More specifically, to sample from \( S_i \) generate \( \tilde{X}_i \) and set \( S_i = S_{i-1} e^{\tilde{X}_i} \), where \( \tilde{X}_i \) is a normal random variable with mean \( \mu \Delta_i \) and variance \( \sigma^2 \Delta_i \).
Proof. We first show that $\text{cov}(V_u, V_t) > 0$ holds under condition 1. That is, $V_t = C_t = B_t E[B_T^{-1} \Pi_T \mid \mathcal{F}_t]$. For any $0 < u < t$ we have

$$
\text{cov}(V_u, V_t) = E \left[ \text{cov}(V_u, V_t \mid \mathcal{F}_u) \right] + \text{cov}(V_u, E[V_t \mid \mathcal{F}_u]),
$$

where the last equality follows from the conditional covariance formula (see Chapter 3 of [21]). It is easy to check that the first term on the right hand side above is zero. Consider the second term and note that

$$
E[V_t \mid \mathcal{F}_u] = E \left[ B_t E[\Pi_T \mid \mathcal{F}_u] \right] = E \left[ B_t \Pi_T \mid \mathcal{F}_u \right],
$$

and so we conclude that for any $0 < u < t$,

$$
\text{cov}(V_u, V_t) = B_u B_t B_T^{-2} \text{Var}(E[\Pi_T \mid \mathcal{F}_u]) > 0.
$$

The second part of the proof, which is based on condition 2, uses standard results on changes of numeraire techniques and Chebyshev’s algebraic inequality (see, for instance, Proposition 2.1 in [8]).

Let $C_t$ denote the time-$t$ value of a derivative contract specified in Assumption 2. Recall Theorem 1 and Theorem 2 of [9], and note that

$$
V_t = C_t = S_1(t)E_{Q^1}[(\alpha_1 + \alpha_2 Z(T))^+ \mid \mathcal{F}_t]
$$

where $Z = S_2/S_1$ and the subscript $Q^1$ refers to expectation under $Q^{S_1}$, i.e. $S_1$ is the numeraire. Note that

$$
\text{cov}(V_u, V_t) = \text{cov}(V_u, E[V_t \mid \mathcal{F}_u])
= \text{cov}(S_1(u)E_{Q^1}[\alpha_1 + \alpha_2 Z(T)]^+ \mid \mathcal{F}_u , E_{Q^1}[S_1(t)(\alpha_1 + \alpha_2 Z(T))^+ \mid \mathcal{F}_u]),
$$

Consider the second term on the right hand side above, and suppose that the transition law of $S_1$ accepts the following specification

$$
S_1(t) = d \beta S_1(u) S_1(\tau), \quad (10)
$$

where $\beta > 0$ is a constant and $t - u \equiv \tau$ (see the last part of the proof on transition law of the numeraire). We, then, have

$$
E_{Q^1}[S_1(t)(\alpha_1 + \alpha_2 Z(T))^+ \mid \mathcal{F}_u] = \beta S_1(u) E_{Q^1}[S_1(\tau)(\alpha_1 + \alpha_2 Z(T))^+ \mid \mathcal{F}_u].
$$

This gives

$$
\text{cov}(V_u, V_t) = \beta S_1^2(u) \text{cov}(E_{Q^T}[\alpha_1 + \alpha_2 Z(T)]^+ \mid \mathcal{F}_u , E_{Q^T}[S_1(\tau)(\alpha_1 + \alpha_2 Z(T))^+ \mid \mathcal{F}_u]).
$$

Note that both expectations above are monotone functions of $Z(u)$ and so using Chebyshev’s algebraic inequality (see, for instance, Proposition 2.1 in [8]) gives $\text{cov}(V_u, V_t) > 0$. 


We now show that (10) holds for the class of stochastic processes considered in this paper for modeling the dynamics of risk factors. The dynamics of risky asset $S_1 > 0$ selected as the numeraire in the proof above is assumed to be modeled by a GBM or a square-root diffusion. In cases where the numeraire is a maturity-$T$ zero-coupon bond, the second part of the proof above is modified as follows. We assume that the zero-coupon bond is modeled such that it possesses an affine term structure, i.e. it has the form $S_1(t) \equiv p(t, T) = e^{A-Bt}$, where $A \equiv A(t, T)$ and $B \equiv B(t, T)$ are deterministic functions and the short rate $r$ is modeled by a Gauss-Markov process or a square-root diffusion as specified before. It is, then, not difficult to see that $p(t, T) = e^{A-Bt} = e^{\beta_1 S_u + \beta_2 (t-u)}$, where $\beta_1$ and $\beta_2$ are constants and $0 < u < t < T$, $t - u \equiv \tau$. Using a similar approach shown in the second part of the proof above, we arrive at $\text{cov}(V_u, V_t) > 0$.

In the case of bilateral transactions for which the exposure process satisfies $V_t = \max\{C_t, 0\} \neq C_t$ for some $0 < t \leq T$, where $C_t$ denotes the time-$t$ value of the derivative contract with payoff function $\Pi_T$. Then, $\text{cov}(V_u, V_t) > 0$.

**Proposition 2.** Consider the credit exposure process, \{V_t; t \geq 0\}, defined on a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq \infty}, P)$, and a $T$-maturity transaction between the financial institution and its counterparty that is bilateral, i.e. the credit exposure process is the price process, $V_t = \max\{C_t, 0\} \neq C_t$ for some $0 < t \leq T$, where $C_t$ denotes the time-$t$ value of the derivative contract with payoff function $\Pi_T$. Then,

$$\text{cov}(V_u, V_t) > 0$$

for any $0 < u < t < T$ under the following condition:

**Numeraire** is the money market account, $B$, with deterministic short rate, $r$, and $\Pi_T$ is a monotone function of a single risky asset whose dynamics is modeled by a GBM, a Gauss-Markov process or a square-root diffusion.

**Proof** Conditioning on $\mathcal{F}^S_u \equiv \mathcal{F}_u$ and using conditional covariance formula gives

$$\text{cov}(V_u, V_t) = \text{cov}(V_u, E[V_t|\mathcal{F}_u]) = \text{cov}(\max\{f(S_u), 0\}, E[\max\{f(S_t), 0\}|\mathcal{F}_u]),$$

for a well-defined function $f$. First consider the first term $\max\{f(S_u), 0\}$ inside the covariance function on the right hand side above. Note that since $f$ is a monotone function, $\max\{f(S_u), 0\} \equiv \tilde{f}(S_u)$ is also a monotone function of $S_u$. Next, consider the second term $E[\max\{f(S_t), 0\}|\mathcal{F}_u]$.

Note that when $S$ is a Gauss-Markov process, the transition law of $S$ implies that for any $0 < u < t$, $S_t = \beta_1 S_u + \beta_2 S_{t-u}$, where $S_u$ and $S_{t-u}$ are independent random variables. Also,

\footnote{For instance, consider the case where $C_t$ represents the time-$t$ value of an interest rate swap. Then, $C_t$ can be negative for some $t > 0$.}
when $S$ is a GBM, for any $0 < u < t$ we have $\log(S_t) = d \log(S_u) + \log(S_{t-u})$, where $S_u$ and $S_{t-u}$ are independent random variables. This follows from the independent and stationary increments properties of Gauss-Markov processes and that their finite dimensional distributions are multivariate normal. When $S$ is a square-root diffusion, $dS_t = \alpha(b - S_t)dt + \sigma\sqrt{S_t}dB_t$ with $B$ a standard one-dimensional Brownian motion and positive constants $\alpha$ and $b$, it can be shown that $S_t$ given $S_u$ is distributed as a positive constant times times a noncentral chi-square random variable with degrees of freedom that depends on $\alpha$, $\sigma$, and $b$, and noncentrality parameter which is an increasing function of $S_u$, (see Chapter 3 of [11]).

This implies that under the class of risk factor models considered in the paper, $E[\max\{f(S_t), 0\}|\mathcal{F}_u]$ is a monotone function of $S_u$. To see this, consider the case where $f$ is an increasing function. Increasing $S_u$ will increase $S_t$; this increases $\max\{f(S_t), 0\}$. So, $E[\max\{f(S_t), 0\}|\mathcal{F}_u] \equiv \tilde{h}(S_u)$ also becomes an increasing function of $S_u$. A similar argument can be used when $f$ is a decreasing function. Consequently, we can write $\text{cov}(V_u, V_i) = \text{cov}(\tilde{f}(S_u), \tilde{h}(S_u))$, where $\tilde{f}$ and $\tilde{h}$ are both either increasing or decreasing functions of $S_u$. Using Chebyshev’s algebraic inequality gives $\text{cov}(V_u, V_i) = \text{cov}(\tilde{f}(S_u), \tilde{h}(S_u)) > 0$.

The monotonicity assumption of the payoff function is satisfied for most of the actively traded OTC derivative contracts; well-known exceptions are Barrier options, (see, for instance, [18]).

Propositions 1 and 2 identify conditions for unilateral and bilateral transactions under which the credit exposure process satisfies $\text{cov}(V_u, V_i) > 0$ for any $0 < u < t < T$. This, then, implies that

$$\text{Var}(\hat{\theta}_{b,m,n,d}) \leq \text{Var}(\hat{\theta}_{b,m,n,p}).$$

(11)

Note that the above inequality holds since

$$\text{Var}(\hat{\theta}_{b,m,n,d}) = \frac{\sum_{i=1}^{n} \text{Var}(V_i) \Delta_i^2}{m} \leq \frac{\sum_{i=1}^{n} \text{Var}(V_i) \Delta_i^2}{m} + \frac{2}{m} \sum_{i<j} \text{cov}(V_i, V_j) \Delta_i \Delta_j = \text{Var}(\hat{\theta}_{b,m,n,p}).$$

(12)

3.1.2 Path Dependent Case

Suppose that $V_t$ is time $t$ value of a maturity-$T$ contract, where the payoff at the time $T$ is a function of $S_1, ..., S_n$, (for instance, an arithmetic Asian option). That is, $V_i = g(S_1, ..., S_i)$, where $g$ is a function from $\mathcal{R}^i$ to $\mathcal{R}$. The DJS sampling method is to make $V_i = g(S_1, ..., S_i)$ and $V_j = g(S_1, ..., S_j)$, $i < j$, uncorrelated random variables. That is, sample from $S_1, ..., S_i$ to generate a single realization of $V_i$. To generate $V_j$, start again from time zero, and sample

---

8 Note that when $S$ is a GBM, the logarithm of $S$ is a Brownian motion, which is a Gauss-Markov process.

9 More specifically, the payoff function of up-and-in and down-and-out European barrier call options are monotone functions of the underlying security prices. This monotonicity assumption does not hold for up-and-out and down-and-in European barrier call options, (see Chapter 6 of [18] and the references there).
from $S_1, \ldots, S_i, \ldots S_j$. Under this DJS-type sampling method, $V_i$ and $V_j$ become uncorrelated, $\text{cov}(V_i, V_j) = 0$. In the PDS-type sampling, given the Monte Carlo realization of $V_i$, to generate $V_j$, one uses the previously generated $S_1, \ldots, S_i$ and only samples from $S_{i+1}, \ldots, S_j$. In this case $V_i$ and $V_j$ are dependent.

Using conditional covariance formula and arguments similar to the ones used in the path independent case, it can be shown that $\text{cov}(V_i, V_j) > 0$, $i \neq j$. More specifically, it can be shown that $\text{cov}(V_i, V_j) > 0$ for unilateral and bilateral transactions under the first condition of Proposition 1 and Proposition 2’s condition, respectively. That is, for the above mentioned covariance function to be positive, we need the numeraire money market account with deterministic short rate in the unilateral case. The bilateral case, additionally, requires monotonicity of the payoff function and its dependence on a single risk factor.

To compare the efficiency of the DJS and PDS-based estimators of $\theta$ in the path dependent case, computing time is also to be considered in parallel with variance of the estimators. More specifically, the estimator with the lower

\[
\text{variance per replication} \times \text{expected computing time},
\]

should be selected (see [13] for the formal formulation of this useful criterion in comparing alternative Monte Carlo estimators). Consider, for instance, arithmetic Asian options. Suppose that the computational time to calculate $\hat{\theta}_{b,m,n,k}$ is proportional to the number of random variables that are to be generated. Let $\text{ct}(\hat{\theta}_{b,m,n,k})$ denote the computational effort associated with $\hat{\theta}_{b,m,n,k}$. Note that,

\[
\frac{\text{ct}(\hat{\theta}_{b,1,n,d})}{\text{ct}(\hat{\theta}_{b,1,n,p})} \approx n \quad \text{and} \quad \frac{\text{Var}(\hat{\theta}_{b,1,n,p})}{\text{Var}(\hat{\theta}_{b,1,n,d})} \approx n. \tag{13}
\]

To see why (13) holds note that to calculate $\hat{\theta}_{b,1,n,d}$, $\frac{n(n+1)}{2}$ random variables are to be generated while $\hat{\theta}_{b,1,n,p}$ requires generating $n$ random variables, (assuming that the calculation of $E[\Pi^A | F_i]$ does not require generating additional random variables). Also, note that as can be seen from (12), variance of the PDS-based estimator is of order $n^2$ because of the covariance terms while the DJS-based estimator has a variance of order $n$. So, $\hat{\theta}_{b,m,n,d}$ and $\hat{\theta}_{b,m,n,p}$ have a similar performance for fixed and sufficiently large $n$. PDS and DJS-based estimators of other derivatives whose payoff depends on the path in a different form can be compared similarly.

3.1.3 Summary of Section 3.1

We summarize the result of Section 3.1 as it is used in the sequel and is directly applied to the efficient CVA$_I$ estimation. To compare the DJS and PDS-based estimators of EPE and CVA$_I$
(both being viewed as weighted sums of expected exposures) variance and computing time of the Monte Carlo estimators are considered. The DJS method induces zero covariance between any two distinct time points of the simulated credit exposure process. So, it remains to look at this covariance function for the credit exposure process under the PDS method. When the dynamics of the risk factors are modeled by the class of continuous time stochastic processes considered in this paper, the covariance function of the credit exposure process under the PDS method becomes positive under conditions of Proposition 1 and 2 for unilateral and bilateral path independent derivatives transactions, respectively. Similar results hold for path dependent derivatives. That is, under conditions of Proposition 1 and 2, DJS-based estimators of EPE and CVA outperform the PDS-based estimators in terms of variance. For path independent derivatives PDS and DJS-based computing times are roughly equal. So, we recommend that the counterparty credit risk modeler uses DJS for path independent derivatives. For path dependent derivatives, DJS-based estimators usually have larger computing times. The criterion introduced above considers the computing time in parallel with variance. There are widely traded path dependent derivatives for which PDS and DJS-based estimators of EPE and CVA perform approximately equally. For instance, for arithmetic Asian options the DJS and PDS-based estimators of EPE and CVA perform similarly.

There are contracts whose payoff function does not exactly match the mathematical conditions of Proposition 1 and 2. For those contracts, a small simulation study could compare the variance of the DJS and PDS-based estimators of EPE and CVA. The Appendix contains numerical examples for EPE estimation of a single interest rate swap, where we conclude that DJS outperforms PDS by at least an order of 10 in terms of variance while the computing times of both Monte Carlo estimators are roughly equal.

Hereafter, we assume that the credit exposure process $V$ satisfies $\text{cov}(V_u, V_s) = 0$ and $\text{cov}(V_u, V_s) > 0$ when simulated under the DJS and PDS methods, respectively, for any $0 < u < t$.

### 3.2 Efficient Monte Carlo EPE Estimation: Biased Estimators

In this subsection, we suppress the subscript $b$ in $\hat{\theta}_{b,m,n,k}$ and instead write $\hat{\theta}_{m,n,k}$ for notational simplicity. We would like to find the number of valuation points, $n$, and the number of simulation runs at each valuation point, $m$, to minimize $\text{MSE}(\hat{\theta}_{m,n,k})$,

$$\text{MSE}(\hat{\theta}_{m,n,k}) = \text{Var}(\hat{\theta}_{m,n,k}) + (E[\hat{\theta}_{m,n,k}] - \text{EPE})^2.$$ 

given a fixed computational budget, denoted by $s$, that is proportional to, $mn$. Also, $k = p$, and $d$ refer to PDS and DJS-based simulation of the credit exposure process on a time grid $0 \equiv t_0 < t_1 < ... < t_n \equiv T$. That is, as shown in the previous section, under PDS sampling and DJS sampling, $\text{cov}(V_i, V_j) > 0$ and $\text{cov}(V_i, V_j) = 0$, respectively, for any $i \neq j$, $i, j = 1, ..., n$.

To formulate and solve this optimization problem, we specify the order of the variance and bias of the Monte Carlo estimator of EPE, $\theta_{m,n,k}$. Note that from basic results on endpoint Reimann sum approximation of integrals, time-discretization bias is of order $1/n$. We are not concerned with deriving sharp estimates of the orders of variance. In fact, our numerical examples indicate that choosing approximately optimal $m$ and $n$ using even very rough approximates
for the orders of variance and bias leads to substantial MSE reduction compared to industry practice.

Suppose that the time grid is equidistant, i.e., $\Delta_i \equiv \Delta = \frac{T}{n}$. We assume that $E[V_t^2] < \infty$ for all $t \in [0, T]$. First, we note that

$$\text{Var}(\hat{\theta}_{m,n,d}) = O\left(\frac{1}{mn}\right). \tag{14}$$

To see this,\(^{11}\) consider $M > 0$ such that $E[V_t^2] \leq M$ for $t \in (0, T]$. Note that,

$$\text{Var}(\hat{\theta}_{m,n,d}) = \Delta^2 \sum_{i=1}^{n} \frac{\text{Var}(V_i)}{m} \leq \left(\frac{T}{n}\right)^2 \sum_{i=1}^{n} \frac{E(V_i^2)}{m} \leq MT^2 \frac{1}{mn}.$$

Now, consider the variance of the PDS-based estimator, $\hat{\theta}_{m,n,p}$,

$$\text{Var}(\hat{\theta}_{m,n,p}) = \Delta^2 \sum_{i=1}^{n} \frac{\text{Var}(V_i)}{m} + \Delta^2 \frac{2}{m} \sum_{i<j} \text{cov}(V_i, V_j).$$

As shown before, the first term above is $O\left(\frac{1}{mn}\right)$. Also, under PDS sampling, the credit exposure process is simulated according to its finite dimensional distributions for which the covariance terms are positive. So, the second term is $O\left(\frac{1}{m}\right)$. This gives,

$$\text{Var}(\hat{\theta}_{m,n,p}) = O\left(\frac{1}{mn} + \frac{1}{m}\right). \tag{15}$$

**PDS-Based Biased Efficient Estimator of EPE** We choose the number of valuation points, $n$, and number of simulation runs at each valuation point, $m$, to minimize the mean square error of the PDS-based estimator, $\hat{\theta}_{m,n,p}$, under a fixed computational budget proportional to $mn$. Approximating the variance of $\hat{\theta}_{m,n,p}$ using (15) leads to the following optimization problems,

$$\min_{m,n} \left(\frac{c_{p,1}}{mn} + \frac{c_{p,2}}{m} + \frac{c_2}{n^2}\right) \quad \text{subject to} \quad s = c_3 mn, \tag{16}$$

for some constants, $c_{p,1}, c_{p,2}, c_2, \text{ and } c_3$. MSE of $\hat{\theta}_{m,n,p}$ is minimized at,

$$m = cs^2 \quad \text{and} \quad n = \tilde{c}s^3, \tag{17}$$

for constants $c$ and $\tilde{c}$.

\(^{11}\)The Landau symbol, $O$, in $f(x, y) = O(g(x, y))$ means that $f(x, y)/g(x, y)$ stays bounded in some limit, say $x, y \to 0$ or $x, y \to \infty$. 

15
DJS-Based Biased Efficient Estimator of EPE  

Let $c_d$ denote a constant. Given (14), we approximate $\text{Var}(\hat{\theta}_{m,n,d})$ with $\frac{c_d}{mn}$ in the MSE minimization problem for the DJS-based estimator,

$$\min_{m,n} \left( \frac{c_d}{mn} + \frac{c_2}{n^2} \right) \quad \text{subject to} \quad s = c_3mn,$$

to which the trivial optimal solution is $m = 1$ and $n = \hat{c}s$ for some constant $\hat{c}$. We note that estimating the various constant parameters appearing in all the above mentioned MSE minimization problems is not possible in practice. In our numerical examples we simply set all these constant parameters equal to 1.

Remark  

We do not claim originality in setting up an MSE minimization problem to derive an optimum balance between variance and bias squared; this can be seen in Chapter 6 of Glasserman [11] and the references there, particularly the paper by Duffie and Glynn [6]. Our contribution is that in our proposed efficient Monte Carlo CCR framework, choosing approximately optimal $m$ and $n$ via solving MSE minimization problems achieve substantial MSE reduction in Monte Carlo estimation of EPE, CVA, (and as will be seen in Section 5), and eEPE. This has neither appeared in the CCR literature nor been applied by practitioners. Moreover, our result that the efficient DJS-based estimator requires all its computational budget allocated to the number of valuation points is surprising. Consider a well defined continuous time stochastic process $S$ whose finite dimensional distributions satisfy $\text{cov}(S_u, S_t) > 0$ for any $0 < u < t$. Suppose that Monte Carlo is to be used to estimate $\theta = E[\int_0^T S_t dt]$ for a given $T > 0$. The purposed efficient Monte Carlo estimator of $\theta$ employs the notion of marginal matching (as opposed to simulating the process based on the finite dimensional distribution of $S$) and uses 1 simulation run at each time point in a discrete time grid given a fixed computational budget all allocated to making the grid as fine as possible.

3.3 Efficient Monte Carlo EPE Estimation: Unbiased Estimators

In this section we derive unbiased estimators of EPE. Specifically, we eliminate the time discretization bias at the expense of introducing additional randomness. To control the variance that would be increased as the result of this new source of randomness, we use stratified sampling. Let $\tau$ denote a $[0, T]$ Uniform random variable that is independent of the credit exposure, $V$. We have,

$$\text{EPE} = TE[V_{\tau}], \quad (18)$$

which simply follows from conditioning on $\tau$, i.e., using $E[V_{\tau}] = E[E[V_{\tau}|\tau]]$, independence of $V$ and $\tau$, and noting that $f(t) = \frac{1}{T}$, $t \in [0, T]$, is the probability density function of $\tau$. Now, consider the following identity,

$$\text{EPE} = TE[V_{\tau}] = T \sum_{i=1}^{n} E[V_{\tau}|\tau \in A_i]p_i = \sum_{i=1}^{n} E[V_{\tau}|\tau \in A_i] \Delta_i, \quad (19)$$
where $A_i = [0, t_i)$, $p_i = P(\tau \in A_i) = \Delta_i$, on the time grid, $0 \equiv t_0 < t_1 < \ldots < t_n \equiv T$, and $\Delta_i = t_i - t_{i-1}$. Assuming $t_i = iT/n$ for all $i = 1, \ldots, n$, our proposed unbiased estimators of EPE use the identity (19) by estimating the conditional expectations, $E[V_{\tau} | \tau \in A_i]$, i.e.,

$$\hat{\theta}_{u,m,k} = \sum_{i=1}^{n} \bar{V}_{\tau_i} \Delta_i,$$

where $\tau_i \equiv \tau | \tau \in A_i$, $\bar{V}_{\tau_i} = \sum_{j=1}^{m} V_{\tau_{ij}} / m$, and $\tau_{i1}, \ldots, \tau_{im}$ are i.i.d. copies of $\tau_i$. That is, to draw a single realization of $V_{\tau_i}$, we first sample from $\tau_i$ conditional on $\tau \in A_i$. Note that $\tau_i$ is a $[t_{i-1}, t_i]$ Uniform random variable. Next, given this realization of $\tau_i$, we generate $V_{\tau_i}$. The subscript $k = p$ and $d$ refer to PDS and DJS sampling, respectively. That is, PDS-based simulation in calculating $\hat{\theta}_{u,m,p}$ implies that $\text{cov}(V_{\tau_i}, V_{\tau_j}) \neq 0$ for $i \neq j$, $i, j = 1, \ldots, n$, and DJS-based simulation in calculating $\hat{\theta}_{u,m,d}$ implies that $\text{cov}(V_{\tau_i}, V_{\tau_j}) = 0$ for $i \neq j$. This immediately implies $\text{Var}(\hat{\theta}_{u,m,d}) \leq \text{Var}(\hat{\theta}_{u,m,p})$. Consider a more general setting that allows different numbers of simulation runs for each stratum. That is, let $m_i$ denote the number of runs used to estimate $E[V_{\tau} | \tau \in A_i]$ and $N = m_1 + \ldots + m_n$ denote the total number simulation runs. Note that our setting with equidistant strata and $m_i \equiv m$, for $i = 1, \ldots, n$ coincides with proportional stratified sampling which uses $m_i = np_i$, (see [22] for results on proportional stratification). This is because $\tau$ is a $[0, T]$ Uniform random variable. In this paper we do not address further possible improvements of our unbiased stratified sampling-based estimators of EPE by attempting to find optimal $m_1, \ldots, m_n$ and $n$ under fixed computational budgets. Our numerical examples indicate that using our unbiased stratified sampling-based estimators by setting $m_i \equiv m$ and choosing $m$ and $n$ as specified in subsection 3.2 leads to substantial MSE reduction when compared to crude biased Monte Carlo estimators of EPE.

Comparing DJS-based Biased and Unbiased estimators Proposition 1 below shows that $\hat{\theta}_{u,m,k}$ and the biased DJS-based estimator of EPE, $\hat{\theta}_{b,m,k}$, are asymptotically equivalent in terms of MSE. This equivalence is further confirmed by our numerical experiments (see the next subsection) in practical settings with fixed and finite computational budgets proportional to $mn$.

Proposition 3. Consider the credit exposure process, $\{V_t : t \geq 0\}$, defined on a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq \infty}, \mathbb{P})$. Suppose that biased and unbiased Monte Carlo estimators of EPE calculated under DJS-sampling,

$$\hat{\theta}_{b,m,k} = \sum_{i=1}^{n} \bar{V}_{\tau_i} \Delta_i, \quad \text{and} \quad \hat{\theta}_{u,m,k} = \sum_{i=1}^{n} \bar{V}_{\tau_i} \Delta_i.$$
are defined on an equi-distant time grid, \( 0 < t_0 < t_1 < ... < t_n \equiv T \), where \( \Delta_i \equiv t_i - t_{i-1} = T/n \equiv \Delta \), \( \tau_i \equiv \tau \mid \tau \in A_i \) and \( A_i = [t_{i-1}, t_i) \). Let \( \bar{V}_i \) and \( \bar{V}_{\tau_i} \) denote the averages of \( m \) Monte Carlo realizations of \( V_i \), and \( V_{\tau_i} \), respectively. That is, the total number of simulation runs is \( N = mn \). We assume that \( E[V_i^2] < \infty \), for all \( i = 1, ..., n \). Asymptotic performance of \( \hat{\theta}_{b,m,n,d} \) and \( \hat{\theta}_{u,m,n,d} \) is equivalent in the following sense,

\[
\lim_{n \to \infty} n \text{MSE}(\hat{\theta}_{b,m,n,d}) = n \text{Var}(\hat{\theta}_{u,m,n,d}) = c \int_0^T \text{Var}(V_t)dt, \tag{22}
\]

where \( c \) is a constant.

**Comparing PDS-based Biased and Unbiased estimators** Analytically comparing the \( \text{MSE}(\hat{\theta}_{b,m,n,p}) \) and \( \text{Var}(\hat{\theta}_{u,m,n,p}) \) is quite difficult due to the presence of the covariance terms. Our numerical examples presented in the next subsection show that the unbiased PDS-based estimator of EPE, \( \hat{\theta}_{u,m,n,p} \), outperforms the efficient biased PDS-estimator, \( \hat{\theta}_{b,m,n,p} \), introduced in the previous section.

### 3.4 Numerical Examples

In this section we use simple numerical examples to illustrate the efficiency of our proposed Monte Carlo estimators of EPE. We consider contract level exposure in a simple setting where \( V_t \equiv S_t \) denotes the value of a geometric Brownian motion at time \( t > 0 \). That is, \( S_t = S_0e^{X_t} \) with \( \{X_t; t \geq 0\} \) being a Brownian motion with drift \( \mu \), and volatility \( \sigma \). This stylized example enables us to calculate the MSE exactly. We consider six different Monte Carlo estimators of EPE in our numerical examples.

Let \( \hat{\theta}_{c,p} \) and \( \hat{\theta}_{c,d} \) denote the “crude” and biased Monte Carlo estimators of EPE under PDS and DJS sampling, respectively. That is,

\[
\hat{\theta}_{c,k} = \sum_{i=1}^{n} \bar{V}_i \Delta_i, \tag{23}
\]

where \( \Delta_i = t_i - t_{i-1}, 0 \equiv t_0 < t_1 < ... < t_n \equiv T \), \( k = p, d \), and \( \bar{V}_i \) is the \( m \)-simulation-run average of \( V_i \). We shall shortly specify the choice of the valuation points.

Let \( \hat{\theta}_{e,b,p} \) and \( \hat{\theta}_{e,b,d} \) denote the efficient and biased Monte Carlo estimators of EPE under PDS and DJS sampling, respectively. In particular, their statistical efficiency is a result of solving the MSE minimization problems in Section 3.2 to derive the (approximately) optimal number of points on the time grid, \( n \), and simulation runs at each of these time points, \( m \), given a fixed computational budget proportional to \( mn \).

Let \( \hat{\theta}_{e,p} \) and \( \hat{\theta}_{e,d} \) denote the unbiased stratified sampling-based Monte Carlo estimators of EPE under PDS and DJS sampling, respectively. That is,

\[
\hat{\theta}_{e,k} = \sum_{i=1}^{n} \bar{V}_{\tau_i} \Delta_i, \tag{24}
\]
where \( V_{\tau_i} = \sum_{j=1}^{m_i} V_{\tau_{ij}} / m_i \) with \( \tau_i \equiv \tau | \tau \in A_i, A_i = [t_{i-1}, t_i], \) and \( k = p, d. \)

We set \( T = 1. \) The crude estimators of EPE are calculated based on 12 valuation points, \( n = 12, \) at \( 1, 2, 3, 4, 8, 12, 18, 21, 24, 36, 49 \) weeks and 1 year. We note that one year, \( T = 1, \) with the number of valuation points fixed at 12, is a setting widely used by financial institutions. There is no mathematical basis for this arrangement of valuation points. It is believed that since some trades have “short” expiration times, having more valuation points earlier would increase the accuracy of the estimators of CCR measures. The time grid used to calculate our efficient estimators of EPE is equidistant, i.e., \( \Delta_i \equiv \Delta = T/n. \) Computational budget, \( s, \) is fixed at \( 12,000 \) and \( 120,000, \) respectively. To calculate \( \hat{\theta}_{e,b,p} \) under these fixed computational budgets, the solution, (17) with both \( c \) and \( \tilde{c} \) set to 1, to the MSE minimization problem of Section 3.2 is used.

This gives, \( n = 23 \) and \( m = 524 \) for \( s = 12,000, \) and \( n = 50, \) and \( m = 2433 \) for \( s = 120,000. \) Similarly, to calculate \( \hat{\theta}_{e,b,d}, \) we use the solution to the MSE minimization problem, (3.2). That is, we set \( n = 12,000, \) and \( m = 1 \) for \( s = 12,000, \) and \( n = 120,000, \) and \( m = 1 \) for \( s = 120,000. \) In calculating the stratified sampling estimators of EPE, \( \hat{\theta}_{a,p} \) and \( \hat{\theta}_{a,d}, \) we do not address the problem of deriving the optimal values of \( n, \) and \( m_1, ..., m_n. \) Instead, we simply use the setting of \( \hat{\theta}_{e,b,p} \) and \( \hat{\theta}_{e,b,d}, \) respectively. That is, to calculate \( \hat{\theta}_{a,p}, \) we set \( n = 23, \) \( m = 524, \) and \( n = 50, \) \( m = 2433, \) under \( s_1 = 12,000 \) and \( s_2 = 120,000, \) respectively. And to calculate \( \hat{\theta}_{a,d}, \) we set \( n = 12,000, \) \( m = 1, \) and \( n = 120,000, \) \( m = 1, \) under \( s_1 = 12,000 \) and \( s_2 = 120,000, \) respectively.

Tables 1 to 4 illustrate that our proposed estimators of EPE lead to substantial MSE reduction when compared to the “crude” Monte Carlo estimators. Comparing the MSE of the PDS-based estimators, \( \hat{\theta}_{c,p}, \hat{\theta}_{e,b,p}, \) and \( \hat{\theta}_{a,p}, \) we find that our proposed stratified sampling-based estimator of EPE leads to an MSE reduction by a factor of up to 100; this unbiased estimator also dominates the efficient biased estimator of EPE, in some cases quite substantially (see Tables 3 and 4). Comparing MSE of the DJS-based Monte Carlo estimators of EPE, \( \hat{\theta}_{c,d}, \hat{\theta}_{e,b,d}, \) and \( \hat{\theta}_{a,d}, \) we observe that the stratified sampling-based estimator of EPE and our efficient biased EPE estimator perform similarly, which suggests that the asymptotic equivalence result in Proposition 3 can hold for even a moderate number of valuation points. Both efficient DJS estimators lead to substantial MSE reduction when compared to the corresponding crude estimator of EPE. Finally, we note that the variance and MSE for the crude estimators do not change much as the computational budget increases from 12,000 to 120,000, whereas those of efficient estimators reduce by up to an order of ten. This contrast yields the simple, yet useful insight that the number of valuation points should vary as the computational budget varies.

4 Efficient Monte Carlo Estimation of Independent CVA

Independent CVA can be viewed as the weighted sum of expected exposures with the weights being default probabilities. Therefore, our results from Section 3 on efficient estimation of EPE immediately apply here (note that for EPE, the weights are subinterval lengths). To summarize our results on efficient Monte Carlo CVA\(_I\) estimation, we suppress the dependence of CVA on
<table>
<thead>
<tr>
<th>$\hat{\theta}_{c,p}$</th>
<th>EPE</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>34.6559</td>
<td>.047219</td>
<td>.48478</td>
<td>.00380</td>
<td></td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>34.6522</td>
<td>.005028</td>
<td>.43768</td>
<td>.00162</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,b,p}$</td>
<td>34.1802</td>
<td>.077212</td>
<td>.1117</td>
<td>.00253</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,b,d}$</td>
<td>33.9955</td>
<td>.004785</td>
<td>.004786</td>
<td>.00174</td>
</tr>
<tr>
<td>$\hat{\theta}_{u,p}$</td>
<td>33.9964</td>
<td>.072068</td>
<td>.072064</td>
<td>.00518</td>
</tr>
<tr>
<td>$\hat{\theta}_{u,d}$</td>
<td>33.9956</td>
<td>.004865</td>
<td>.004866</td>
<td>.00335</td>
</tr>
</tbody>
</table>

Table 1: $S_0 = 30, \mu = .2, \sigma = .3, s = 12,000$

<table>
<thead>
<tr>
<th>$\hat{\theta}_{c,p}$</th>
<th>EPE</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>34.652</td>
<td>.004791</td>
<td>.4372</td>
<td>.03887</td>
<td></td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>34.6521</td>
<td>.000501</td>
<td>.43303</td>
<td>.01564</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,b,p}$</td>
<td>34.0798</td>
<td>.016741</td>
<td>.024026</td>
<td>.02299</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,b,d}$</td>
<td>33.9948</td>
<td>.000483</td>
<td>.000483</td>
<td>.02409</td>
</tr>
<tr>
<td>$\hat{\theta}_{u,p}$</td>
<td>33.9957</td>
<td>.015533</td>
<td>.015533</td>
<td>.04420</td>
</tr>
<tr>
<td>$\hat{\theta}_{u,d}$</td>
<td>33.9945</td>
<td>.000486</td>
<td>.000486</td>
<td>.03426</td>
</tr>
</tbody>
</table>

Table 2: $S_0 = 30, \mu = .2, \sigma = .3, s = 120,000$

<table>
<thead>
<tr>
<th>$\hat{\theta}_{c,p}$</th>
<th>EPE</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>57.7556</td>
<td>.16106</td>
<td>23.5389</td>
<td>.00389</td>
<td></td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>57.7598</td>
<td>.01628</td>
<td>23.4351</td>
<td>.00189</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,b,p}$</td>
<td>54.1296</td>
<td>.23369</td>
<td>1.6954</td>
<td>.00270</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,b,d}$</td>
<td>52.9238</td>
<td>.015853</td>
<td>.015862</td>
<td>.00189</td>
</tr>
<tr>
<td>$\hat{\theta}_{u,p}$</td>
<td>52.9226</td>
<td>.217</td>
<td>.21698</td>
<td>.00516</td>
</tr>
<tr>
<td>$\hat{\theta}_{u,d}$</td>
<td>52.9198</td>
<td>.015796</td>
<td>.015796</td>
<td>.00390</td>
</tr>
</tbody>
</table>

Table 3: $S_0 = 30, \mu = 1, \sigma = .3, s = 12,000$

<table>
<thead>
<tr>
<th>$\hat{\theta}_{c,p}$</th>
<th>EPE</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>57.7579</td>
<td>.016112</td>
<td>23.4159</td>
<td>.03891</td>
<td></td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>57.7591</td>
<td>.001616</td>
<td>23.4136</td>
<td>.01661</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,b,p}$</td>
<td>53.4783</td>
<td>.047841</td>
<td>.35899</td>
<td>.02412</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,b,d}$</td>
<td>52.9212</td>
<td>.001563</td>
<td>.001564</td>
<td>.02627</td>
</tr>
<tr>
<td>$\hat{\theta}_{u,p}$</td>
<td>52.9189</td>
<td>.045783</td>
<td>.045781</td>
<td>.04657</td>
</tr>
<tr>
<td>$\hat{\theta}_{u,d}$</td>
<td>52.9203</td>
<td>.001565</td>
<td>.001565</td>
<td>.03598</td>
</tr>
</tbody>
</table>

Table 4: $S_0 = 30, \mu = 1, \sigma = .3, s = 120,000$
the stochastic discount factor by assuming zero short rate,

\[ \text{CVA}_I = E \left[ E[V_T \mathbb{1}\{\tau \leq T\}|\tau] \right] = \int_0^T E[V_t]dF_t, \]  

(25)

where \( F \) denotes the cumulative distribution function of \( \tau \), which is assumed to be known (market observable) from, for instance, credit default swap spreads of the counterparty, (see, for instance, [16]).

**Efficient Biased Estimators of CVA\(_I\)** We can employ our MSE minimization formulation to first specify the approximately optimal \( n \) and \( m \) under a fixed computational budget, and then estimate CVA\(_I\) with

\[ \xi_{b,k} = \sum_{i=1}^{n} \bar{V}_i \Delta F_i, \]  

(26)

where \( k = p,d \) denotes PDS and DJS sampling, respectively, \( \bar{V}_i = \sum_{j=1}^{m_i} V_{ij}/m_i \) as defined in Section 3, and \( \Delta F_i \equiv F(t_i) - F(t_{i-1}) \). (We have suppressed the dependence of \( \xi_{b,k} \) on \( m \) and \( n \), i.e., \( \xi_{b,k} \equiv \xi_{b,m,n,k} \)).

**Efficient Unbiased Estimators of CVA\(_I\)** Note that

\[ E[V_T \mathbb{1}\{\tau \leq T\}] = \sum_{i=1}^{n} E[V_t|\tau \in A_i]P(\tau \in A_i), \]  

(27)

where stratum \( i \) is \( A_i = [t_{i-1}, t_i) \). Let \( m_i \), \( i = 1,2,...,n \) denote the number of simulation runs used to estimate \( E[V_t] \), where \( V_i \equiv V_{t_i} \), \( t_0 \equiv 0 \), and \( t_n = T \). Also, \( N = \sum_{i=1}^{n} m_i \) denotes the total number of simulation runs used in estimating CVA\(_I\). Using \( \tau \) as the stratification variable and the identity (27), the stratified sampling estimator of CVA\(_I\) is

\[ \xi_{u,k} = \sum_{i=1}^{n} \bar{V}_{\tau_i} p_i, \]  

(28)

where \( k = p,d \) denotes PDS and DJS sampling, respectively. Also, \( p_i \equiv P(\tau \in A_i) = \Delta F_i \), \( \tau_i \equiv \tau|\tau \in A_i \), and \( \bar{V}_{\tau_i} = \sum_{j=1}^{m_i} V_{\tau_{ij}}/m_i \). That is, to draw a single realization of \( V_{\tau_i} \), we first sample from \( \tau \) conditional on \( \tau \in A_i \); next, given this realization of \( \tau_i \), we generate \( V_{\tau_i} \). In terms of computing time, \( \xi_{b,k} \) requires generating \( N \) realizations of \( V_i \) and \( \xi_{u,k} \) requires \( N \) additional samples from the truncated \( \tau \) based on the strata defined above. Note that since generating \( V_i \) is computationally much more intensive than the truncated \( \tau \), \( \xi_{b,k} \) outperforms \( \xi_{u,k} \) merely marginally in terms of the computational time.

As mentioned before, proportional stratified sampling sets \( m_i = NP(\tau \in A_i) \). For CVA\(_I\) estimation, even if we assume \( A_i \)'s to be equidistant strata, \( P(\tau \in A_i) \)'s are not equal in general. Therefore, proportional stratified sampling does not lead to an equal number of simulation runs at all the valuation points, as is the case for EPE estimation (see Section 3.3).
Similar to our numerical examples in Section 3.4, we have empirically observed that \( \xi_{u,p} \) outperforms \( \xi_{b,p} \) in terms of mean square error.\(^{13}\) Lemma 1 below shows an asymptotic equivalence between \( \xi_{b,d} \) and \( \xi_{u,d} \). The proof of Lemma 1 is similar to Proposition 3, and so it is omitted. Also, similar to our numerical examples in Section 3.4, we have observed that DJS-based biased and unbiased estimators of CVA\(_I\) are equivalent in terms of MSE for large \( n \), which confirms the result of Lemma 1 below.

**Lemma 1.** Consider the proposed estimators of CVA\(_I\), \( \xi_{b,d} \) and \( \xi_{u,d} \) as defined in (26) and (28), respectively. Suppose that proportional sampling is used for both estimators, i.e., \( m_i = N p_i \), and \( \sum_{i=1}^{n} m_i = N \), \( i = 1, \ldots, n \). We assume that \( E[V_i^2] < \infty \), \( i = 1, \ldots, n \). Note that DJS sampling gives \( \text{cov}(V_i, V_j) = 0 \) for all \( i \neq j \) and \( i, j = 1, \ldots, n \). Then the following holds,

\[
\lim_{n \to \infty} n \text{Var}(\xi_{u,d}) = n \text{MSE}(\xi_{b,d}) = c \int_{0}^{T} \text{Var}(V_i) dF(t) \tag{29}
\]

where \( c \) is a constant and \( F \) is the cumulative distribution function of \( \tau \). That is, \( \xi_{b,d} \) and \( \xi_{u,d} \) perform similarly in terms of asymptotic MSE.

Similar to our proposed efficient EPE estimation framework, when the DJS method is chosen for sampling from the credit exposure process, we recommend using efficient biased CVA estimation where the approximately optimal \( m \) and \( n \) are derived via solving the MSE minimization problem discussed in Section 3.2. Note the surprising approximately optimal solution sets \( m = 1 \) and allocates the total computational budget to the number of valuation points. When the PDS method simulates the credit exposure process, we suggest our unbiased stratified sampling based estimator of CVA\(_I\), where the number of strata and simulation runs at each stratum are specified based on the solution to the MSE minimization problem as in Section 3.2.

### 5 Efficient Monte Carlo Estimation of eEPE

In this section we discuss efficient Monte Carlo estimation of effective expected positive exposure, eEPE,

\[
eEPE = \int_{0}^{T} \max_{0 \leq u \leq t} E[V_u] dt,
\]

where \( \{V_t : t \geq 0\} \) denotes the credit exposure process, and \( T \) denotes the expiration time of the transaction with the longest maturity in a portfolio of OTC derivatives held by a financial institution with its counterparty. Consider the time grid, \( 0 \equiv t_0 < t_1 < \ldots < t_n \equiv T \). Set \( \Delta_i \equiv t_i - t_{i-1}, i = 1, \ldots, n \). Monte Carlo estimators of eEPE are,

\[
\hat{\theta}_{m,n,k} = \sum_{i=1}^{n} \max_{1 \leq j \leq t_i} \{\bar{V}_j\} \Delta_i, \tag{30}
\]
where \( \tilde{V}_j \) denotes the \( m \)-simulation run average of the i.i.d. random variables, \( V_{j1}, \ldots, V_{jm} \). The subscript \( k = p \) and \( d \) denote PDS and DJS sampling, respectively. That is, under \( k = p \) (\( k = d \)), \( V_j \)'s are positively correlated (uncorrelated). Consider the mean square error of \( \hat{\theta}_{m,n,k} \),

\[
\text{MSE}(\hat{\theta}_{m,n,k}) = \text{Var}\left( \sum_{i=1}^{n} \max_{1 \leq j \leq i} \{ \tilde{V}_j \} \Delta_i \right) + \left( \sum_{i=1}^{n} E[\max_{1 \leq j \leq i} \{ \tilde{V}_j \}] \Delta_i - eEPE \right)^2.
\]  

(31)

**Bias Decomposition** It is useful to differentiate the following two sources of bias,

\[
\left( \sum_{i=1}^{n} E[\max_{1 \leq j \leq i} \{ \tilde{V}_j \}] \Delta_i - \sum_{i=1}^{n} \max_{1 \leq j \leq i} E[\tilde{V}_j] \Delta_i \right) - \left( eEPE - \sum_{i=1}^{n} \max_{1 \leq j \leq i} E[\tilde{V}_j] \Delta_i \right).
\]  

(32)

That is, the first part of the bias is due to the presence of the maximum operator and the second part is time-discretization bias. Note that for a fixed \( n \), variance of \( \hat{\theta}_{m,n,k} \) converges to zero as \( m \to \infty \). Now, consider Proposition 4 below whose proof is in the Appendix.

**Proposition 4.** Let \( \{ V_t; t \geq 0 \} \) denote the credit exposure process. Let,

\[ M_{n,m,k} \equiv \max\{ \widetilde{V}_1, \ldots, \widetilde{V}_n \}, \]

where \( V_i \equiv V_{ti} \) on the time grid \( 0 \equiv t_0 < t_1 < \ldots < t_n \equiv T \), and \( \tilde{V}_i = \frac{1}{m} \sum_{j=1}^{m} V_{ij} \), \( V_{i1}, \ldots, V_{im} \) are i.i.d random variables. Also, \( k = d \) and \( k = p \) refer to the cases where \( V_i \) are uncorrelated and positively correlated, respectively, resulting from DJS and PDS-based simulation of \( V \). Assume that \( E[V_i^2] < \infty \) for all \( i = 1, \ldots, n \). Let \( M_n \equiv \max\{ E[V_1], \ldots, E[V_n] \} \). Then, as \( m \to \infty \),

\[ M_{n,m,k} \to M_n \text{ a.s.}, \]  

(33)

where a.s. stands for almost surely.

Note that dominated convergence theorem and Proposition 4 give \( E[M_{n,m,k}] \to M_n \) as \( m \to \infty \). So, the first part of the bias

\[
\sum_{i=1}^{n} E[\max_{1 \leq j \leq i} \{ \tilde{V}_j \}] \Delta_i - \sum_{i=1}^{n} \max_{1 \leq j \leq i} E[\tilde{V}_j] \Delta_i
\]

converges to zero as \( m \to \infty \). That is, \( \hat{\theta}_{m,n,d} \) and \( \hat{\theta}_{m,n,p} \) are consistent estimators of \( eEPE_{dst} \) for a fixed \( n \).

In what follows we first show that for a fixed \( n \) and sufficiently large \( m \), \( \hat{\theta}_{m,n,d} \) outperforms \( \hat{\theta}_{m,n,p} \) in terms of variance. Next, after specifying approximates for the order of variance and bias of \( \hat{\theta}_{m,n,k} \), we formulate an MSE minimization problem over \( m \) and \( n \) given a fixed computational budget. Our numerical results indicate that our proposed estimators of \( eEPE \), which use approximately optimal \( m \) and \( n \), lead to substantial MSE reduction when compared to the crude estimators.

---

\(^{14}\)Note that \( M_{n,m,k} \leq \sum_{i=1}^{n} \tilde{V}_i \) and Proposition 4 assumes \( E[\tilde{V}_i] = E[V_i] < \infty \).
5.1 Comparing PDS and DJS-based Monte Carlo Estimators of eEPE

We are to compare the variance of \( \hat{\theta}_{m,n,p} \) and \( \hat{\theta}_{m,n,d} \) for a fixed \( n \) and sufficiently large \( m \). Set

\[
\theta \equiv \sum_{i=1}^{n} \max_{1 \leq j \leq i} E[V_j] \Delta_i, \quad \hat{\theta}_{m,n,k} \equiv \sum_{i=1}^{n} \max_{1 \leq j \leq i} \{ \bar{V}_j \} \Delta_i,
\]

where \( k = p \) (\( k = d \)) refer to the cases where \( V_i \) and \( V_j \) for any \( i \neq j \), are positively correlated (uncorrelated). In what follows we find it useful to append a second subscript \( m \) to emphasize that the average is based on \( m \) i.i.d random variables and a third subscript \( k = d \) or \( p \) to indicate DJS or PDS.

Denote by \( \delta_{i,j} \equiv E[V_i] - E[V_j] \) and \( \delta \equiv \min\{|\delta_{i,j}| : i \neq j, i,j = 1,...,n\} \). Without loss of generality, assume \( \delta > 0 \). Let \( \sigma_{i,j,k} \) denote the standard deviation of \( V_i - V_j \) under estimation method type \( k \) and \( \sigma_{\max} \equiv \max\{\sigma_{i,j,k} : i,j = 1,...,n, k = d,p\} \). For \( i = 1,...,n \), let \( \tau_i \) denote the index for which \( \max\{E[V_1],...,E[V_i]\} \) is attained and \( \tau_{i,m,k} \) be the index for which \( \max\{\bar{V}_{1,m,k},...,\bar{V}_{i,m,k}\} \) is achieved. It then follows from these definitions that

\[
\theta = \sum_{i=1}^{n} E[V_{\tau_i}] \Delta_i \quad \text{and} \quad \hat{\theta}_{m,n,k} = \sum_{i=1}^{n} \bar{V}_{\tau_{i,m,k},m,k} \Delta_i. \tag{34}
\]

For \( k = d \) or \( p \) and \( i = 2,...,n \), the probability that simulations do not yield the right \( \tau_i \) can be bounded from above as follows

\[
P(\tau_{i,m,k} \neq \tau_i) \leq \sum_{j \neq \tau_i, j=1,...,i} P(\bar{V}_{\tau_{i,m,k}} - \bar{V}_{j,m,k} < 0)
\]

\[
= \sum_{j \neq \tau_i, j=1,...,i} P(\bar{V}_{\tau_{i,m,k}} - \bar{V}_{j,m,k} - \delta_{\tau_i,j} < -\delta_{\tau_i,j})
\]

\[
< \sum_{j \neq \tau_i, j=1,...,i} P(\bar{V}_{\tau_{i,m,k}} - \bar{V}_{j,m,k} - \delta_{\tau_i,j} < -\delta)
\]

\[
< \sum_{j \neq \tau_i, j=1,...,i} P(|\bar{V}_{\tau_{i,m,k}} - \bar{V}_{j,m,k} - \delta_{\tau_i,j}| > \delta)
\]

\[
\leq \sum_{j \neq \tau_i, j=1,...,i} \frac{\sigma_{\tau_{i,j},k}^2}{m\delta^2}
\]

\[
\leq \frac{(i-1) \cdot \sigma_{\max}^2}{m\delta^2}, \tag{35}
\]

where (35) follows from the Chebyshev’s inequality.

Consider the event \( B_m = \{ \tau_i = \tau_{i,m,d} = \tau_{i,m,p}, \text{for all } i = 1,...,n \} \). It makes sense to call \( B_m \) the desirable event and \( B_m^c \) the undesirable event. Let \( \hat{\theta}_{m,n,k,B_m} \) denote \( \hat{\theta}_{m,n,k} \) conditional on the event \( B_m \). We have that

\[
\Var(\hat{\theta}_{m,n,d,B_m}) < \Var(\hat{\theta}_{m,n,p,B_m}). \tag{36}
\]
This order can be established by first noting that

\[ \hat{\theta}_{m,n,k,B} = \sum_{i=1}^{n} V_{\tau_{i,m,k}} \Delta_i. \]  

Then since \( V_i \) and \( V_j \), for any \( i \neq j \), are positively correlated (uncorrelated) under PDS (DJS) sampling, the variance of expression (37) is lower under DJS than under PDS.

Note that:

\[ P(B^c_m) \leq \sum_{i=2}^{n} P(\tau_{i,m,d} \neq \tau_i) + \sum_{i=2}^{n} P(\tau_{i,m,p} \neq \tau_i) \]  

\[ < 2 \sum_{i=2}^{n} \frac{(i - 1) \cdot \sigma_{\text{max}}^2}{m \delta^2}, \]  

The above argument leads to the following result.

**Proposition 5.** Consider the desirable event \( B_m \) as defined above. First, conditional on this event, (36) holds. Secondly, the desirable event occurs asymptotically almost surely as \( m \to \infty \). That is, \( \lim_{m \to \infty} P(B_m) = 1 \). More specifically, \( P(B^c_m) \) goes to zero at rate \( 1/m \) as \( m \to \infty \).

Proposition 5 suggests that for sufficiently large \( m \), \( \text{Var}(\hat{\theta}_{m,n,d}) \leq \text{Var}(\hat{\theta}_{m,n,p}) \). Our various numerical examples of Section 5.3 use \( m > 400 \); they all indicate that \( \text{Var}(\hat{\theta}_{m,n,d}) \leq \text{Var}(\hat{\theta}_{m,n,p}) \).

### 5.2 Efficient Monte Carlo Estimation of eEPE

Similar to our approach in subsection 3.2, we would like to find the number of valuation points, \( n \), and the number of simulation runs at each valuation point, \( m \), to minimize \( \text{MSE}(\hat{\theta}_{m,n,k}) \) given a fixed computational budget, \( s \), that is proportional to, \( mn \). To do so, we need to specify the order of the variance and bias of the Monte Carlo estimator of eEPE, \( \hat{\theta}_{m,n,k} \). We are not concerned with deriving sharp estimates of the orders of variance and bias. In fact, our numerical examples indicate that choosing approximately optimal \( m \) and \( n \) using even very rough approximates for the orders of variance and bias lead to substantial MSE reduction. The following is used to formulate our MSE minimization problem: for \( k = p \) or \( d \),

\[ \text{Var}(\hat{\theta}_{m,n,k}) \approx \frac{c_{1,k}}{mn} + \frac{c_{2,k}}{m} \quad \text{and} \quad \text{Bias}(\hat{\theta}_{m,n,k}) \approx \frac{c_{3,k}}{m} + \frac{c_{3}}{n}, \]  

for some constants \( c_{1,k}, c_{2,k}, c_{3} \). The above approximation of the order of bias uses (32) and Proposition 4. Note that our rough approximate of the order of variance, applicable to both \( \hat{\theta}_{m,n,d} \) and \( \hat{\theta}_{m,n,p} \), is similar to (15). This is because of the presence of the maximum operators that leads to positive covariance terms. To see this, let \( \bar{V}_i \) denote the \( m \)-simulation-run average of the i.i.d random variables, \( V_{i1}, ..., V_{im}, i = 1, ..., n \), and consider an equidistant time grid with \( n \) time points, \( \Delta = T/n \). Note that,

\[ \text{Var}(\hat{\theta}_{m,n,k}) = \Delta^2 \text{Var} \left( \bar{V}_1 + \max\{\bar{V}_1, \bar{V}_2\} + ... + \max\{\bar{V}_1, ..., \bar{V}_n\} \right), \]
is equal to $\Delta^2 = \frac{T^2}{n^2}$ times the sum of $n$ non-zero variance terms and $n(n-1)/2$ positive covariance terms both for $k = d$ and $k = p$. This leads to a result similar to (15).

Given (40), we recommend solving the following MSE minimization problem to specify the approximately optimal $m$ and $n$,

$$\min_{m,n} \left( \frac{c_1}{mn} + \frac{c_2}{m} + \left( \frac{c_3}{m} + \frac{c_4}{n} \right)^2 \right) \quad \text{subject to} \quad s = cmn,$$

for some constants $c_1, c_2, c_3, c_4,$ and $c$.

### 5.3 Numerical Examples

Our numerical examples presented below illustrate the efficiency of our proposed estimators of eEPE.\textsuperscript{15} As in Section 3.4, we consider the geometric-Brownian-motion stylized example, where $V_t \equiv S_t = S_0 e^{X_t}$ with $X$ being a Brownian motion with drift $\mu$ and volatility $\sigma$. Let $\hat{\theta}_{c,p}$ and $\hat{\theta}_{c,d}$ denote the “crude” Monte Carlo estimators of eEPE under PDS and DJS sampling, respectively. That is,

$$\hat{\theta}_{c,k} = \sum_{i=1}^{n} \max_{1 \leq j \leq i} \bar{V}_j \Delta_i,$$

where $k = p, d$ and $\Delta_i = t_i - t_{i-1}$ and the $t_i$’s are 1, 2, 3, 4, 8, 12, 18, 24, 36, 49 weeks and 1 year, with $t_{12} = T = 1$ year. Let $\hat{\theta}_{e,p}$ and $\hat{\theta}_{e,d}$ denote the efficient Monte Carlo estimators of eEPE under PDS and DJS sampling, respectively, based on an equidistant time grid, i.e., expression (42) with $\Delta_i \equiv \Delta = T/n$) and resulting from solving the MSE minimization problem (41) (with constants $c_i$, $i = 1, 2, 3,$ and $c$ therein set to 1) in Section 5.2. In particular, under $s = 12,000$, the optimal $n = 29$ and $m = 414$, and under $s = 120,000$, the optimal $n = 62$ and $m = 1935$. Our various numerical examples result in findings similar to those for the EPE estimation. For example, Tables 5 to 8, all based on $10^4$ replications, show that the variance of the DJS-based estimators are much lower than that of the corresponding PDS-based estimators. Also, our proposed estimators of eEPE substantially outperform the crude Monte Carlo estimators in terms of MSE; for instance, MSE is reduced by a factor of 100 in Table 8.

\textsuperscript{15}We refer the reader to Section (D) of the Appendix for a discussion on eEPE\textsubscript{dst} and numerical illustrations of Propositions 4 and 5.
<table>
<thead>
<tr>
<th>eEPE</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}_{c,p}$</td>
<td>57.2278</td>
<td>.10931</td>
<td>22.4936</td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>57.2233</td>
<td>.034659</td>
<td>22.3768</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,p}$</td>
<td>53.4344</td>
<td>.19358</td>
<td>1.0731</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,d}$</td>
<td>53.4379</td>
<td>.011188</td>
<td>.89722</td>
</tr>
</tbody>
</table>

Table 5: $S_0 = 30, \mu = 1, \sigma = .25, s = 12,000$

<table>
<thead>
<tr>
<th>eEPE</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}_{c,p}$</td>
<td>57.2277</td>
<td>.010824</td>
<td>22.3945</td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>57.2262</td>
<td>.003591</td>
<td>22.3734</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,p}$</td>
<td>52.9363</td>
<td>.03962</td>
<td>23301</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,d}$</td>
<td>52.9354</td>
<td>.001083</td>
<td>.19367</td>
</tr>
</tbody>
</table>

Table 6: $S_0 = 30, \mu = 1, \sigma = .25, s = 120,000$

<table>
<thead>
<tr>
<th>eEPE</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}_{c,p}$</td>
<td>81.0388</td>
<td>.24286</td>
<td>101.0233</td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>81.0309</td>
<td>.0843</td>
<td>100.7055</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,p}$</td>
<td>72.8899</td>
<td>.3986</td>
<td>3.9705</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,d}$</td>
<td>72.8885</td>
<td>.024652</td>
<td>3.5914</td>
</tr>
</tbody>
</table>

Table 7: $S_0 = 30, \mu = 1.5, \sigma = .25, s = 12,000$

<table>
<thead>
<tr>
<th>eEPE</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}_{c,p}$</td>
<td>81.0332</td>
<td>.024156</td>
<td>100.692</td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>81.0302</td>
<td>.008395</td>
<td>100.6154</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,p}$</td>
<td>71.8779</td>
<td>.083579</td>
<td>0.85454</td>
</tr>
<tr>
<td>$\hat{\theta}_{e,d}$</td>
<td>71.8807</td>
<td>.002425</td>
<td>0.77819</td>
</tr>
</tbody>
</table>

Table 8: $S_0 = 30, \mu = 1.5, \sigma = .25, s = 120,000$
6 Conclusion: A Two-Step Monte Carlo CCR Framework

It has become increasingly crucial for financial institutions to actively manage their counterparty credit risk. Proper counterparty credit risk management is challenging and computationally intensive. Monte Carlo simulation is often used for CCR pricing and measurement. Poor Monte Carlo CCR estimation can lead to overestimation or underestimation of CCR risk. For instance, Basel III CVA capital charges or the newly devised BCBS capital charges on banks against their central counterparty credit risk could be underestimated (overestimated) significantly. We improve the existing widely used Monte Carlo CCR frameworks by substantially increasing the efficiency of Monte Carlo estimators of the key CCR measures: EPE, CVA, and eEPE.

In our two-step proposed framework, the counterparty credit risk modeler first needs to choose between the two credit exposure sampling methods: PDS or DJS. Introducing and using the notion of marginal matching, we identify conditions under which the so-called path dependent simulation (PDS) method, which simulates the credit exposure process based on the finite dimensional distributions of the underlying risk factors, leads to CCR estimators whose variance is substantially larger than the variance of the CCR estimators calculated based on the so-called direct jump to simulation date (DJS) method. Taking into account the computational time in parallel with the mean square error, we demonstrate that DJS sampling is preferable to PDS sampling for path independent derivatives. For path dependent derivatives since the computational time of the DJS-based estimator usually exceeds that of the PDS-based estimator, the two sampling methods could become approximately equivalent in some cases.

Next, in the second step, the CCR modeler needs to choose the number of valuation points and simulation runs at each valuation point for efficient EPE, CVA$_I$, and eEPE estimations. We show that the mean square error (MSE) of the crude Monte Carlo estimators of EPE, CVA, and eEPE can be substantially reduced by solving approximate MSE minimization problems that specify how to achieve an approximately optimal balance between bias squared and variance. These MSE minimization problems can be easily solved after approximate orders of variance and bias for each of the above mentioned CCR measures under the PDS and DJS methods are derived.

For efficient EPE and CVA$_I$ estimation, if the PDS method has been chosen in the first step above, we recommend employing our unbiased stratified sampling-based estimators of EPE and CVA$_I$. These unbiased estimators of EPE and CVA$_I$ use stratified sampling with the number of strata and simulation runs (allocated to each stratum) being chosen based on the solution to the above mentioned MSE minimization problems.

Our numerical examples indicate that unbiased PDS-based estimators of EPE and CVA are preferable to the efficient biased PDS-based estimators of EPE and CVA. Our analytical results suggest that unbiased and biased DJS-based estimators of EPE and CVA are asymptotically equivalent. Our numerical examples confirm this approximate asymptotic equivalence.

Finally, an interesting case arises when the CCR modeler chooses the DJS method for estimating EPE and CVA$_I$. In this case, our proposed efficient EPE and CVA$_I$ estimators use 1 simulation run at each valuation point and the total computational budget is allocated to making the discrete time grid (the set of valuation points) as fine as possible. Our various numerical
examples illustrate that employing this two-step Monte Carlo CCR framework will substantially increase the efficiency of the existing Monte Carlo CCR “engines”.

Appendix

A Proof of Proposition 3

In this proof, for notational simplicity, we suppress the dependence of \( \hat{\theta}_{b,m,n,d} \) and \( \hat{\theta}_{u,m,n,d} \) on \( m \) and \( n \). Note that,

\[
nMSE(\hat{\theta}_{b,d}) = \frac{T}{m} \sum_{i=1}^{n} \text{Var}(V_i) \Delta_i + n \left( \sum_{i=1}^{n} E[V_i] \Delta_i - \int_{0}^{T} E[V_t] dt \right)^2,
\]

where the first term on the right-hand side of the above equality uses \( \text{Var}(\bar{V}_i) = \text{Var}(V_i)/m \). So, \( nMSE(\hat{\theta}_{b,d}) \) converges to \( c \int_{0}^{T} \text{Var}(V_t) dt \) as \( n \to \infty \).

Now, consider \( \text{Var}(\hat{\theta}_{u,d}) \), and let \( I_n = I_n(\tau) \in \{1, ..., n\} \) denote the index of the stratum containing \( \tau \). Set \( p_i = P(\tau \in A_i) = \Delta_i/T \). From standard results on stratified sampling we have,

\[
\text{Var}(\hat{\theta}_{u,d}) = \frac{T^2}{mn} \sum_{i=1}^{n} \text{Var}(V_\tau | \tau \in A_i)p_i = \frac{T^2}{mn} E[\text{Var}(V_\tau | I_n)].
\]

Since \( \int_{0}^{T} \text{Var}(V_t) dt = T E[\text{Var}(V_\tau | \tau)] \), to complete the proof, it suffice to show that, as \( n \to \infty \),

\[
E[\text{Var}(V_\tau | I_n)] \longrightarrow E[\text{Var}(V_\tau | \tau)].
\]

From the formula for the conditional variance, to show the convergence in (45), it suffice to show that, as \( n \to \infty \),

\[
E \left[ (E[\text{Var}(V_\tau | I_n)]^2 \right] \longrightarrow E \left[ (E[\text{Var}(V_\tau | \tau)]^2 \right].
\]

Set \( X = E[\text{Var}(V_\tau | \tau) \text{ and } X_n = E[\text{Var}(V_\tau | I_n). \text{ Note that } X_n \text{ is a martingale because as } n \text{ increases } I_n \text{ generate increasing family of sigma-algebras. We can use martingale convergence theorem (see Chapter 4 of [7]) to conclude that } X_n \text{ converges to } X \text{ almost surely as } n \to \infty. \text{ Using continuous mapping theorem and dominated convergence theorem (see Chapter 1 of [7]) we conclude that, } E[X_n^2] \text{ converges to } E[X^2] \text{ almost surely, and so (46) holds. This completes the proof of Lemma 1.}^{16}

\[^{16}\text{The probabilistic arguments used in second part of the proof are similar to the ones used in the proof of Lemma 4.1 in [12].}\]
B  Proof of Proposition 4

We first consider $M_{2,m,k}$. Let us assume that $M_2 = E[V_2]$ without loss of generality. Note that,

$$\max\{\bar{V}_1, \bar{V}_2\} - E[V_2] = \bar{V}_11\{\bar{V}_1 > \bar{V}_2\} + (\bar{V}_2 - E[V_2])1\{\bar{V}_2 > \bar{V}_1\} - E[V_2]1\{\bar{V}_1 > \bar{V}_2\}. \tag{47}$$

First, consider the indicator random variable, $1\{\bar{V}_1 > \bar{V}_2\}$; the dependence of $\bar{V}_i$ on $m$ is suppressed for notational simplicity. Set $W^k \equiv V_1 - V_2$, where $k = d, p$ refer to the cases where $V_1$ and $V_2$ are uncorrelated and positively correlated, respectively. Note that $1\{\bar{V}_1 > \bar{V}_2\} \leq 1\{W^k > E[W^k]\}$; $W^k_1, \ldots, W^k_m$ are i.i.d random variables and $\bar{W}^k$ is their average. It is well known that $1\{\bar{V}_1 > \bar{V}_2\} \to 0$ a.s. if and only if for all $\epsilon > 0$,

$$P(1\{\bar{V}_1 > \bar{V}_2\} > \epsilon \text{ i.o.}) = 0, \tag{48}$$

where i.o. stands for infinitely often. To see that (48) holds, note that,

$$P(1\{\bar{V}_1 > \bar{V}_2\} > \epsilon) \leq P(1\{\bar{W}^k > E[W^k]\} > \epsilon) \leq \frac{P(|\bar{W}^k - E[W^k]| > \tilde{\epsilon})}{\epsilon^2}, \tag{49}$$

for all $\tilde{\epsilon} > 0$. To derive the last inequality above the Chebyshev’s inequality is used. Then, (49), almost sure convergence of $W^k \to E[W^k]$ following from the strong law of large numbers (SLLN), and Kolmogorov’s 0-1 law, (see Theorem 8.1 of [7]), give (48).

Now, consider the first term on the right side of (47). Given that $\bar{V}_1$ and $1\{\bar{V}_1 > \bar{V}_2\}$, almost surely converge to $E[V_1]$ and zero, respectively, it is not difficult to show that

$$\bar{V}_11\{\bar{V}_1 > \bar{V}_2\} \to 0 \text{ a.s..}$$

To see this, it suffices to write

$$\bar{V}_11\{\bar{V}_1 > \bar{V}_2\} = (\bar{V}_1 - E[V_1])1\{\bar{V}_1 > \bar{V}_2\} + E[V_1]1\{\bar{V}_1 > \bar{V}_2\},$$

and use SLLN for the sequence of indicator random variables and $\bar{V}_1$. The last term on the right side of (47) converges to zero a.s. based on (48). Analogous arguments led to (48) show that the second term on the right side of (47) converges to zero a.s. This completes the proof for $n = 2$. Induction and analogous arguments are employed for the general case.

Suppose that $M_{n-1,m,k} \to M_{n-1}$, a.s. as $m \to \infty$. Assume that $M_n = E[V_n]$. Then, we need to show that a similar almost sure convergence holds for $M_{n,m,k}$. To see this, it suffices to note that, for all $\epsilon > 0$ and $\tilde{\epsilon} > 0$,

$$P(1\{\bar{V}_1 > \max\{\bar{V}_2, \ldots, \bar{V}_n\}\} > \epsilon) \leq \frac{P\left(|\bar{V}_1 - E[V_1]| - (\max\{\bar{V}_2, \ldots, \bar{V}_n\} - E[V_n]) > \tilde{\epsilon}\right)}{\epsilon^2}$$

$$\leq \frac{P\left(|\bar{V}_1 - E[V_1]| > \tilde{\epsilon}\right)}{\epsilon^2} + \frac{P\left(|\max\{\bar{V}_2, \ldots, \bar{V}_n\} - E[V_n]| > \tilde{\epsilon}\right)}{\epsilon^2},$$

which is then used to show that

$$1\{\bar{V}_1 > \max\{\bar{V}_2, \ldots, \bar{V}_n\}\} \to 0 \text{ a.s..}$$

This completes the proof.
Table 9: $\alpha = 0.01, \beta = 0.0004, \sigma = 0.3, r_0 = 0.02$

Table 10: $\alpha = 0.01, \beta = 0.0002, \sigma = 0.3, r_0 = 0.04$

Table 11: $\alpha = 0.01, \beta = 0.0003, \sigma = 0.4, r_0 = 0.03$

C Numerical Examples for Interest Rate Swaps

Tables 9-11 below compare the variance and computing time of the DJS and PDS-based Monte Carlo estimators of EPE for a single interest rate swap contract in the Vasicek short rate framework (see, for instance, [2]). Specifically, the short rate, denoted by $r$, is modeled by $dr_t = (b - ar_t)dt + \sigma dB_t$, where $B$ is a standard Brownian motion and $a > 0$. Recall that standard results from affine term structure modeling expresses the time-$t$ value of a $T$-maturity zero coupon bond as $p(t, T) = e^{A(t, T) - B(t, T)r_t}$, where $A$ and $B$ are given deterministic functions of time, $a$, $b$, and $\sigma$. The interest rate swap in our examples is a forward swap settled in arrears (see, for instance, [2]) with quarterly payments, a principal value of 100, and maturity 1 year. We let $m = 10^4$ and $n = 12$. In particular, the 12 valuation points are equally spaced, i.e., monthly, within the one-year interval. Variances of the DJS and PDS-based estimators $\hat{\theta}_{b,m,n,d}$ and $\hat{\theta}_{b,m,n,p}$ (column 2 of Tables 9-11) are estimated based on 1000 simulation runs. Tables 9-11 are instances of our numerical examples for interest rate swap where the DJS-based estimator of EPE outperforms the PDS-based estimator in terms of variance (by an order of at least 10) while the computing times are approximately equal.

D Numerical Examples for eEPE_{dst}

The numerical results presented in this section demonstrate the consistency of PDS and DJS estimators for eEPE_{dst} and the asymptotic efficiency of DJS over PDS. In particular, they support our Propositions 4 and 5.
We consider the simple forward contract and the underlying security price process following a geometric Brownian motion with initial value $S_0 = 30$, drift $\mu = 0.01$, and volatility $\sigma = 1$ here. We compare the crude PDS and DJS estimators $\hat{\theta}_{c,p}$ and $\hat{\theta}_{c,d}$ as defined in Section 5.3. Each estimation procedure is replicated $10,000$ times to produce the estimates.

In Tables 12 to 15, in addition to presenting the estimator value, variance, MSE (which, unlike that in Section 5.3, is defined with respect to the estimand of $\text{eEPE}_{dst}$), and CPU time, we also include a column named “WrongOrderProb”, which gives the estimate for the probability that the indices at which the running maximums are achieved ever go wrong, i.e., using the notation in Section 5.1 of the main paper, $P(\tau_{i,m,k} \neq \tau_i, \text{for some } i), k = p \text{ or } d$, corresponding to PDS and DJS respectively. The sum of these two probabilities provides an upper bound for $P(B^c_m)$ in the statement of Proposition 5. As these four tables show, this upper bound converges to zero as $m$ increases, which implies $\lim_{m \to \infty} P(B^c_m) = 0$. Also, the bias of both estimators vanishes as $m$ increases; this is consistent with Proposition 4.

<table>
<thead>
<tr>
<th>$\text{eEPE}_{dst}$</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
<th>WrongOrderProb</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}_{c,p}$</td>
<td>41.0514</td>
<td>20.6228</td>
<td>20.7011</td>
<td>0.000329</td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>42.4019</td>
<td>7.901</td>
<td>10.5697</td>
<td>0.000431</td>
</tr>
</tbody>
</table>

Table 12: $m = 50$

<table>
<thead>
<tr>
<th>$\text{eEPE}_{dst}$</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
<th>WrongOrderProb</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}_{c,p}$</td>
<td>40.7481</td>
<td>2.1847</td>
<td>2.1849</td>
<td>0.00156</td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>40.8833</td>
<td>0.77441</td>
<td>0.78761</td>
<td>0.00121</td>
</tr>
</tbody>
</table>

Table 13: $m = 500$

<table>
<thead>
<tr>
<th>$\text{eEPE}_{dst}$</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
<th>WrongOrderProb</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}_{c,p}$</td>
<td>40.7756</td>
<td>0.21446</td>
<td>0.2146</td>
<td>0.0173</td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>40.7708</td>
<td>0.07379</td>
<td>0.07379</td>
<td>0.0090</td>
</tr>
</tbody>
</table>

Table 14: $m = 5000$

<table>
<thead>
<tr>
<th>$\text{eEPE}_{dst}$</th>
<th>Variance</th>
<th>MSE</th>
<th>CPU Time</th>
<th>WrongOrderProb</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}_{c,p}$</td>
<td>40.7708</td>
<td>0.02156</td>
<td>0.02156</td>
<td>0.1708</td>
</tr>
<tr>
<td>$\hat{\theta}_{c,d}$</td>
<td>40.7682</td>
<td>0.00711</td>
<td>0.00711</td>
<td>0.0932</td>
</tr>
</tbody>
</table>

Table 15: $m = 50,000$
References


