EFFICIENT MONTE CARLO FOR DISCRETE VARIANCE CONTRACTS

Nicolas Merener *
Business School, Universidad Torcuato Di Tella
1010 Saenz Valiente, Buenos Aires, (1428), Argentina.
nmerener@utdt.edu. Corresponding author.

Leonardo Vicchi
Instituto Nacional de Matematica Pura e Aplicada
Estrada Dona Castorina 110, Rio de Janeiro, Brasil, CEP 22460-320
lvicchi@gmail.com

Abstract

We develop an efficient Monte Carlo method for the valuation of financial contracts on discretely realized variance. We work with a general stochastic volatility model that makes realized variance dependent on the full path of the asset price. The variance contract price is a high dimensional integral over the fundamental sources of randomness. We identify a two-dimensional manifold that drives most of the uncertainty in realized variance and compute the contract price by combining precise integration over this manifold, implemented as fine stratification or deterministic sampling with quasi-random numbers, with conditional Monte Carlo on the remaining dimensions. For a subclass of models and a class of nonlinear payoffs we derive approximate theoretical results that quantify the variance reduction achieved by our method. Numerical tests for the discretized versions of the widely used Hull-White and Heston models show that the algorithm performs significantly better than standard Monte Carlo, even for fixed computational budgets.

Keywords: Variance contracts; Monte Carlo simulation; Variance reduction techniques

*We thank Jose Blanchet, attendees at ALIO Informs 2010, SIAM Financial Mathematics 2010, Research In Options 2010, and the Fields Institute Mathematical Finance seminar for valuable comments We also thank anonymous referee and Associate Editor for helping us improve this paper. All remaining errors are our own.
1 Introduction

1.1 Our contribution

We develop in this paper an efficient Monte Carlo method for the valuation of derivative contracts on realized variance. Best known is the variance swap, in which two parties agree to exchange a future payment proportional to the variance realized by an asset price over a certain time interval, against a fixed payment agreed at inception. The fair price of future realized variance for the SP500 index is the well known VIX index, a widely followed indicator of uncertainty in the market in the aftermath of the financial crisis of 2008 as discussed in Whaley (2000), Carr and Lee (2009), Carr and Wu (2009). After variance swaps, which have a linear payoff in realized variance, the market has begun to trade more complex contracts. These deliver a stochastic cash flow at a fixed future time $T$, that is a function of the empirical variance $\Omega$, realized between 0 and $T$. In the absence of arbitrage, the initial price can be written as

$$C = B(0,T)E[g(\Omega)]$$  (1)

where $g : \mathbb{R} \to \mathbb{R}$ is a payoff function, the expectation is taken with respect to a risk-neutral measure, and $B(0,T)$ is a deterministic discount factor. For example, the valuation of a volatility swap is associated to a square root function $g$ in (1) and a European call option takes $g(\Omega) = \max\{\Omega - k, 0\}$. The initial length of most variance contracts is of several months. Following market practice we focus on discretely realized variance $\Omega$, defined as the sum of squared daily increments of the logarithm of the asset price over $N$ consecutive days. This is closely related to the sum of squared daily returns.

We assume in this paper that asset returns are generated by a random shocks modulated by an autonomous, but possibly correlated stochastic volatility process. This includes, as special cases, the Hull and White (1987), Heston (1993), and SABR (Hagan et al. (2002)) models among others. We approximate the discrete returns implied by the continuous time models through Euler discretizations (Kloeden and Platen, (1999)) driven by Gaussian shocks. In our general setting realized variance depends on the joint realization of the shocks to the asset price and the path of the variance process, and (1) is a high dimensional integral. This rules out a closed form solution for the price of an arbitrary payoff of realized variance. A numerical method is the natural tool for this problem.

We develop a computational method for (1) that exploits a structural property of payoffs that depend on realized variance. We identify a two-
dimensional manifold, defined by the sum of squared shocks to the asset price and the average level of the variance process over the life of the contract, that largely drives uncertainty in realized variance. Our numerical algorithm combines precise numerical integration over this low dimensional manifold, implemented as finely stratified Monte Carlo or deterministic sampling with quasi-random numbers, with repeated Monte Carlo sampling conditional on the values of the variables at the low dimensional manifold. Allocating extensive computational effort to this low dimensional manifold leads to a price estimator with significantly lower variance than a straightforward Monte Carlo simulation. For a subclass of models and a class of nonlinear payoffs including put and call options we obtain an approximate theoretical result that quantifies the variance reduction achieved by the method.

We implement our method and present numerical evidence of its performance in experiments with realistic parameters under the discretized Hull-White and Heston models. We find that the method achieves significantly lower variance than a straightforward Monte Carlo simulation even for the same computational budget.

1.2 Related Work

Much work has been done recently on variance derivatives. As discussed in Carr and Lee (2009) most of the literature assumes that the realized variance in (1) is recorded continuously in time. Notable exceptions are Broadie and Jain (2008), Windcliff et al. (2006), Carr et al. (2010), Keller-Ressel and Muhle-Karbe (2010), Sepp (2011) and Drimus and Farkas (2011), who preserve the fact that, in practice, realized variance is defined in terms of discrete time returns or logarithmic increments. This literature finds that the difference between discretely and continuously realized variance is economically important for volatility swaps and variance options, with nonlinear payoffs including put and call options we obtain an approximate theoretical result that quantifies the variance reduction achieved by the method.

The approach in Carr et al. (2010) values variance swaps under a general time changed Levy process. This is a nonparametric approach, desirable from a practitioner’s point of view because it eliminates model risk, but restricted to a specific payoff. The analytical transform based method in Keller-Ressel and Muhle-Karbe (2010) is fast and can handle general payoffs, but extending it to general stochastic volatility models as in our paper is not trivial. The approach in Sepp (2011) develops an accurate approximation for the pricing of options on discrete variance under a Heston model with jumps. Drimus and Farkas (2011), in work developed contemporaneously with our paper, introduce a pricing algorithm for discrete variance
contracts that conditions on the continuously integrated variance and which captures most of the correction to realized variance that arises from discrete sampling. An efficient simulation method, as we propose, can be used for an arbitrary European payoff on discretely realized variance, under a wide class of discretized stochastic volatility models, and it involves no approximation beyond the time discretization of the underlying stochastic differential equations.

We achieve variance reduction in the estimation of a high dimensional expectation by integrating deterministically, or stratifying, along a low dimensional manifold, followed by conditional Monte Carlo. Our work is close to the stratification approach in Glasserman et al. (1999) for path dependent derivatives, and Glasserman et al. (2000), in which radial stratification, similar to the integration over the norm of a vector of Gaussian shocks that we propose, is used in a Value-at-Risk application. Related work on variance reduction for stochastic volatility models includes Ben Ameur et al. (1999), but this does not exploit the structure of variance contracts. In Hull and White (1987), pricing under a stochastic volatility model is implemented by simulation of the variance process and closed form pricing conditional on the path of realized volatility. This is also discussed by Boyle et al. (1997) and exploited by Drimus and Farkas (2011). Our approach is different in that we first perform numerical integration on a low dimensional manifold, and then generate Monte Carlo samples.

Some stochastic volatility models, including the Heston model, can be simulated exactly (Broadie and Kaya (2006), Glasserman and Kim (2010)) with non-Gaussian increments. Their emphasis is in finding a discretization that preserves the exact law of the continuous time model. We explore a different question. We start with a continuous time model, adopt the Euler scheme with smallest bias according to Lord et al. (2010), and develop an algorithm that generates samples for the discrete process leading to an estimator with significantly less variance than standard Monte Carlo. The Euler scheme is helped by the very small time step implied by the daily recording frequency in real contracts. Our approach is that of a market participant that consistently uses the same fine Euler discretization for pricing complex contracts and in the calibration to vanilla options. This user is effectively replacing the continuous time model by its discrete counterpart, a common practice in the financial industry. In this setting, we aim to develop a method that efficiently prices the uncertainty in realized variance.

The paper is structured as follows. Section 2 introduces the class of models with autonomous but correlated stochastic volatility. Section 3 discusses the properties of an estimator of an expectation constructed as the
combination of integration on a low dimensional manifold and conditional random sampling. The implementation of the algorithm is in section 4, then tested in realistic examples in section 5. We conclude in section 6.

2 Models

2.1 Continuous time models

Let \( z_t \) and \( w_t \) be independent standard Brownian Motions, and consider an underlying \( S \) following

\[
dS_t = S_t r dt + \sigma(S_t) \sqrt{V_t} dz_t
\]

\[
dV_t = \beta(V_t) dt + \gamma(V_t) \rho dz_t + \sqrt{1 - \rho^2} dw_t
\]

under the risk neutral measure and appropriate technical conditions for \( \sigma, \beta \) and \( \gamma \). This is a general stochastic volatility model in continuous time, with autonomous stochastic volatility. Some special cases of (2) include the Hull and White (1987) model with

\[
dV_t = \mu V_t dt + \sigma V_t \rho dz_t + \sqrt{1 - \rho^2} dw_t
\]

the model by Heston (1993) with

\[
dV_t = \kappa (\theta - V_t) dt + \sigma \sqrt{V_t} dz_t + \sqrt{1 - \rho^2} dw_t
\]

and the SABR model (Hagan et al. (2002)), which is usually expressed in terms of the dynamics of the underlying \( S \) and its stochastic volatility \( \sigma \)

\[
dS_t = S_t r dt + S_t^b \sigma_t dz_t
\]

\[
d\sigma_t = \alpha \sigma_t \rho dz_t + \sqrt{1 - \rho^2} dw_t
\]

The SABR model can be cast in the form of (2) by introducing \( V_t \equiv \sigma_t^2 \) and applying Ito’s Lemma.
2.2 Discrete time schemes

We compute realized variance from discrete time observations of the process (2). Recordings are approximated by an Euler discretization \{\hat{S}, \hat{V}\} over a time grid \(t_0 < t_1 < ... < t_N\) that coincides with the recording times prescribed by the actual contract. This implies a fine time grid, with \(t_{i+1} - t_i = \Delta = 1/252\), because actual contracts are written on daily recordings. The simplest Euler scheme is a first order expansion on \(S\) and \(V\) leading to

\[
\begin{align*}
\hat{S}_{i+1} &= \hat{S}_i + \hat{S}_i \Delta + \sigma(\hat{S}_i)\sqrt{\hat{V}_i}\sqrt{\Delta}Z_{i+1} \\
\hat{V}_{i+1} &= \hat{V}_i + \beta(\hat{V}_i)\Delta + \gamma(\hat{V}_i)\sqrt{\Delta} [\rho Z_{i+1} + \sqrt{1-\rho^2}W_{i+1}] \\
\end{align*}
\] (6)

with \(Z_1, ..., Z_N\) and \(W_1, ..., W_N\) the independent standard normal components of vectors \(\{Z, W\}\). In the Hull-White model (3) it is convenient to apply an Euler rule to the logarithm of the variable of interest and then exponentiate to get

\[
\begin{align*}
\hat{S}_{i+1} &= \hat{S}_i \exp\{(r - \hat{V}_i/2)\Delta + \sqrt{\hat{V}_i}\sqrt{\Delta}Z_{i+1}\} \\
\hat{V}_{i+1} &= \hat{V}_i \exp\{(\mu - \sigma^2/2)\Delta + \sigma\sqrt{\Delta} [\rho Z_{i+1} + \sqrt{1-\rho^2}W_{i+1}]\} \\
\end{align*}
\] (7)

The Euler scheme in (6) allows \(\hat{V}\) to become negative, which in the Heston (4) model leads to a nonsensical solution. This is avoided in some alternative Euler schemes, compared by Lord et al. (2010). Among these, the one with smallest bias for the Heston model is the full truncation scheme

\[
\begin{align*}
\hat{S}_{i+1} &= \hat{S}_i \exp\{(r - \hat{V}_i/2)\Delta + \sqrt{\hat{V}_i}\sqrt{\Delta}Z_{i+1}\} \\
\tilde{V}_{i+1} &= \tilde{V}_i + \kappa(\theta - \max\{\tilde{V}_i, 0\})\Delta + \sqrt{\max\{\tilde{V}_i, 0\}}\sqrt{\Delta} [\rho Z_{i+1} + \sqrt{1-\rho^2}W_{i+1}] \\
\hat{V}_{i+1} &= \max\{\tilde{V}_{i+1}, 0\} \\
\end{align*}
\] (8)

2.3 Realized variance

Our goal is to compute the expectation of a function of discretely recorded variance. This is defined as

\[
\Omega \equiv \sum_{i=0}^{N-1} \log\left(\frac{S_{(i+1)/252}}{S_{i/252}}\right)^2
\] (9)
for recordings of the true continuous time process (2), approximated in this paper by an Euler rule. For notational simplicity, in the remainder of the paper $S_i$ will denote a discretization scheme applied to (2). Under the simplest Euler scheme (6), realized variance (9) can be expressed as a deterministic function of the stochastic shocks, $RV(Z, W)$,

$$RV(Z, W) \equiv \sum_{i=0}^{N-1} \log(1 + r\Delta + \frac{\sigma(S_i)}{S_i}\sqrt{V_i}\sqrt{\Delta}Z_{i+1})^2$$

(10)

where $V_i$ is some function of $\{Z, W\}$ implied by model of choice that we do not need to write explicitly. The price we want to compute is

$$C = B(0, T)E[g(RV(Z, W))]$$

(11)

which is a high dimensional integral in the Gaussian components of $\{Z, W\}$.

3 Combining Integration with Random Sampling

3.1 The general setting

In this section we present a formal algorithm to compute (11) by combining numerical integration over a low dimensional manifold with Monte Carlo (MC) sampling over the remaining dimensions, and characterize the variance reduction achieved by the method. For context, we first consider a simple MC estimator for the expectation of a function of realized variance. This is straightforward (we omit the trivial discount factor): generate $M$ independent paths, each formed by an independent realization of $Z$ and $W$, and compute

$$\hat{C} = \frac{1}{M} \sum_{j=1}^{M} g(RV(Z^j, W^j))$$

(12)

The estimator $\hat{C}$ is unbiased, in the sense that $E[\hat{C}] = E[g(RV(Z, W))]$ and its variance is

$$\text{Var}[\hat{C}] = \frac{\text{Var}[g(RV(Z, W))]}{M}.$$  

(13)

An alternative way of computing (11) is as a deterministic integral over the $2N$-dimensional space spanned by the Gaussian vectors $Z$ and $W$. We write
\[
C = \int_{z_1, \ldots, z_N, w_1, \ldots, w_N} g(RV(z, w)) \prod_{j=1}^{N} n(z_j)n(w_j)dz_jdw_j \tag{14}
\]

with \(n(u)\) the univariate standard normal density. This is exact, but (14) is an integral over a high dimensional space that can rarely be solved analytically. And its numerical solution through simple quadrature suffers from the curse of dimensionality, being too expensive beyond dimension 3 or 4.

The method in this paper lies between (12) and (14) in the sense that integration is performed over few dimensions that substantially contribute to the noise of the estimator, and MC samples are generated conditional on the values of variables on the low dimensional manifold.

Let \(\pi\) be a vector-valued random variable with components \(\pi_1, \ldots, \pi_p\) defined by the action of deterministic functions of \(\mathbb{R}^{2N} \to \mathbb{R}\) on the Gaussian vectors \(\{Z, W\}\). For example, later in this section we take \(\pi = \{Z'Z, v'W\}\) for a constant vector \(v\) so that \(\pi_1\) has \(\chi^2_N\) distribution and \(\pi_2\) is Gaussian, and in section 4 we transform these variables to make them uniformly distributed in \([0, 1]\). Let \(\eta\) be the probability density associated to \(\pi\). We combine numerical integration over the values of the components of \(\pi\) with conditional MC sampling. For simplicity in this section we adopt a Cartesian grid, with step size \(\Delta\), over the values of the components of \(\pi\), although other grids can be implemented. We assume that the payoff function \(g\) and the construction of \(\pi\) are suitable for integration by quadrature on a finite grid, possibly with a small error due to the tail of \(\eta\) and the discrete grid.

Let \(\pi(I)\) be the value of \(\pi\) at the lowest corner of the quadrature cell indexed by \(I = i_1, \ldots, i_p\), and \(M(I)\) the number of MC samples for that cell. The estimator is

\[
\hat{C} = \sum_{i_1=1}^{N_1} \ldots \sum_{i_p=1}^{N_p} \eta(\pi(I))\Delta^p \frac{1}{M(I)} \sum_{j=1}^{M(I)} g(RV(Z^j, W^j|\pi(I))) \tag{15}
\]

where \(Z^j, W^j|\pi(I)\) indicates conditional sampling. We impose that \(Z^i, W^j|\pi(I)\) and \(Z^j, W^j|\pi(I)\) are independent for \(i \neq j\), and \(Z^i, W^j|\pi(I = x)\) and \(Z^j, W^j|\pi(I = y)\) are independent for \(x \neq y\). Then the variance of \(\hat{C}\) is

\[
\text{Var}[\hat{C}] = \sum_{i_1=1}^{N_1} \ldots \sum_{i_p=1}^{N_p} \eta(\pi(I))^2\Delta^{2p} \frac{1}{M(I)} \text{Var}[g(RV(Z^j, W^j|\pi(I)))]
\]

As shown in Chapter 4 in Glasserman (2004) it is convenient to take a proportional allocation of paths \(M(I) = M\eta(\pi(I))\Delta^p\) (or the closest integer,
assuming $M$ large enough for this to be positive). As $\eta$ integrates to 1, the total number of paths used in $\hat{C}$ is close to $M$. This leads to

$$\text{Var}[\hat{C}] \approx \frac{1}{M} \sum_{i_1=1}^{N_1} \cdots \sum_{i_p=1}^{N_p} \eta(\pi(I)) \Delta^p \text{Var}[g(RV(Z^I, W^I|\pi(I)))]$$ (16)

which is a quadrature based approximation for $E[\text{Var}\{g(RV(Z, W))|\pi\}]$. Therefore, for a fixed number of paths $M$, the variance of the estimator $\hat{C}$ is essentially determined by $E[\text{Var}[g(RV(Z, W))|\pi]]$. It is informative to compare this result with (13). By the variance decomposition formula it holds that

$$\text{Var}[g(RV(Z, W))] = \text{Var}[E[g(RV(Z, W))|\pi]] + E[\text{Var}[g(RV(Z, W))|\pi]]$$ (17)

The left side of (17) is independent of $\pi$ and equal to $M$ times (13). The two terms on the right are nonnegative and the second term coincides with $M$ times (16) except for a small quadrature error. Therefore, for a total number of paths $M$, a MC method with proportional allocation can never increase variance relative to the unconditional case, regardless of the specific choice of $\pi$. In practice, however, we want to decrease the variance of an estimator subject to a finite computational budget. Conditioning on a bad choice of $\pi$ might not be useful if the increase in computing time associated to the numerical integral is larger than the speed gain from the lower variance estimator. These observations indicate that a successful implementation of (15) should attempt to identify $\pi_1, \ldots, \pi_p$ with:

- Small dimension $p$, for fast numerical integration.
- A practical method to sample $W$ and $Z$ conditional on $\pi$.
- $E[\text{Var}[g(RV(Z, W))|\pi]]$ much smaller than $\text{Var}[g(RV(Z, W))]$ to achieve significant variance reduction.

We introduce notation. Let $\|V\|_1 \equiv \sum_{i=0}^{N-1} V_i$, be the average of the discretized variance process over a single path, and $\|V\|_2^2 \equiv \sum_{i=0}^{N-1} V_i^2$. Let $\|Z\|_2^2 \equiv \sum_{i=0}^{N-1} Z_{i+1}^2$, be the squared norm of the vector of shocks to the asset price.
3.2 Monte Carlo conditional on $\|Z\|_2^2$ and $\|V\|_1$

In order to identify a good choice of $\pi$ we do a linear expansion of (10)

$$RV(Z, W) \approx \sum_{i=0}^{N-1} (r^2 \Delta^2 + 2 \frac{\sigma(S_i)}{S_i} \sqrt{V_i} \Delta \frac{3}{2} Z_i + \frac{\sigma(S_i)}{S_i}^2 V_i \Delta Z_i^2)$$

(18)

and notice that, for realistic parameters ($r = 0.05, \frac{\sigma(S_i)}{S_i} \sqrt{V_i} = 0.3, \Delta = 1/252$), the terms above are of orders $10^{-8}, 10^{-6},$ and $10^{-3}$ because the random shocks dominate over the drift term. Therefore discrete realized variance is approximately

$$\overline{RV}(Z, W) \equiv \sum_{i=0}^{N-1} \Delta \left( \frac{\sigma(S_i)}{S_i} \right)^2 V_i Z_i^2$$

(19)

Consider now the variance reduction achieved by performing MC conditional on $\pi = \{\pi_1, \pi_2\} = \{\|Z\|_2^2, \|V\|_1\}$. To the extent that significant variability of (19) is explained by $\|Z\|_2^2$ and $\|V\|_1$, then

$$E[\text{Var}[g(\overline{RV}(Z, W))]|\pi]$$

(20)

should be small relative to $\text{Var}[g(\overline{RV}(Z, W))]$. For example, in the special case with linear $\sigma(S)$ in (2), and $V_i = V_0 \forall i < N$, (19) implies that

$$E[\text{Var}[g(\overline{RV}(Z, W))]|\pi] = E[\text{Var}[g(\pi_1 \frac{\pi_2}{N} \Delta)]|\pi] = 0$$

(21)

Moreover, the closeness of (10) and (19) implies that small (20) leads to

$$E[\text{Var}[g(RV(Z, W))]|\pi]$$

also small, as needed for effective variance reduction for exact realized variance (10). Taking $\pi = \{\|Z\|_2^2, \|V\|_1\}$ is successful in eliminating variance because it contains two main sources of stochasticity in $\overline{RV}$: the sum of squared shocks to the underlying, and the path average of the modulating variance process. We stress that the actual computational algorithm will use (10) without neglecting the small terms that led to (19). We only use (19) to derive a result that approximately quantifies the variance reduction achieved by the method.

**Assumption 3.1** For a model (2) and an associated Euler scheme we assume that:

\[10\]
• \( V \) is independent of \( Z \)
• The function \( \sigma \) in (2) is linear

We are assuming that the variance process is not only autonomous but also independent of shocks to the process \( S \). A linear \( \sigma \) is a feature of the original versions of the Hull-White and Heston models.

**Lemma 1** It holds that

\[
\frac{1}{\Delta} E[RV] = E[\|V\|_1] \tag{22}
\]

\[
\frac{1}{\Delta} E[RV|\pi] = \|Z\|^2 \frac{\|V\|_1}{N} \tag{23}
\]

\[
\frac{1}{\Delta^2} \text{Var}[RV] = 2E[\|V\|_2] + \text{Var}[\|V\|_1] \tag{24}
\]

\[
\frac{1}{\Delta^2} \text{Var}[RV|\pi] = -\frac{2}{N} E[\|V\|_2] + 2E[\|V\|_2] \tag{25}
\]

**Proof:**

From (19) and Assumption 3.1 \( \frac{RV(Z, W)}{\Delta} \equiv \sum_{i=0}^{N-1} V_i Z_{i+1}^2 \)

Next, for proof of (22) we have

\[
E[\sum_{i=0}^{N-1} V_i Z_{i+1}^2] = \sum_{i=0}^{N-1} E[V_i] E[Z_{i+1}^2] = \sum_{i=0}^{N-1} V_i = E[\|V\|_1]
\]

For proof of (23) note that \( Z_1, \ldots, Z_N \) are independent and identically distributed, implying \( E[Z_i^2] = \frac{\|Z\|^2}{N} \) for all \( i \). Using independence of \( Z \) and \( V \) together with the law of iterated expectations we get

\[
\frac{1}{\Delta} E[RV|\pi] = E[E[\sum_{i=0}^{N-1} V_i Z_{i+1}^2|V, \|Z\|^2_2]|\pi] = E[\sum_{i=0}^{N-1} V_i E[Z_{i+1}^2|V, \|Z\|^2_2]|\pi] = E\left[ \sum_{i=0}^{N-1} V_i \|Z\|^2_2 \right] = \frac{\|Z\|^2_2}{N} \left( \sum_{i=0}^{N-1} V_i \right) = \|Z\|^2 \frac{\|V\|_1}{N}
\]

For proof of (24) we have that \( \text{Var}[RV] = \frac{1}{\Delta^2} \text{Var}[\sum_{i=0}^{N-1} V_i Z_{i+1}^2] \), therefore
\[
\frac{1}{\Delta^2} \text{Var}[\widehat{R\mathcal{V}}] = \sum_{i=0}^{N-1} \text{Var}[V_i Z_{i+1}^2] + \sum_{i,j=0, i \neq j}^{N-1} \text{Covar}[V_i Z_{i+1}^2, V_j Z_{j+1}^2]
\] (26)

For the variance terms in (26) we have

\[
\] (27)

For the covariance terms in (26), independence of \(V\) and \(Z\) implies that

\[
\text{Covar}[V_i Z_{i+1}^2, V_j Z_{j+1}^2] = \text{Covar}[V_i, V_j]
\] (28)

From (26), (27) and (28) we obtain

\[
\frac{1}{\Delta^2} \text{Var}[\widehat{R\mathcal{V}}] = \sum_{i=0}^{N-1} (3E[V_i^2] - E[V_i]) + \sum_{i,j=0, i \neq j}^{N-1} \text{Covar}[V_i, V_j]
\]

\[
= 2E[\|V\|_2^2] + \text{Var}[\|V\|_1]
\] (29)

For proof of (25) we invoke (17) to write

\[
E[\text{Var}[\widehat{R\mathcal{V}}(Z, W)|\pi]] = -\text{Var}[E[\widehat{R\mathcal{V}}(Z, W)|\pi]] + \text{Var}[\widehat{R\mathcal{V}}(Z, W)]
\] (30)

By (23) it holds that

\[
\frac{1}{\Delta^2} \text{Var}[E[\widehat{R\mathcal{V}}(Z, W)|\pi]] = \text{Var}[\|Z\|_2^2 \frac{\|V\|_1}{N}]
\]

\[
= \frac{1}{N^2} (E[\|V\|_1^2] E[\|Z\|_2^2] - E[\|V\|_1] E[\|Z\|_2])
\]

\[
= \frac{1}{N^2} ((N(N - 1) + N3) E[\|V\|_1^2] - N^2 E[\|V\|_1])
\]

\[
= \text{Var}[\|V\|_1] + \frac{2}{N} E[\|V\|_1^2]
\] (31)

therefore, subtracting (31) from (24) we get that (30) becomes

\[
\frac{1}{\Delta^2} \text{Var}[\widehat{R\mathcal{V}}(Z, W)|\pi] = -\frac{2}{N} E[\|V\|_1^2] + 2E[\|V\|_2^2] \quad \square
\] (32)

Lemma 1 will be used next to derive a result on the magnitude of variance reduction obtained by conditioning on \(\|Z\|_2^2\) and \(\|V\|_1\). Motivated by (17) define residual variance as
\[
\frac{E[\text{Var}[g(\hat{RV}(Z, W))]|\pi]}{\text{Var}[g(\hat{RV}(Z, W))]} \tag{33}
\]

This is the proportion of the variance of standard MC that survives after performing deterministic integration on the low dimensional manifold.

**Assumption 3.2** For a constant \(A\) and \(\epsilon(x) : \mathbb{R} \to \mathbb{R}\) we assume that the payoff function \(g(x)\) satisfies any of the following two conditions,

\[g(x) = A + x - \epsilon(x), \text{ for nondecreasing } g(x) \text{ and } \epsilon(x)\]
\[g(x) = A - x - \epsilon(x), \text{ for nonincreasing } g(x) \text{ and } \epsilon(x)\]

**Remark:** The class of payoff functions \(g(x)\) that satisfy Assumption 3.2 includes calls and puts with strike \(k\), which can be written as

- **call**: \(\max\{x - k, 0\} = x - \epsilon(x), \text{ for } \epsilon(x) = \min\{x, k\}\)
- **put**: \(\max\{k - x, 0\} = k - x - \epsilon(x), \text{ for } \epsilon(x) = -\max\{x - k, 0\}\)

**Theorem 2** Under Assumptions 3.1 and 3.2, the residual variance after integrating on \(\pi = \{\|Z\|^2_2, \|V\|_1\}\) satisfies:

\[
\text{Residual Variance} \leq \frac{(2E[\|V\|_2^2 - \frac{1}{N}\|V\|_1^2])\Delta^2}{\text{Var}[g(\hat{RV}(Z, W))]} - \frac{\text{Var}[\hat{RV}(Z, W)] - \text{Var}[\|Z\|^2_2 \|V\|_1 N] \Delta^2}{\text{Var}[g(\hat{RV}(Z, W))]} \tag{34}
\]

**Proof:**

Under Assumption 3.2 it holds that

\[
\text{Var}[\hat{RV}(Z, W)|\pi] = \text{Var}[g(\hat{RV}(Z, W))|\pi] + \text{Var}[\epsilon(\hat{RV}(Z, W))|\pi] + 2\text{Covar}[g(\hat{RV}(Z, W)), \epsilon(\hat{RV}(Z, W))|\pi] \tag{35}
\]

and the covariance term is nonnegative (Schmidt, 2003) because \(g\) and \(\epsilon\) are both nondecreasing or nonincreasing. Therefore

\[
\text{Var}[g(\hat{RV}(Z, W))|\pi] \leq \text{Var}[\hat{RV}(Z, W)|\pi] \tag{36}
\]

implying that

\[
E[\text{Var}[g(\hat{RV}(Z, W))|\pi]] \leq E[\text{Var}[\hat{RV}(Z, W)|\pi]] \tag{37}
\]
We bound residual variance (33) using (25) and (37) to get

\[
\text{Residual Variance} \leq \frac{E[\text{Var}[\tilde{RV}(Z, W))|\pi]]}{\text{Var}[g(RV(Z, W))]}
\]

which proves the first inequality in (34). For the last equality in (34) we use (17) and (38) to write

\[
\text{Residual Variance} \leq \frac{\text{Var}[\tilde{RV}(Z, W))] - \text{Var}[E[\tilde{RV}(Z, W))|\pi]]}{\text{Var}[g(RV(Z, W))]}
\]

\[
\text{Residual Variance} = \frac{\text{Var}[\tilde{RV}(Z, W))] - \text{Var}[\|Z\|^2_2 \|W\|_N^1] \Delta^2}{\text{Var}[g(RV(Z, W))]}
\]

(39)

Theorem 2 states in the last equality in (34) that the efficiency of the proposed method in reducing variance depends on how close the second moment of \(\|Z\|^2_2 \|W\|_N\) is to the second moment of \(\tilde{RV}\).

The first inequality in (34), has a geometrical interpretation. The quantity inside the expectation in the numerator is the square of the Euclidean distance of a point \(V \in \mathbb{R}^N\) to the subspace \(L = \{ (x_1, ..., x_N) \in \mathbb{R}^N : x_1 = x_2 = ... = x_N \}\). The residual variance is zero only if the discrete time process \(V\) is a constant a.s. By continuity (the expression is a polynomial over the coordinates of \(V\)), it is clear that the closer the process \(V\) is to a constant in a pathwise sense, the greater the variance reduction.

We have shown that the proposed algorithm leads to quantifiable variance reduction. However, its implementation requires the knowledge of the distribution of the path average of \(V\), which in general is not known. In order to address this issue we consider a closely related alternative set of conditioning variables.

3.3 Monte Carlo conditional on \(\|Z\|^2_2\) and \(v^\prime W\)

We continue to assume that Assumptions 3.1 and 3.2 hold. We propose to combine numerical integration and MC in an estimator of the form (15), now taking \(\tilde{\pi} \equiv \{ \|Z\|^2_2, v^\prime W \}\) as conditioning variables. The new conditioning variable, \(v^\prime W\), is a linear combination of shocks \(W\) with weights defined by \(v \in \mathbb{R}^N\), chosen to be close to the path average of \(V\) in a suitable sense, therefore preserving the efficiency gains achieved by taking \(\|V\|_1\) as conditioning variable. Yet, by being a linear combination of shocks \(W\), this
new conditioning variable has a well known Gaussian distribution. Moreover, the components of $W$ are Gaussian even after conditioning on a linear combination of them, therefore sampling is trivial.

The argument that led to (36) in the proof of Theorem 2 continues to hold for the new conditioning set, therefore

$$E[\operatorname{Var}[g(\hat{R}V(Z, W))|\hat{\pi}]] \leq E[\operatorname{Var}[\hat{R}V(Z, W)|\hat{\pi}]] \quad (40)$$

The aim is to find $v$ with $\|v\|_2 = 1$, such that by conditioning on $\pi_2 = v'W$ the residual variance is small. Taking advantage of (40) we choose to minimize $E[\operatorname{Var}[\hat{R}V(Z, W)|\hat{\pi}]]$, which, by (17), is equivalent to maximizing $\operatorname{Var}[E[\hat{R}V|\hat{\pi}]]$. We are looking for

$$\arg \max_{\{v\}} \operatorname{Var}[E[\sum_{i=0}^{N-1} V_i(W)Z_{i+1}^2||Z||_2^2, v'W]] \quad (41)$$

Solving (41) is a daunting task because, in principle, each $V_i$ arising from an Euler rule on (2) is a nonlinear function of $W$. However, for relatively small $N$, which is indeed the setting of importance for this paper, the behavior of the coefficients of (2) as functions of $V$ suggests linearizing $V_i$ as a function of the underlying shocks $W$ to get

$$\arg \max_{\{v\}} \operatorname{Var}[E[\sum_{i=0}^{N-1} (V_0 + b_i^wW)Z_{i+1}^2||Z||_2^2, v'W]]$$

where $b_i^w \equiv \nabla_w V_i$ evaluated at $W = 0$. By independence of $Z$ and $W$ we obtain

$$\arg \max_{\{v\}} \operatorname{Var}[Y(Z)E[\sum_{i=0}^{N-1} b_i^wW|v'W]]$$

for $Y$ a generic function of $Z$. Recalling that $E[E[\sum_{i=0}^{N-1} b_i^wW|v'W]] = 0$, we get

$$\arg \max_{\{v\}} E[E[\sum_{i=0}^{N-1} b_i^wW|v'W]] \quad (42)$$

Let $b \equiv \nabla_w \sum_{i=0}^{N-1} V_i(W = 0)$. Then (42) is equivalent to

$$\arg \max_{\{v\}} \operatorname{Var}[E[b'W|v'W]] \quad (43)$$
and its solution is attained at \( v = \frac{b}{\|b\|_2} = \frac{\nabla w \sum_{i=0}^{N-1} V_i(W=0)}{\|\nabla w \sum_{i=0}^{N-1} V_i(W=0)\|_2} \).

The optimal vector of weights in \( v'W \) is proportional to the gradient of the path average of \( V \). It is in this sense that the optimal linear combination is close to the exact path average \( \|v\|_1 \).

The optimality rule in (43) is analogous to the optimal stratification direction in Glasserman et al., (1999). This led us to consider, following Glasserman et al. (1999), an alternative optimal \( v \) defined as the solution of

\[
\arg\max_{\{v\}} \text{Var}[E[b'W + \frac{1}{2}(W'AW)]|v'W|] (44)
\]

which keeps terms up to order two in the expansion of \( V \) in powers of \( W \), \( f \) is defined by \( f(W) = \|V\|_1 \) and \( H_f \) is the Hessian of \( f \). The solution to (44) is characterized by:

**Proposition 3.1** Let \( \hat{f}(W) = b'W + \frac{1}{2}(W'AW) \), for \( b \in \mathbb{R}^N \) and \( A \in \mathbb{R}^{N \times N} \). Then the optimal stratification direction is the solution to

\[
\arg\max_{\{v\}} (b'v)^2 + \frac{1}{2}(v'Av)^2 (45)
\]

with \( \|v\|_2 = 1 \).

The proof of Proposition 3.1, follows closely along the lines in Glasserman et al. (1999) so we omit it in the interest of space. In unreported numerical experiments we found that including second order terms does not lead to significant variance reduction and that it has a cost in terms of overhead for the algorithm. For this reason we advocate adopting the weights given by the gradient of \( \|v\|_1 \). We follow this prescription in the rest of the paper.

### 4 Implementation of the Algorithm

We implement the algorithm from section 3 without requiring Assumptions 3.1 and 3.2 to hold. We lift now the requirement of independent \( Z \) and \( V \), allowing for nonzero \( \rho \) in (2). By analogy to the solution of (43) we implement the method on a low dimensional manifold spanned by \( \|Z\|_2^2 \) and \( v'W \) where \( v = \nabla_w \|V(Z,W)\|_1 \) evaluated at \( \{Z = 0, W = 0\} \). We do not claim that this conditioning variable remains optimal in the correlated case, but we stress that it is tractable and that introduces no bias in the algorithm.
The initial step of the algorithm is the computation of $v$ proportional to $\nabla_w \|V(Z,W)\|_1$. As an example, we show the case of the Hull-White model (3) with $\rho = 0$ (the case of $\rho \neq 0$ follows exactly along the same lines, and the Heston case is similar). The Euler discretization (7) leads to

$$\|V\|_1 = V_0 + (1 + e^{(\mu - \frac{1}{2}\sigma^2)\Delta + \sigma \sqrt{\Delta} W_1 + e^{2(\mu - \frac{1}{2}\sigma^2)\Delta + \sigma \sqrt{\Delta} (W_1 + W_2)}} + \ldots + e^{(N-1)(\mu - \frac{1}{2}\sigma^2)\Delta + \sigma \sqrt{\Delta} (W_1 + \ldots + W_{N-1})})$$

which leads to

$$\partial_{w_i} \|V\|_1 = (\sigma \sqrt{\Delta}) \sum_{k=1}^{N-1} e^{k(\mu - \frac{1}{2}\sigma^2)\Delta + \sigma \sqrt{\Delta} (W_1 + \ldots + W_k)}$$

and then

$$\nabla_i \|V\|_1(W = 0) = (\sigma \sqrt{\Delta}) \left( \frac{e^{(\mu - \frac{1}{2}\sigma^2)\Delta N - 1} - e^{(\mu - \frac{1}{2}\sigma^2)\Delta - 1}}{e^{(\mu - \frac{1}{2}\sigma^2)\Delta - 1}} \right)$$

Then, the optimal vector $v$ is used in two alternative numerical implementations of the integration over the low dimensional manifold.

4.1 Quasi-MC + conditioning

- Integrate deterministically over $\|Z\|^2_2$ and $v'W$ using quasi-random numbers, as discussed in Chapter 5 in Glasserman [9]. After skipping the first 128 points in a Sobol sequence, we take 8192 points in the square $[0,1] \times [0,1]$. Independent samples of $\|Z\|^2_2$ and $v'W$ are obtained by the application of the inverse $\chi^2_N$ and standard normal distributions on $\{u_1, u_2\}$.

- For each value of $\|Z\|^2_2$ generate the conditional vector $Z$, writing $Z = \|Z\|^2_2 \frac{\xi}{\|Z\|^2_2}$, where $\xi \sim N(0,I)$.

- For each value of $v'W$ generate the conditional vector $W$ using the fact that if $\xi \sim N(0, \Sigma)$ in $\mathbb{R}^d$ with $Y = v'\xi$ stratified for some $v \in \mathbb{R}^d$, then $(\xi|Y = y) \sim N(\Sigma vy, \Sigma - \Sigma vv')$. Consequently, sample $W$ as $W = vv'W + (I - vv')\Lambda$, where $\Lambda \sim N(0,I)$ and $\|v\| = 1$.  

17
• Compute the realized variance and derivative payoff for each sample of \( \{Z, W\} \), and take the simple average over the two dimensional manifold associated to \( \|Z\|^2 \) and \( v'W \).

4.2 Stratification + conditioning

• Stratify \( \|Z\|^2 \) and \( v'W \) by inversion, as explained in Chapter 4 in Glasserman [9]. We partition the square \([0, 1] \times [0, 1]\) through a cartesian grid with \(90 \times 90\) equally sized cells. Within each cell we independently sample a uniformly distributed two dimensional vector \( \{u_1, u_2\} \). Samples of \( \|Z\|^2 \) and \( v'W \) are obtained by the application of the inverse \( \chi^2_N \) and standard normal distributions on \( \{u_1, u_2\} \).

• Same as second and third steps in the QMC based implementation.

• Compute realized variance and the derivative payoff for each sample of \( \{Z, W\} \), and take simple average over the stratification grid for the manifold spanned by \( \|Z\|^2 \) and \( v'W \).

5 Numerical results

We test the algorithms of section 4 implemented in MATLAB on a desktop PC running Windows XP Professional with an Intel Core 2 Duo CPU with 2.80 GHz and 1.87 GB of RAM. We have discussed two implementations of the conditional approach: a possibly biased scheme based on integration through quasi-random numbers and a stratification scheme that is unbiased by construction. We run experiments for the discretized Hull-White and Heston models and investigate the presence of bias, and efficiency in reducing variance. Relative bias is defined as the absolute difference between standard MC and our efficient MC estimates for some function of interest, divided by the standard estimate. In the few cases in which the standard errors are larger than this difference, we conservatively bound the bias by the sum of standard errors although this is not an almost-sure bound. The reduction in variance is quantified as the square of the ratio of standard errors from competing methods. We compute prices for expected realized variance (a linear payoff), variance of realized variance, at the money options and out of the money options. Deeply in the money options are close to linear payoffs so we omit them for brevity.
Table 1: Long computations for expected payoffs of realized variance: linear, variance, at the money call and out of the money call. Designed to quantify bias in the QMC implementation for the Hull-White model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Payoff</th>
<th>Standard MC</th>
<th>QMC + Cond.</th>
<th>Rel. Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=10, σ = 0.001</td>
<td>E</td>
<td>15.869 (0.002)</td>
<td>15.8722 (0.0001)</td>
<td>0.0002</td>
</tr>
<tr>
<td>N=10, σ = 0.001</td>
<td>Var</td>
<td>50.46 (0.03)</td>
<td>50.451 (0.002)</td>
<td>&lt; 0.0007</td>
</tr>
<tr>
<td>N=10, σ = 0.001</td>
<td>ATM</td>
<td>2.788 (0.002)</td>
<td>2.7970 (0.0001)</td>
<td>0.003</td>
</tr>
<tr>
<td>N=10, σ = 0.001</td>
<td>OTM</td>
<td>0.810 (0.001)</td>
<td>0.8052 (0.0000)</td>
<td>0.006</td>
</tr>
<tr>
<td>N=10, σ = 1.0</td>
<td>E</td>
<td>15.876 (0.002)</td>
<td>15.8762 (0.0004)</td>
<td>&lt; 0.0002</td>
</tr>
<tr>
<td>N=10, σ = 1.0</td>
<td>Var</td>
<td>54.27 (0.03)</td>
<td>54.261 (0.005)</td>
<td>&lt; 0.0007</td>
</tr>
<tr>
<td>N=10, σ = 1.0</td>
<td>ATM</td>
<td>2.876 (0.002)</td>
<td>2.8725 (0.0004)</td>
<td>0.001</td>
</tr>
<tr>
<td>N=10, σ = 1.0</td>
<td>OTM</td>
<td>0.854 (0.001)</td>
<td>0.8516 (0.0002)</td>
<td>0.003</td>
</tr>
<tr>
<td>N=50, σ = 0.001</td>
<td>E</td>
<td>79.379 (0.006)</td>
<td>79.3811 (0.0002)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>N=50, σ = 0.001</td>
<td>Var</td>
<td>253.0 (0.1)</td>
<td>252.61 (0.004)</td>
<td>0.002</td>
</tr>
<tr>
<td>N=50, σ = 0.001</td>
<td>ATM</td>
<td>6.301 (0.003)</td>
<td>6.303 (0.0002)</td>
<td>&lt; 0.0006</td>
</tr>
<tr>
<td>N=50, σ = 0.001</td>
<td>OTM</td>
<td>1.553 (0.001)</td>
<td>1.5396 (0.0001)</td>
<td>0.009</td>
</tr>
<tr>
<td>N=50, σ = 1.0</td>
<td>E</td>
<td>79.352 (0.008)</td>
<td>79.372 (0.003)</td>
<td>0.0003</td>
</tr>
<tr>
<td>N=50, σ = 1.0</td>
<td>Var</td>
<td>704.5 (0.5)</td>
<td>703.7 (0.4)</td>
<td>0.002</td>
</tr>
<tr>
<td>N=50, σ = 1.0</td>
<td>ATM</td>
<td>10.211 (0.006)</td>
<td>10.205 (0.004)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>N=50, σ = 1.0</td>
<td>OTM</td>
<td>3.049 (0.003)</td>
<td>3.050 (0.003)</td>
<td>&lt; 0.002</td>
</tr>
</tbody>
</table>

5.1 Hull-White model

Our first set of results for an Euler discretization of the logarithm of S and V in the Hull-White model (7) are long computations with standard errors low enough to identify the bias present in the implementation that relies on quasi-MC for the low dimensional integral. We use 10 million paths for standard MC and 250 conditional random paths for each of the 4096 Sobol points. The computations for options with $N = 10$ days to expiry take about 820 seconds under standard MC, and about 100 seconds for conditional quasi-MC. The out of the money calls are struck one standard deviation above expected future variance. All runs begin from $V_0 = 0.04$, which corresponds to a 20% initial volatility.

We report in Table 1 the mean and standard error of standard MC and conditional QMC estimators in a variety of experiments with correlation $\rho = 0$. For $N = 10$ we show moments for almost zero volatility of variance ($\sigma = 0.001$), and economically very large volatility of variance ($\sigma = 1$) which corresponds to a %100 annual volatility of variance. These cases can be seen as bounds in the range of likely realistic parameter values. The first moment
is virtually identical in both cases: a higher $\sigma$ widens the distribution of realized variance but does not change its mean. Notably, from $\sigma = 0.001$ to $\sigma = 1$, the second moment only increases from 50.46 to 54.27. This is because, for relatively small $N$, most of the variance of variance comes from the finite number of shocks in the construction of realized variance. In $N = 10$ days the variance process $V$ does not depart much from its initial value $V_0$, therefore $RV$ is close to a sum of equally weighted random Gaussian shocks. This suggests that integrating on $||Z||_2^2$ will be helpful in reducing variance for small $N$.

For $N = 50$, Table 1 shows that the second moment of $RV$ is 253.0 for $\sigma = 0.001$ and 704.5 for $\sigma = 1$. A larger number of days increases the importance of large $\sigma$ because $V$ has more time to diffuse away from $V_0$ and also because, as $N$ increases, the variance of a $\chi^2_N$ random variable decreases relative to its mean. This suggests that integration on the path average of $V$ will be helpful in reducing variance for larger $N$.

The last column in Table 1 displays the relative bias associated with the QMC based algorithm for the Hull-White model. The bid-ask spread in the VIX index in the market in recent years was about one percentage volatility point (Carr and Wu (2009)). A conservatively large value for the VIX index is 50%, only surpassed on a few days during the 2008 crisis. Relative bid-ask spreads are therefore rarely smaller than 0.02. Moreover, the algorithms in this paper are to be used for more exotic contracts on realized variance, therefore with relative bid-ask spreads wider than 0.02. The relative biases we report in Table 1 are smaller than 0.02, indicating that the algorithm is sufficiently accurate for practical use.

A theoretical result (17) guarantees that, for same number of paths and proportional allocation, the variance of conditional MC can not be larger than that of standard MC. However, this result is silent about the computational cost required by the low dimensional numerical integration. A fair comparison is achieved by running both methods with the same computational budget. We use 12000 paths for standard MC, 4096 paths for the QMC algorithm, and $90 \times 90$ strata with 1 path per stratum for the stratified implementation. With these parameters the computing costs in all three methods coincide, at 1 second for contracts with $N = 10$ days to expiration and 3 seconds for $N = 50$.

We price ATM calls on variance and OTM calls and puts on variance. Runs with $\sigma = 0.1$ correspond to a 10% volatility of variance, or a mildly stochastic volatility regime, and a second set of runs with $\sigma = 1.0$ corresponds to very high volatility of volatility, in both cases with correlation $\rho = 0$. Results in Table 2 show that the variance reduction is particularly
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N=10, ( \sigma = 0.1 )</td>
<td>OTM Put</td>
<td>0.726 (0.01)</td>
<td>0.7344 (0.0004)</td>
<td>625</td>
</tr>
<tr>
<td>N=10, ( \sigma = 0.1 )</td>
<td>ATM Call</td>
<td>2.858 (0.05)</td>
<td>2.786 (0.0009)</td>
<td>3086</td>
</tr>
<tr>
<td>N=10, ( \sigma = 0.1 )</td>
<td>OTM Call</td>
<td>0.818 (0.03)</td>
<td>0.800 (0.0007)</td>
<td>1837</td>
</tr>
<tr>
<td>N=10, ( \sigma = 1.0 )</td>
<td>OTM Put</td>
<td>0.763 (0.01)</td>
<td>0.750 (0.002)</td>
<td>25</td>
</tr>
<tr>
<td>N=10, ( \sigma = 1.0 )</td>
<td>ATM Call</td>
<td>2.855 (0.05)</td>
<td>2.862 (0.005)</td>
<td>96</td>
</tr>
<tr>
<td>N=10, ( \sigma = 1.0 )</td>
<td>OTM Call</td>
<td>0.861 (0.03)</td>
<td>0.846 (0.004)</td>
<td>59</td>
</tr>
<tr>
<td>N=50, ( \sigma = 0.1 )</td>
<td>OTM Put</td>
<td>1.176 (0.03)</td>
<td>1.157 (0.002)</td>
<td>352</td>
</tr>
<tr>
<td>N=50, ( \sigma = 0.1 )</td>
<td>ATM Call</td>
<td>6.411 (0.09)</td>
<td>6.365 (0.003)</td>
<td>744</td>
</tr>
<tr>
<td>N=50, ( \sigma = 0.1 )</td>
<td>OTM Call</td>
<td>1.384 (0.04)</td>
<td>1.337 (0.002)</td>
<td>494</td>
</tr>
<tr>
<td>N=50, ( \sigma = 1.0 )</td>
<td>OTM Put</td>
<td>2.594 (0.05)</td>
<td>2.609 (0.009)</td>
<td>32</td>
</tr>
<tr>
<td>N=50, ( \sigma = 1.0 )</td>
<td>ATM Call</td>
<td>10.259 (0.2)</td>
<td>10.206 (0.032)</td>
<td>38</td>
</tr>
<tr>
<td>N=50, ( \sigma = 1.0 )</td>
<td>OTM Call</td>
<td>3.049 (0.1)</td>
<td>3.053 (0.016)</td>
<td>32</td>
</tr>
</tbody>
</table>

Table 2: Variance reduction by conditional QMC for the Hull-White model.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N=10, ( \sigma = 0.1 )</td>
<td>OTM Put</td>
<td>0.726 (0.01)</td>
<td>0.734 (0.001)</td>
<td>100</td>
</tr>
<tr>
<td>N=10, ( \sigma = 0.1 )</td>
<td>ATM Call</td>
<td>2.858 (0.05)</td>
<td>2.786 (0.007)</td>
<td>51</td>
</tr>
<tr>
<td>N=10, ( \sigma = 0.1 )</td>
<td>OTM Call</td>
<td>0.818 (0.03)</td>
<td>0.801 (0.006)</td>
<td>25</td>
</tr>
<tr>
<td>N=10, ( \sigma = 1.0 )</td>
<td>OTM Put</td>
<td>0.763 (0.01)</td>
<td>0.750 (0.002)</td>
<td>25</td>
</tr>
<tr>
<td>N=10, ( \sigma = 1.0 )</td>
<td>ATM Call</td>
<td>2.855 (0.05)</td>
<td>2.871 (0.009)</td>
<td>31</td>
</tr>
<tr>
<td>N=10, ( \sigma = 1.0 )</td>
<td>OTM Call</td>
<td>0.861 (0.03)</td>
<td>0.852 (0.007)</td>
<td>18</td>
</tr>
<tr>
<td>N=50, ( \sigma = 0.1 )</td>
<td>OTM Put</td>
<td>1.176 (0.03)</td>
<td>1.158 (0.005)</td>
<td>36</td>
</tr>
<tr>
<td>N=50, ( \sigma = 0.1 )</td>
<td>ATM Call</td>
<td>6.411 (0.09)</td>
<td>6.371 (0.01)</td>
<td>81</td>
</tr>
<tr>
<td>N=50, ( \sigma = 0.1 )</td>
<td>OTM Call</td>
<td>1.384 (0.04)</td>
<td>1.347 (0.009)</td>
<td>20</td>
</tr>
<tr>
<td>N=50, ( \sigma = 1.0 )</td>
<td>OTM Put</td>
<td>2.594 (0.05)</td>
<td>2.608 (0.01)</td>
<td>25</td>
</tr>
<tr>
<td>N=50, ( \sigma = 1.0 )</td>
<td>ATM Call</td>
<td>10.259 (0.2)</td>
<td>10.215 (0.03)</td>
<td>44</td>
</tr>
<tr>
<td>N=50, ( \sigma = 1.0 )</td>
<td>OTM Call</td>
<td>3.049 (0.1)</td>
<td>3.066 (0.03)</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 3: Variance reduction by stratification and conditioning for the Hull-White model.
<table>
<thead>
<tr>
<th>Parameters</th>
<th>QMC + Cond.</th>
<th>Strat. MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N=10, \sigma = 0.1, \rho = -0.75$</td>
<td>315</td>
<td>44</td>
</tr>
<tr>
<td>$N=10, \sigma = 0.1, \rho = -0.25$</td>
<td>816</td>
<td>64</td>
</tr>
<tr>
<td>$N=10, \sigma = 0.1, \rho = 0.00$</td>
<td>3086</td>
<td>51</td>
</tr>
<tr>
<td>$N=10, \sigma = 0.1, \rho = 0.25$</td>
<td>984</td>
<td>64</td>
</tr>
<tr>
<td>$N=10, \sigma = 0.1, \rho = 0.75$</td>
<td>420</td>
<td>33</td>
</tr>
<tr>
<td>$N=10, \sigma = 1.0, \rho = -0.75$</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>$N=10, \sigma = 1.0, \rho = -0.25$</td>
<td>46</td>
<td>39</td>
</tr>
<tr>
<td>$N=10, \sigma = 1.0, \rho = 0.00$</td>
<td>96</td>
<td>31</td>
</tr>
<tr>
<td>$N=10, \sigma = 1.0, \rho = 0.25$</td>
<td>61</td>
<td>51</td>
</tr>
<tr>
<td>$N=10, \sigma = 1.0, \rho = 0.75$</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>$N=50, \sigma = 0.1, \rho = -0.75$</td>
<td>37</td>
<td>20</td>
</tr>
<tr>
<td>$N=50, \sigma = 0.1, \rho = -0.25$</td>
<td>256</td>
<td>81</td>
</tr>
<tr>
<td>$N=50, \sigma = 0.1, \rho = 0.00$</td>
<td>744</td>
<td>81</td>
</tr>
<tr>
<td>$N=50, \sigma = 0.1, \rho = 0.25$</td>
<td>433</td>
<td>81</td>
</tr>
<tr>
<td>$N=50, \sigma = 0.1, \rho = 0.75$</td>
<td>44</td>
<td>20</td>
</tr>
<tr>
<td>$N=50, \sigma = 1.0, \rho = -0.75$</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>$N=50, \sigma = 1.0, \rho = -0.25$</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>$N=50, \sigma = 1.0, \rho = 0.00$</td>
<td>38</td>
<td>44</td>
</tr>
<tr>
<td>$N=50, \sigma = 1.0, \rho = 0.25$</td>
<td>17</td>
<td>11</td>
</tr>
<tr>
<td>$N=50, \sigma = 1.0, \rho = 0.75$</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 4: The effect of correlation on variance reduction for ATM call options under the Hull-White model.

High for short term contracts with mild stochasticity in the variance process, as expected from integrating on $\|Z\|_2^2$. Higher volatility of volatility, or a longer duration, tend to decrease the reduction in variance, but in all cases this is significant. Table 3 shows the variance reduction achieved for the same set of experiments and budgets by the stratified conditional MC implementation of the Hull-White model. Efficiency gains are not as large as those reported in Table 2, but still significant.

In Table 4 we quantify the effect of nonzero correlation on the conditional QMC, and stratified conditional MC methods. Variance reduction is very significant for values of the correlation parameter $\rho$ relatively close to zero. As expected, correlations close to 1 or -1 diminish the effectiveness of our algorithms because the projection of the path average of $V$ on the space spanned by $W$ becomes quite different from the true path path average.
Table 5: Long computations for expected payoffs of realized variance: linear, at the money call and out of the money call. Designed to quantify bias in the QMC implementation for the Heston model.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Payoff</th>
<th>Standard MC</th>
<th>QMC + Cond.</th>
<th>Rel. Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma = 0.2, \kappa = 10$</td>
<td>$E$</td>
<td>42.855 (0.004)</td>
<td>42.860 (0.001)</td>
<td>0.0001</td>
</tr>
<tr>
<td>$\sigma = 0.2, \kappa = 10$</td>
<td>ATM</td>
<td>5.687 (0.003)</td>
<td>5.682 (0.001)</td>
<td>0.0009</td>
</tr>
<tr>
<td>$\sigma = 0.2, \kappa = 10$</td>
<td>OTM</td>
<td>1.579 (0.002)</td>
<td>1.578 (0.001)</td>
<td>&lt; 0.002</td>
</tr>
<tr>
<td>$\sigma = 0.001, \kappa = 20$</td>
<td>$E$</td>
<td>39.841 (0.004)</td>
<td>39.837 (0.001)</td>
<td>&lt; 0.0002</td>
</tr>
<tr>
<td>$\sigma = 0.001, \kappa = 20$</td>
<td>ATM</td>
<td>5.003 (0.003)</td>
<td>5.001 (0.003)</td>
<td>&lt; 0.002</td>
</tr>
<tr>
<td>$\sigma = 0.001, \kappa = 20$</td>
<td>OTM</td>
<td>1.347 (0.001)</td>
<td>1.345 (0.003)</td>
<td>&lt; 0.003</td>
</tr>
</tbody>
</table>

5.2 Heston model

We first quantify bias in the QMC implementation of the Heston model discretized by (8). We use 10 million paths for standard MC and a total of 250*4096 paths for the QMC plus conditional sampling algorithm. Results are shown in Table 5 for contracts with $N = 20$ days to expiration. For the experiments considered, the relative bias under the Heston model implemented through QMC is much smaller than the relative bid-ask spread of 0.02 that is relevant for practical purposes.

Results in Table 6 show the variance reduction for computational budgets of 1.5 seconds, and $N = 20$ days under QMC and stratification implementations. We use a total of 4096 paths for the QMC implementation, 90 × 90 strata for stratified MC, and 12000 paths for standard MC. The initial variance is $\nu_0 = 0.06$, well above the mean variance $\theta = 0.04$. For $\rho = 0$ we consider two cases: $\sigma = 0.2$, which corresponds to a roughly 100% per year volatility of variance in (4) and mild mean reversion, and a second case with very low volatility of variance and higher mean reversion rate. In the first case the variability of $\nu$ is due to its stochasticity, and in the second case the path of $\nu$ is trending strongly, hence also far from constant. The variance reduction is significant for all experiments in Tables 6, although not as large as in the Hull-White model.

Our final set of experiments evaluates the impact of correlation $\rho$ in (3) for the pricing of at the money call options on realized variance. We take parameter values from Broadie and Kaya (2006), calibrated to market data for SP500 options of November 2, 1993. These are $\nu_0 = 0.01020, \kappa = 6.21, \theta = 0.019, \sigma = 0.61, r = 0.0319$. Results in Table 7 show significant variance reduction for various levels of correlation.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma = 0.2, \kappa = 10$</td>
<td>OTM Put</td>
<td>1.399 (0.03)</td>
<td>1.429 (0.005)</td>
<td>39</td>
</tr>
<tr>
<td>$\sigma = 0.2, \kappa = 10$</td>
<td>ATM Call</td>
<td>5.578 (0.09)</td>
<td>5.681 (0.011)</td>
<td>67</td>
</tr>
<tr>
<td>$\sigma = 0.2, \kappa = 10$</td>
<td>OTM Call</td>
<td>1.571 (0.05)</td>
<td>1.571 (0.008)</td>
<td>41</td>
</tr>
<tr>
<td>$\sigma = 0.001, \kappa = 20$</td>
<td>OTM Put</td>
<td>1.243 (0.03)</td>
<td>1.255 (0.004)</td>
<td>56</td>
</tr>
<tr>
<td>$\sigma = 0.001, \kappa = 20$</td>
<td>ATM Call</td>
<td>4.957 (0.08)</td>
<td>4.996 (0.01)</td>
<td>64</td>
</tr>
<tr>
<td>$\sigma = 0.001, \kappa = 20$</td>
<td>OTM Call</td>
<td>1.371 (0.04)</td>
<td>1.339 (0.008)</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 6: Variance reduction through the QMC and stratified implementations of conditional MC in the Heston model.

<table>
<thead>
<tr>
<th>Correlation</th>
<th>QMC + Cond.</th>
<th>Strat. MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho = -0.7$</td>
<td>285</td>
<td>28</td>
</tr>
<tr>
<td>$\rho = 0.0$</td>
<td>460</td>
<td>81</td>
</tr>
<tr>
<td>$\rho = 0.7$</td>
<td>522</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 7: Effect of correlation on variance reduction in the pricing of at the money calls on realized variance under the Heston model. $N = 20, V_0 = 0.010201, \kappa = 6.21, \theta = 0.019, \sigma = 0.61, r = 0.0319$

6 Conclusions

We developed an efficient pricing algorithm for arbitrary European contracts on realized variance under a wide class of discretized stochastic volatility models. The simulation algorithm combines integration by quasi-random numbers, or stratification, over a two dimensional manifold that explains most of the variance of the MC price estimator, followed by exact conditional random sampling of realized variance. Numerical tests show significant variance reduction relative to standard MC for fixed computational budgets. The implementation based on stratification is exact, and the one that relies on QMC shows a bias that is economically small for current applications.
References


