

## IMPORTANCE SAMPLING FOR TAIL RISK IN DISCRETELY REBALANCED PORTFOLIOS

Paul Glasserman

Columbia University  
3022 Broadway, Uris Hall  
New York, NY 10027, U.S.A.

Xingbo Xu

Columbia University  
500 West 120th Street, S. W. Mudd Building  
New York, NY 10027, U.S.A.

### ABSTRACT

We develop an importance sampling (IS) algorithm to estimate the lower tail of the distribution of returns for a discretely rebalanced portfolio — one in which portfolio weights are reset at regular intervals. We use a more tractable continuously rebalanced portfolio to design the IS estimator. We analyze a limiting regime based on estimating probabilities farther in the tail while letting the rebalancing frequency increase. We show that the estimator is asymptotically efficient for this sequence of problems; its relative error grows in proportion to the fourth root of the number of rebalancing dates.

### 1 INTRODUCTION

In contrast to many hedging strategies, which are built upon the assumption of continuous trading, in practice one can only trade discretely. The study of discretely rebalanced portfolios arises naturally in models of transaction costs and discrete hedging, such as [Bertsimas, Kogan, and Lo \(2000\)](#), [Boyle and Emanuel \(1980\)](#), [Duffie and Sun \(1990\)](#), [Leland \(1985\)](#) and [Morton and Pliska \(1995\)](#). The difference between a discretely rebalanced portfolio and its continuous counterpart has received extensive study. For example, [Tankov and Voltchkova \(2009\)](#) study it under jump-diffusion models, drawing on tools from the simulation literature for discrete approximation of continuous processes, such as [Jacod and Protter \(1998\)](#).

The tail distribution of a portfolio's return is a focus of portfolio risk management. To estimate a distribution far in the tail, techniques of rare event simulation have been well developed in various contexts; see [Asmussen and Glynn \(2007\)](#) for background. Applications in finance have been explored, such as the single-period problem in [Glasserman, Heidelberg, and Shahabuddin \(2000\)](#) and the credit risk application in [Glasserman and Li \(2003\)](#). Here we address a dynamic problem of estimating the lower tail of a portfolio that evolves over time and is periodically rebalanced to a fixed set of weights. We develop an importance sampling estimator using a more tractable continuously rebalanced portfolio to design the change in sampling distribution. [Glasserman \(2009\)](#) analyzed the asymptotic behavior of the difference between the discretely rebalanced portfolio and its continuous counterpart as the number of rebalancing dates increases. Here we extend a conditional limiting result in [Glasserman \(2009\)](#) to develop and analyze our importance sampling technique.

In more detail, we derive a limiting result for the discretely rebalanced portfolio conditional on the logarithmic value of the continuous portfolio being a large negative number of order  $O(\sqrt{N})$ , where  $N$  is the number of rebalancing dates over a fixed horizon. We apply an exponential change of measure to the continuous portfolio to set its mean at a target level corresponding to the tail threshold of the discrete portfolio, and we carry out the importance sampling under the new measure. We prove that this algorithm is logarithmically efficient, with a relative error that is  $O(N^{1/4})$ .

Section 2 describes the setting in which we work. Section 3 provides the main theorem for the conditional limit of the discretely rebalanced portfolio, and Section 4 gives the estimator. In Subsections 5.2 and 5.3, a lower bound of the mean of the estimator and an upper bound of its variance are derived, respectively, which leads to the main result of efficiency in Subsection 5.1. In Section 6 a numerical experiment is presented.

### 2 SETTING

We consider a portfolio consisting of  $d$  assets, with a fixed vector  $w = (w_1, \dots, w_d)^\top$  of weights, such that  $\sum_{i=1}^d w_i = 1$ . Each  $w_i$  is the target fraction of the portfolio's value invested in the  $i^{\text{th}}$  asset. A  $d$ -dimensional standard Brownian

motion  $W = (W_1, \dots, W_d)^\top$  drives the prices of the  $d$  assets with dynamics

$$\frac{dS_i(t)}{S_i(t)} = \mu_i dt + \sum_{i=1}^d \sigma_i^\top dW(t), \quad i = 1, \dots, d.$$

The drifts  $\mu_i$  and volatility vectors  $\sigma_i = (\sigma_{i1}, \dots, \sigma_{id})^\top$  are constant. In a continuously rebalanced portfolio, the portfolio weights  $w$  can be maintained throughout. Under such continuous rebalancing, the portfolio value  $V$  has the following dynamics:

$$\frac{dV(t)}{V(t)} = \sum_{i=1}^d w_i \frac{dS_i(t)}{S_i(t)} = \mu_w dt + \sum_{i=1}^d \bar{\sigma}^\top dW(t),$$

where  $\mu_w = \sum_i w_i \mu_i$  and  $\bar{\sigma} = \sum_i w_i \sigma_i$ . Denote the portfolio's volatility by  $\sigma_w = \|\bar{\sigma}\|$ , with  $\|\cdot\|$  being the  $l_2$ -norm. Then  $V$  has a closed form solution

$$V(t) = V(0) \exp\left\{(\mu_w - \frac{1}{2}\sigma_w^2)t + \bar{\sigma}^\top W(t)\right\}. \tag{1}$$

Now consider a risk horizon  $T = 1$ , and rebalancing horizon  $\Delta t = T/N$ . Assume we only rebalance the portfolio at times  $n\Delta t$  for  $n = 1, \dots, N$ , so that after each rebalancing the portfolio weights are equal to  $w$ , while between two rebalancing dates we do nothing to the portfolio. This discretely rebalanced portfolio  $\hat{V}$  evolves from  $n\Delta t$  to  $(n+1)\Delta t$  according to

$$\hat{V}_{n+1} = \hat{V}_n \left( \sum_{i=1}^d w_i \frac{S_i((n+1)\Delta t)}{S_i(n\Delta t)} \right) = \hat{V}_n \left( \sum_{i=1}^d w_i \exp\left\{(\mu_i - \frac{1}{2}\|\sigma_i\|^2)\Delta t + \sigma_i^\top \Delta W(n)\right\} \right).$$

We write  $\hat{V}_n = \hat{V}(n\Delta t)$ ,  $\Delta W(n) = W((n+1)\Delta t) - W(n\Delta t)$ , and normalize both portfolios to  $V(0) = \hat{V}(0) = 1$ . Hence, we can write

$$\hat{V}_N = V(T) \prod_{n=1}^N \frac{\hat{R}_{n,N}}{R_{n,N}},$$

where 
$$\frac{\hat{R}_{n,N}}{R_{n,N}} = \sum_{i=1}^d w_i \exp\left\{(\mu_i - \mu_w - \frac{1}{2}\|\sigma_i\|^2 + \frac{1}{2}\sigma_w^2)\Delta t + (\sigma_i - \bar{\sigma})^\top \Delta W(n\Delta t)\right\}. \tag{2}$$

We focus on estimating the lower tail of the distribution of  $\hat{V}_N$ . We model this by focusing on  $P(\hat{V}_N < v_N)$ , where  $v_N$  is decreasing and converges to 0 as  $N \rightarrow \infty$ . Note that the continuous value  $V$  in (1) is always positive. Since  $\hat{V}_N$  is close to  $V(T)$  as shown in Theorem 1 in Glasserman (2009) as well as Theorem 1 later in this paper, together with the fact that  $\log V(T)$  has a normal distribution, such  $v_N$  makes the lower tail of  $\hat{V}_N$  a rare event. Particularly, our work considers the case where  $-\log v_N = O(\sqrt{N})$ .

To design an efficient importance sampling algorithm, we will use  $V(T)$  to define a change of probability, and then analyze the efficiency of the algorithm as  $N \rightarrow \infty$  in Sections 5.1–5.3. In order to do that, we will first look at the conditional limit distribution of  $\hat{V}_N$  given  $-\log V(T) = O(\sqrt{N})$  in Section 3.

### 3 CONDITIONAL LIMIT DISTRIBUTION

In this section, we provide a theorem that will be essential for building the importance sampling algorithm, as well as the analysis of the efficiency of the algorithm.

Theorem 1 in Glasserman (2009) shows that as  $N \rightarrow \infty$ , the continuous and discrete portfolios get closer at such a rate that  $\sqrt{N}(\hat{V}(T) - V(T))/V(T)$  converges to a normal distribution. Furthermore, it shows that when conditioning on  $-\log V(T)$  being a very large number of order  $\Theta(\sqrt{N})$ , the ratio  $\log(\hat{V}(T)/V(T))$  converges to a nonzero constant, that is,  $-\log \hat{V}(T)$  is of order  $\Theta(\sqrt{N})$  as well. To make things explicit, we condition on  $\log V(T) = \hat{y}_N = x\sqrt{N} + \mu_w T - \sigma_w^2 T/2$ . From (1) we can see that this is equivalent to conditioning on the Brownian motion  $\bar{\sigma}^\top W(T) = x\sqrt{N}$ .

The following theorem explores the conditional distribution of the  $\hat{V}(T)$  (or say  $\hat{V}(T)/V(T)$ ), which tells more details about the rate of the convergence.

**Theorem 1.** For fixed  $x < 0$ ,  $\hat{y}_N = x\sqrt{N} + \mu_w T - \sigma_w^2 T/2$ ,

$$\sqrt{N}(\log \frac{\hat{V}_N}{V(T)} - a_x) | (\log V(T) = \hat{y}_N) \Rightarrow N(b_x, \bar{\sigma}^2), \text{ for some } \bar{\sigma} > 0, \quad (3)$$

where

$$a_x = \frac{1}{2} \sum_i w_i [(\sigma_i - \bar{\sigma})^\top \bar{\sigma}]^2 x^2 / \sigma_w^4, \quad (4)$$

and

$$b_x = \frac{1}{6} \sum_i w_i \left( \left[ \frac{(\sigma_i - \bar{\sigma})^\top \bar{\sigma}}{\sigma_w^2} \right]^3 x^3 + \frac{(\sigma_i - \bar{\sigma})^\top \bar{\sigma}}{\sigma_w^2} [3\|\sigma_i - \bar{\sigma}\|^2 + 4(\mu_i - \mu_w - \frac{1}{2}(\|\sigma_i\|^2 - \sigma_w^2))] x \right). \quad (5)$$

Here  $\Rightarrow$  means convergence in distribution.

**Remark** The discretely rebalanced portfolio can be negative, that is,  $\hat{V}_N < 0$ , however, as we prove in Appendix A.1 the limit is almost surely non-negative, and the possibility that  $\hat{V}_N < 0$  is negligible. We can simply assign any value to the logarithm function for negative  $\hat{V}_N$  and not change the limit distribution.

**Remark** The theorem can be viewed as an extension of Theorem 2 of Glasserman (2009), which only shows the conditional limit of  $\log(\hat{V}_N/V(T))$ , while the theorem above also shows the speed of convergence. This theorem is also closely related to Theorem 1 of Glasserman (2009) when  $x = 0$  and  $a_0 = b_0 = 0$ . If we consider  $\log(\hat{V}_N/V(T)) \approx \hat{V}_N/V(T) - 1$ , then it has a similar result of convergence to a normal distribution with zero mean.

#### 4 IMPORTANCE SAMPLING ESTIMATOR

Following Theorem 1, we have the intuition that, conditional on an extreme value of  $V(T)$ ,

$$\hat{V}_N | (\log V(T) = \hat{y}_N) \approx H_N(\hat{y}_N) = \exp\{\hat{y}_N + a_x + b_x/\sqrt{N}\}.$$

So, we parameterize the loss threshold as  $v_N = H_N(\hat{y}_N)$ . Then in order to compute  $P(\hat{V}_N < v_N)$ , it becomes natural to choose a change of measure that sets the mean of  $\log V(T)$  to  $\hat{y}_N$ .

From the expression for  $V(T)$  in (1),  $\log V(T) \sim N((\mu_w - \sigma_w^2/2)T, \sigma_w^2)$ , so an exponential change of measure using  $\log V(T)$  can be easily defined

$$\frac{dP_\theta}{dP} = \exp\{\theta \log V(T) - \Psi_{\log V(T)}(\theta)\},$$

where  $\Psi_{\log V(T)}(\theta) = \log E[\exp\{\theta \log V(T)\}]$  is the cumulant generating function of  $\log V(T)$ . We use the subscript  $\theta$  to indicate the new measure. And  $\theta$  can be chosen so that the mean of  $\log V(T)$  under the new measure is matched to  $\hat{y}_N$ :

$$E_\theta[\log V(T)] = \Psi'_{\log V(T)}(\theta) = \hat{y}_N = x\sqrt{N} + \mu_w T - \frac{1}{2}\sigma_w^2 T. \quad (6)$$

By substituting the explicit form of  $\Psi_{\log V(T)}(\theta)$ , we can get  $\theta_N = x\sqrt{N}/\sigma_w^2$ , which also solves

$$\Psi^*(\hat{y}_N) = \sup_\theta \{\theta y - \Psi_{\log V(T)}(\theta)\} = \frac{x^2 N}{2\sigma_w^2}.$$

Hence, under the  $P_{\theta_N}$ ,  $\Delta W(n)$  becomes  $\Delta \tilde{W}(n) + \bar{\sigma}x/(\sqrt{\Delta t}\sigma_w^2)$ , where  $\tilde{W}$  is a Brownian motion under  $P_{\theta_N}$ . So we can simply replace  $\Delta W(n)$  with  $\Delta \tilde{W}(n) + \bar{\sigma}x/(\sqrt{\Delta t}\sigma_w^2)$  in (2) to simulate  $\hat{V}_N$ .

The identity

$$P(\hat{V}_N \leq H_N(\hat{y}_N)) = E_{\theta_N}[I_{\{\hat{V}_N \leq H_N(\hat{y}_N)\}} \exp(-\theta_N \log V_N + \Psi_{\log V}(\theta_N))],$$

yields the estimator

$$\hat{p} = I_{\{\hat{V}_N \leq H_N(\hat{y}_N)\}} \exp(-\theta_N \log V_N + \Psi_{\log V}(\theta_N)), \quad \text{under } P_{\theta_N}. \tag{7}$$

## 5 EFFICIENCY OF THE ESTIMATOR

### 5.1 Main Result

To analyze the efficiency of the estimator (7), we examine the squared coefficient of variation  $\text{Var}_{\theta_N}(\hat{p})/E_{\theta_N}[\hat{p}]^2$  or its logarithmic version  $\log(\text{Var}_{\theta_N}(\hat{p}))/\log(E_{\theta_N}[\hat{p}]^2)$ , as  $N \rightarrow \infty$ . To make our analysis easier to follow, we will first present the result, and leave the main parts of the proof to Sections 5.2 and 5.3.

**Theorem 2.** *The estimator  $\hat{p}$  is logarithmically efficient*

$$\lim_{N \rightarrow \infty} \frac{\log E_{\theta_N}[\hat{p}^2]}{\log E_{\theta_N}[\hat{p}]^2} = \frac{1}{2}.$$

In fact, we have

$$\frac{\text{Var}_{\theta_N}[\hat{p}]}{E_{\theta_N}[\hat{p}]^2} = O(\sqrt{N}), \text{ as } N \rightarrow \infty.$$

The result follows immediately from the following two convergence results, which will be derived in the Subsection 5.2 and 5.3

$$E_{\theta_N}[\hat{p}] \geq \Theta\left(\frac{1}{\sqrt{N}}\right) \exp\{-\Psi^*(\hat{y}_N)\}; \tag{8}$$

$$\text{and } E_{\theta_N}[\hat{p}^2] = O\left(\frac{1}{\sqrt{N}}\right) \exp\{-2\Psi^*(\hat{y}_N)\}. \tag{9}$$

**Remark** In fact, as observed by Jose Blanchet, this is the rate one would expect by applying the same change of measure to estimate the continuous quantity  $P(\log V(T) < \hat{y}_N)$ , so it seems reasonable to expect that this is the best one can do with a path-independent change of measure for  $\hat{V}_N$ . It may be possible to improve the order of the relative error through path-dependent importance sampling.

### 5.2 Lower Bound for the Mean

By the definition of  $\Psi^*(\cdot)$ , we have the lower bound

$$\begin{aligned} E_{\theta_N}[\hat{p}] &= E_{\theta_N}[I_{\{\hat{V}_N \leq H_N(\hat{y}_N)\}} \exp(-\theta_N \log V_N + \Psi_{\log V}(\theta_N))] \\ &\geq P_{\theta_N}(\hat{V}_N \leq H_N(\hat{y}_N), \log V(T) \geq \hat{y}_N) \exp\{-\theta_N \hat{y}_N + \Psi_{\log V}(\theta_N)\} \\ &\geq P_{\theta_N}(\hat{V}_N \leq H_N(\hat{y}_N), \log V(T) \geq \hat{y}_N) \exp\{-\Psi^*(\hat{y}_N)\}. \end{aligned} \tag{10}$$

So, we only need to analyze  $P_{\theta_N}(\hat{V}_N \leq H_N(\hat{y}_N), \log V(T) \geq \hat{y}_N)$ , which can be written in an integral form

$$P_{\theta_N}(\hat{V}_N \leq H_N(\hat{y}_N), \log V(T) \geq \hat{y}_N) = \int_0^\infty P_{\theta_N}(\hat{V}_N \leq H_N(\hat{y}_N) | \log V(T) = \hat{y}_N + y) P_{\theta_N}(\log V(T) - \hat{y}_N \in dy). \tag{11}$$

Denote  $a_x = \alpha x^2$  and  $b_x = \beta_1 x^3 + \beta_2 x$  for constants  $\alpha, \beta_1$  and  $\beta_2$  implicitly defined by (4) and (5). Then

$$\begin{aligned} &P_{\theta_N}(\hat{V}_N \leq H_N(\hat{y}_N) | \log V(T) = \hat{y}_N + y) \\ &= P_{\theta_N}(\sqrt{N}(\log \frac{\hat{V}_N}{V(T)} - a_x) \leq b_x - \sqrt{N}y | \log V(T) = \hat{y}_N + y) \\ &= P_{\theta_N}(\sqrt{N}(\log \frac{\hat{V}_N}{V(T)} - a_{x+y/\sqrt{N}}) \leq b_{x+y/\sqrt{N}} - \sqrt{N}y - (2\alpha x + 3\beta_1 x^2 + \beta_2) \frac{y}{\sqrt{N}} - (\alpha + 3\beta_1 x) \frac{y^2}{N} \\ &\quad - \beta_1 \frac{y^3}{N^{3/2}} | \log V(T) = \hat{y}_N + y), \end{aligned} \tag{12}$$

where last equality is the result of evaluating  $a$  and  $b$  at  $x + y/\sqrt{N}$ .

The conditional distribution of  $\Delta W(n)$  under  $P_{\theta_N}$  is the same as the conditional distribution in Theorem 1 in (20), so the result of Theorem 1 still holds under  $P_{\theta_N}$ . Also, the Central Limit Theorem used in the Theorem 1 satisfies the Lyapunov criterion, so we can apply the error estimation for the Central Limit Theorem in Theorem 7.4.1 of Chung (2000) to (12), leading to

$$P_{\theta_N}(\hat{V}_N \leq H_N(\hat{y}_N) | \log V(T) = \hat{y}_N + y) \geq (1 + \varepsilon_N) \Phi\left(-\frac{\sqrt{N}y + (2\alpha x + 3\beta_1 x^2 + \beta_2)yN^{-1/2} + (\alpha + 3\beta_1 x)y^2N^{-1} + \beta_1 y^3 N^{-3/2}}{\tilde{\sigma}}\right), \tag{13}$$

where  $\Phi(\cdot)$  is the standard normal distribution function, and  $\varepsilon_N \rightarrow 0$  as  $N \rightarrow \infty$ .

Then using a lower bound for the normal distribution function in (25) in Appendix A.2, we have that for any fixed  $\delta > 0$ , there exists some constant  $C'$ , such that

$$\begin{aligned} & \Phi\left(-\frac{\sqrt{N}y + (2\alpha x + 3\beta_1 x^2 + \beta_2)yN^{-1/2} + (\alpha + 3\beta_1 x)y^2N^{-1} + \beta_1 y^3 N^{-3/2}}{\tilde{\sigma}}\right) \\ & \geq C \exp\left\{-\frac{1}{2}\left(\sqrt{N}y + (2\alpha x + 3\beta_1 x^2 + \beta_2)\frac{y}{\sqrt{N}} + (\alpha + 3\beta_1 x)\frac{y^2}{N} + \beta_1 \frac{y^3}{N^{3/2}}\right)^2 \tilde{\sigma}^{-2}\right. \\ & \quad \left. - \delta \left|\sqrt{N}y + (2\alpha x + 3\beta_1 x^2 + \beta_2)\frac{y}{\sqrt{N}} + (\alpha + 3\beta_1 x)\frac{y^2}{N} + \beta_1 \frac{y^3}{N^{3/2}}\right| \tilde{\sigma}^{-1}\right\} \\ & \geq C \exp\{-C'(\sqrt{N}y + Ny^2 + y^3 + N^{-1}y^4 + N^{-2}y^5 + N^{-3}y^6)\}, \end{aligned} \tag{14}$$

where (14) hold for sufficient large  $N$ . Now we substitute (12), (13) and (14) back into (11)

$$\begin{aligned} & P_{\theta_N}(\hat{V}_N \leq H_N(\hat{y}_N), \log V(T) \geq \hat{y}_N) \\ & \geq (1 + \varepsilon_N) C \int_0^\infty \exp\{-C'(\sqrt{N}y + Ny^2 + y^3 + N^{-1}y^4 + N^{-2}y^5 + N^{-3}y^6)\} P_{\theta_N}(\log V(T) - \hat{y}_N \in y) \\ & = (1 + \varepsilon_N) \tilde{C} \int_0^\infty \exp\{-\tilde{C}'(\sqrt{N}y + Ny^2 + y^3 + N^{-1}y^4 + N^{-2}y^5 + N^{-3}y^6)\} dy, \end{aligned}$$

for some constants  $\tilde{C}$  and  $\tilde{C}'$ , and the last equality holds because  $\log V(T) - \hat{y}_N \sim N(0, \sigma_w^2)$  under  $P_{\theta_N}$ .

Next, we divide the integral into three parts, and only evaluate the central one to get a lower bound. In the middle interval  $y \in [1/\sqrt{N}, \sqrt{N}]$ , where  $Ny^2 + \sqrt{N}y$  dominates all other terms in the power. Then there exists some constant  $\tilde{C}''$  such that

$$\begin{aligned} & P_{\theta_N}(\hat{V}_N \leq H_N(\hat{y}_N), \log V(T) \geq \hat{y}_N) \\ & \geq (1 + \varepsilon_N) \tilde{C} \left( \int_0^{\frac{1}{\sqrt{N}}} + \int_{\frac{1}{\sqrt{N}}}^{\sqrt{N}} + \int_{\sqrt{N}}^\infty \right) \exp\{-\tilde{C}'(\sqrt{N}y + Ny^2 + y^3 + N^{-1}y^4 + N^{-2}y^5 + N^{-3}y^6)\} dy \\ & \geq (1 + \varepsilon_N) \tilde{C} \int_{\frac{1}{\sqrt{N}}}^{\sqrt{N}} \exp\{-\tilde{C}''(\sqrt{N}y + Ny^2)\} dy \\ & \geq (1 + \varepsilon_N) \frac{\tilde{C}}{\sqrt{N}} \int_1^N \exp\{-\tilde{C}''(y + y^2)\} dy \\ & = \Theta\left(\frac{1}{\sqrt{N}}\right). \end{aligned}$$

This proves (8).

### 5.3 Upper Bound for the Variance

We can write the second moment of the estimator as follows

$$\begin{aligned} E_{\theta_N}[\hat{\rho}^2] & = E_{\theta_N} [I_{\{\hat{V}_N \leq H_N(\hat{y}_N), \log V_N > \hat{y}_N\}} + I_{\{\hat{V}_N \leq H_N(\hat{y}_N), \log V_N \leq \hat{y}_N\}}] \exp(-2\theta_N \log V_N + 2\Psi_{\log V}(\theta_N)) \\ & = I + II. \end{aligned}$$

We can write the second term as

$$\begin{aligned} II &= E_{\theta_N} [I_{\{\hat{V}_N \leq H_N(\hat{y}_N), \log V_N \leq \hat{y}_N\}} \exp(-2\theta_N \log V_N + 2\Psi_{\log V}(\theta_N))] \\ &= E [I_{\{\hat{V}_N \leq H_N(\hat{y}_N), \log V_N \leq \hat{y}_N\}} \exp(-\theta_N \log V_N + \Psi_{\log V}(\theta_N))] \\ &\leq \exp\{-\Psi^*(\log(\hat{y}_N))\} P(\hat{V}_N \leq H_N(\hat{y}_N), \log V_N \leq \hat{y}_N) \\ &\leq \exp\{-\Psi^*(\log(\hat{y}_N))\} P(\log V_N \leq \hat{y}_N). \end{aligned}$$

In the second line, we change the computation back to the original measure.

Denote  $\phi(\cdot)$  as the density function for standard normal distribution. Using the property of the normal tail distribution in (27) in Appendix A.2, we can get an upper bound for  $II$  for sufficiently large  $N$ : there exists  $\tilde{\epsilon}_N \rightarrow 0$  as  $N \rightarrow \infty$  so that

$$II \leq \exp\{-\Psi^*(\hat{y}_N)\} (1 + \tilde{\epsilon}_N) \frac{\phi(x\sqrt{N}/\sigma_w)}{x\sqrt{N}/\sigma_w} = \Theta\left(\frac{1}{\sqrt{N}}\right) \exp\{-2\Psi^*(\hat{y}_N)\}. \tag{15}$$

For the first term  $I$ , we have

$$\begin{aligned} I &= E_{\theta_N} [I_{\{\hat{V}_N \leq H_N(\hat{y}_N), \log V_N > \hat{y}_N\}} \exp\{-2\theta_N \log V_N + 2\Psi_{\log V}(\theta_N)\}] \\ &= E_{\theta_N} [I_{\{\hat{V}_N \leq H_N(\hat{y}_N), \log V_N > \hat{y}_N\}} \exp\{-2\theta_N(\log V_N - \hat{y}_N)\} \exp\{-2\Psi^*(\log(\hat{y}_N))\}]. \end{aligned} \tag{16}$$

So we only need to analyze  $E_{\theta_N} [I_{\{\hat{V}_N \leq H_N(\hat{y}_N), \log V_N > \hat{y}_N\}} \exp\{-2\theta_N(\log V_N - \hat{y}_N)\}]$ .

First, we can write it in integral form as

$$\begin{aligned} &E_{\theta_N} [I_{\{\hat{V}_N \leq H_N(\hat{y}_N), \log V_N > \hat{y}_N\}} \exp\{-2\theta_N(\log V_N - \hat{y}_N)\}] \\ &= \int_0^\infty E_{\theta_N} [I_{\{\sqrt{N}(\log \frac{\hat{V}_N}{V(T)} - a_x) \leq b_x - \sqrt{N}y\}} \exp\{-2\theta_N y\} | \log V(T) = \hat{y}_N + y] P_{\theta_N}(\log V(T) - \hat{y}_N \in dy), \end{aligned} \tag{17}$$

then we rewrite the integrand to use the result of Theorem 1:

$$\begin{aligned} &E_{\theta_N} [I_{\{\sqrt{N}(\log \frac{\hat{V}_N}{V(T)} - a_x) \leq b_x - \sqrt{N}y\}} \exp\{-2\theta_N y\} | \log V(T) = \hat{y}_N + y] \\ &= \exp\{-2\theta_N y\} P_{\theta_N}(\{\sqrt{N}(\log \frac{\hat{V}_N}{V(T)} - a_{x+y/\sqrt{N}}) \leq b_{x+y/\sqrt{N}} - \sqrt{N}y - (2\alpha x + 3\beta_1 x^2 + \beta_2) \frac{y}{\sqrt{N}} \\ &\quad - (\alpha + 3\beta_1 x) \frac{y^2}{N} - \beta_1 \frac{y^3}{N^{3/2}} | \log V(T) = \hat{y}_N + y\}. \end{aligned}$$

Then using Theorem 1 and the error estimate for the Lyapunov Central Limit Theorem in Theorem 7.4.1 of Chung (2000) as in Subsection 5.2, we have

$$\begin{aligned} &E_{\theta_N} [I_{\{\sqrt{N}(\log \frac{\hat{V}_N}{V(T)} - a_x) \leq b_x - \sqrt{N}y\}} \exp\{-2\theta_N y\} | \log V(T) = \hat{y}_N + y] \\ &\leq (1 + \epsilon_N) \exp\{-2\theta_N y\} \Phi\left(-\frac{\sqrt{N}y + (2\alpha x + 3\beta_1 x^2 - \beta_2)yN^{-1/2} + (\alpha + 3\beta_1 x)y^2N^{-1} + \beta_1 y^3N^{-3/2}}{\tilde{\sigma}}\right) \\ &\leq (1 + \epsilon_N) \exp\left\{-\frac{2x\sqrt{N}y}{\tilde{\sigma}^2} - \frac{(\sqrt{N}y + (2\alpha x + 3\beta_1 x^2 - \beta_2)yN^{-1/2} + (\alpha + 3\beta_1 x)y^2N^{-1} + \beta_1 y^3N^{-3/2})^2}{2\tilde{\sigma}}\right\}. \end{aligned} \tag{18}$$

The last inequality used the result  $\theta_N = x\sqrt{N}/\sigma_w^2$  (which was derived earlier when the change of measure was first defined in Section 2) and an upper bound for the normal distribution in (26) in Appendix A.2.

Substituting (18) into (17), and dividing the integral into three parts, we get

$$\begin{aligned} &E_{\theta_N} [I_{\{\hat{V}_N \leq H_N(\hat{y}_N), \log V_N > \hat{y}_N\}} \exp\{-2\theta_N(\log V_N - \hat{y}_N)\}] \\ &\leq (1 + \epsilon_N) \left( \int_0^{\frac{1}{\sqrt{N}}} + \int_{\frac{1}{\sqrt{N}}}^{\sqrt{N}} + \int_{\sqrt{N}}^\infty \right) \\ &\exp\left\{-\frac{2x\sqrt{N}y}{\tilde{\sigma}^2} - \frac{1}{2\tilde{\sigma}^2} (\sqrt{N}y + (2\alpha x + 3\beta_1 x^2 + \beta_2) \frac{y}{\sqrt{N}} + (\alpha + 3\beta_1 x) \frac{y^2}{N} + \beta_1 \frac{y^3}{N^{3/2}})^2\right\} \phi\left(\frac{y}{\sigma_w}\right) dy. \end{aligned}$$

In the first integral, the power of the exponential function is bounded as  $y \in [0, 1/\sqrt{N}]$ , so the whole integrand is bounded. As a result, the first integral is of order  $O(\sqrt{N})$ . In the third integral, where  $y > \sqrt{N}$ ,  $-2x\sqrt{N}y/\bar{\sigma}^2 - Ny^2/(2\bar{\sigma}^2) < 0$  for sufficient large  $N$ . So the exponential function is smaller than 1 for large  $N$ , which means the third integral is bounded by the tail distribution  $\Phi(-\sqrt{N}/\sigma_w)$  which is of order  $O(1/\sqrt{N})$  from (27). And in the second integral,  $\sqrt{N}y$  and  $Ny^2$  dominate all other terms in the power. Then we have for some constant  $\bar{C}$

$$\begin{aligned} & E_{\theta_N} [I_{\{\hat{V}_N \leq H_N(\hat{y}_N), \log V_N > \hat{y}_N\}} \exp\{-2\theta_N(\log V_N - \hat{y}_N)\}] \\ & \leq O\left(\frac{1}{\sqrt{N}}\right) + \exp\{-|\Omega(\sqrt{N})|\} + (1 + \varepsilon_N) \int_{\frac{1}{\sqrt{N}}}^{\sqrt{N}} \exp\left\{-\frac{2x\sqrt{N}y}{\bar{\sigma}^2} - \bar{C}Ny^2\right\} dy \\ & \leq O\left(\frac{1}{\sqrt{N}}\right) + \exp\{-|\Omega(\sqrt{N})|\} + \frac{1 + \varepsilon_N}{\sqrt{N}} \int_1^N \exp\left\{-\frac{2xy}{\bar{\sigma}^2} - \bar{C}y^2\right\} dy \\ & = \Theta\left(\frac{1}{\sqrt{N}}\right). \end{aligned} \tag{19}$$

So using (19), we have for (16) that

$$I = O\left(\frac{1}{\sqrt{N}}\right) \exp\{-2\Psi^*(\log(\hat{y}_N))\},$$

Together with (15), this proves (9).

### 6 NUMERICAL TEST

We now illustrate the effectiveness of this importance sampling algorithm with a numerical example.

Let  $d = 10$  be the number of assets and set the total time horizon  $T = 1$ . Assume the  $i^{th}$  asset has volatility  $\|\sigma_i\| = 0.025 + 0.0125(i - 1)$  for  $i = 1, \dots, d$ , and all assets are equally correlated with correlation  $\rho = 0.2$ . We consider those cases when the time horizon is divided equally into  $N$  intervals, where  $N = 3, 5, 10, 15, 20$  and  $25$ . Set the coefficient of the condition  $x = -1$ , where  $x$  is as defined in (6). For each  $N$ , we use  $M = 5000$  replications. We use  $c_v^2 = \text{Var}_{\theta_N}(\hat{p}) / (E_{\theta_N}[\hat{p}])^2$  as an indicator for the efficiency of the algorithm.

Table 1 shows the results of the implementation. We denote by  $\bar{p}$  the plain Monte Carlo estimator, whose variance is  $p - p^2$ , where  $p = P(\hat{V}_N \leq H_N(\hat{y}_N))$ . An easy way to check the correctness of the tail probability is that given  $\hat{y}_N = O(\sqrt{N})$ ,  $\log P(\hat{V}_N \leq H_N(\hat{y}_N))$  should be roughly linear in  $N$ . This can be observed by looking at the power term in the first column of Table 1. The last column in Table 1 is the ratio of variances between plain Monte Carlo estimator and our importance sampling estimator. This variance ratio gives the number of plain Monte Carlo replications required for each importance sampling replication in order that the two methods yield the same precision. So the larger is the value of this ratio, the greater is the variance reduction. The results in Table 1 show a very large reduction.

As expected, from Table 1 we can see that  $c_v^2/\sqrt{N}$  is of order  $O(1)$ . From Figure 1, where we run the algorithm for some larger values of  $N$  as well, we can see that this ratio decreases relatively fast for small  $N$ , indicating that the numerical performance is better than the limiting analysis predicts; the ratio stabilizes as  $N \rightarrow \infty$ , as guaranteed by the analysis. The figure includes a graph of  $1/\sqrt{N}$  for comparison.

Table 1: Variance reduction for increasing  $N$ .

| $N$ | $c_v^2/\sqrt{N}$ | $P(\hat{V}_N \leq H_N(\hat{y}_N))$ (95% confidence) | $\text{Var}(\bar{p})/\text{Var}_{\theta_N}(\hat{p})$ |
|-----|------------------|---|--|
| 3   | 0.56             | $(2.29, 2.39) \times 10^{-2}$                       | 42.7   |
| 5   | 0.44             | $(3.03, 3.17) \times 10^{-3}$                       | 327  |
| 10  | 0.33             | $(1.93, 2.02) \times 10^{-5}$                       | $5.09 \times 10^4$                                   |
| 15  | 0.27             | $(1.25, 1.32) \times 10^{-7}$                       | $7.45 \times 10^6$                                   |
| 20  | 0.23             | $(8.40, 8.82) \times 10^{-10}$                      | $1.10 \times 10^9$                                   |
| 25  | 0.20             | $(5.66, 5.93) \times 10^{-12}$                      | $1.65 \times 10^{11}$                                |

### 7 CONCLUDING REMARKS

We have developed, analyzed and tested an importance sampling estimator for the tail of the return distribution of a discretely rebalanced portfolio. Our method uses a more tractable continuously rebalanced portfolio to design the estimator, and our analysis is based on the conditional convergence of the difference between the discrete and continuous cases. The method is asymptotically efficient as we move farther into the tail while increasing the number of rebalancing dates.

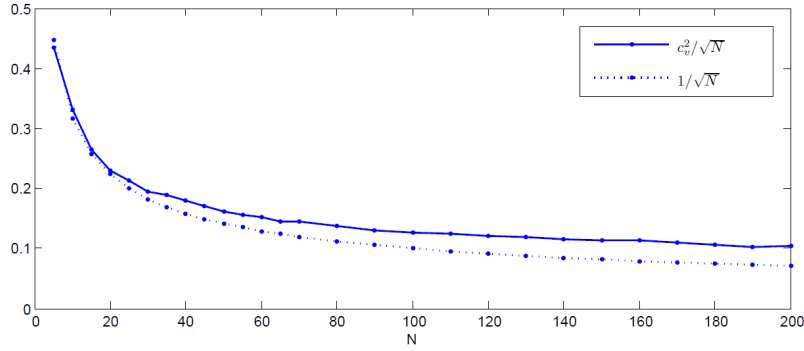


Figure 1: Plot of  $c_v^2/\sqrt{N}$ .

The discussion here is limited to a simple model of the dynamics of the underlying assets and a simple type of portfolio specified by a fixed set of weights. Potential extensions of the approach developed here include relaxing these restrictions.

**ACKNOWLEDGMENTS**

We thank Jose Blanchet for helpful discussions. This work is supported in part by NSF grant DMS0914539.

**A APPENDICES**

**A.1 Proof of Theorem 1**

*Proof.* Since  $\log V(T) = (\mu_w - \sigma_w^2/2)T + \sum_{n=1}^N \bar{\sigma}^\top \Delta W(n)$  is a sum of i.i.d. normal random variables, when conditioned on the sum, each normal increment will have the following distribution (as in, e.g., Theorem 2.5.1 of Anderson (1984)),

$$\Delta W(n) | (\log V(T) = \hat{y}_N) \stackrel{d}{=} \frac{\bar{\sigma}}{\sigma_w^2} x \sqrt{\Delta t} + \Delta \tilde{W}(n) - \frac{\bar{\sigma} \bar{\sigma}^\top \tilde{W}(T)}{\sigma_w^2} \Delta t, \tag{20}$$

where  $\tilde{W}$  is a standard  $d$ -dimensional Brownian motion, and  $\stackrel{d}{=}$  means equality in distribution. Then we have

$$\frac{\hat{V}_N}{V(T)} | (\log V(T) = \hat{y}_N) \stackrel{d}{=} \prod_{n=1}^N \left( \sum_i w_i h_{N,n,i}(x, \sqrt{\Delta t}) \right),$$

where

$$h_{N,n,i}(x, \sqrt{\Delta t}) = \exp\left\{ \frac{(\sigma_i - \bar{\sigma})^\top \bar{\sigma}}{\sigma_w^2} x \sqrt{\Delta t} + (\mu_i - \mu_w - \frac{1}{2}(\|\sigma_i\|^2 - \sigma_w^2))\Delta t + (\sigma_i - \bar{\sigma})^\top Z_{N,n} \sqrt{\Delta t} - \frac{(\sigma_i - \bar{\sigma})^\top \bar{\sigma} \bar{\sigma}^\top \sum_{k=1}^N Z_{N,k}}{\sigma_w^2} \Delta t \right\}$$

for i.i.d.  $Z_{N,n} \sim N(0, I)$ . Then we can obtain a Taylor approximation for each  $h_{N,n,i}(x, \cdot)$  as follows

$$\sum_i w_i h_{N,n,i}(x, \sqrt{\Delta t}) = 1 + Y_{N,n} + X_{N,n} + r_{N,n},$$

where the term of degree one is zero,  $Y_{N,n}$  and  $X_{N,n}$  are the terms of degree two and three, respectively, and  $r_{N,n}$  is the remainder, such that

$$\|Y_{N,n}\| = O(\Delta t), \|X_{N,n}\| = O(\Delta t^{3/2}), \|r_{N,n}\| = O(\Delta t^2).$$



With some algebraic simplification, we obtain an explicit expression for  $Y_{N,n}$  as

$$Y_{N,n} = \left( \frac{1}{2} \sum_i w_i ((\sigma_w^2 - \|\sigma_i\|^2) + [(\sigma_i - \bar{\sigma})Z_{N,n}]^2) + \sum_i w_i (\sigma_i - \bar{\sigma}^\top) Z_{N,n} (\sigma_i - \bar{\sigma})^\top \frac{x}{\sigma_w^2} + a_x \right) \Delta t,$$

where  $a_x$  is defined in (4). Then  $\sqrt{N} \sum_{n=1}^N (Y_{N,n} - a_x/N)$  has mean zero. Also because  $Y_{N,n} - a_x/N$  is a sum of independent copies of polynomials of normal random variables, each of which has  $L_p$ -norm  $O(\Delta t)$  for any  $p \geq 1$ , the Lyapunov Central Limit Theorem in Theorem 7.1.2 of Chung (2000), yields

$$\sqrt{N} \sum_{n=1}^N (Y_{N,n} - \frac{a_x}{N}) \Rightarrow N(0, \sigma'^2), \text{ for some } \sigma' > 0. \tag{21}$$

We can also derive  $X_{N,n}$  with the following form

$$\begin{aligned} X_{N,n} = & \frac{1}{6} g_{N,n}^{(3)}(0) \Delta t^{3/2} + I_{N,n} \Delta t^{3/2} + \frac{1}{6} \sum_i w_i \left( \left[ \frac{(\sigma_i - \bar{\sigma})^\top \bar{\sigma}}{\sigma_w^2} x \right]^3 \Delta t^{3/2} + 3 \left[ \frac{(\sigma_i - \bar{\sigma})^\top \bar{\sigma}}{\sigma_w^2} x \right]^2 (\sigma_i - \bar{\sigma})^\top Z_{N,n} \right. \\ & \left. + \frac{1}{2} \frac{(\sigma_i - \bar{\sigma})^\top \bar{\sigma}}{\sigma_w^2} x [(\sigma_i - \bar{\sigma})^\top Z_{N,n}]^2 + \frac{2}{3} \frac{(\sigma_i - \bar{\sigma})^\top \bar{\sigma}}{\sigma_w^2} x (\mu_w - \mu_w - \frac{1}{2} (\|\sigma_i\|^2 - \sigma_w^2)) \right) \Delta t^{3/2}, \end{aligned}$$

where

$$g_{N,n}(\sqrt{\Delta t}) = \sum_i w_i \exp\{(\mu_i - \mu_w)\Delta t - \frac{1}{2}(\|\sigma_i\|^2 - \sigma_w^2)\Delta t + (\sigma_i - \bar{\sigma})^\top Z_{N,n} \sqrt{\Delta t}\} - 1,$$

and

$$I_{N,n} = -\frac{2}{9} \sum_i w_i \frac{\|\sigma_i - \bar{\sigma}\|^2 \bar{\sigma}^\top \sum_{k=1}^N Z_{N,k} / \sqrt{N}}{\sigma_w^2} \Delta t^{3/2}.$$

Here  $g_{N,n}(\cdot)$  is the same as defined in Glasserman (2009), and  $g_{N,n}^{(3)}(0)$  consists of odd powers of normal distributions, and does not depend on  $x$ . So we can divide  $X_{N,n}$  into two parts. The first part is  $(\frac{1}{6} g_{N,n}^{(3)}(0) + I_{N,n}) \Delta t^{3/2}$ , which does not contain  $x$  and has mean zero and  $L_p$ -norm of order  $O(\Delta t^{3/2})$  for any  $p > 1$ . The second part has mean  $b_x \Delta t^{3/2}$  as defined in (5), and  $L_p$ -norm of order  $O(\Delta t^{3/2})$  for any  $p > 1$ . So for the same reason as for (21), we have

$$\sqrt{N} \sum_{n=1}^N X_{N,n} \Rightarrow N(b_x, \sigma''^2), \text{ for some } \sigma'' > 0. \tag{22}$$

Based on the fact that convergence to zero in  $L_p$  (here we can easily prove for  $p = 1$ ) implies convergence to zero in distribution, we have

$$\sqrt{N} \sum_{n=1}^N (Y_{N,n} - a_x \Delta t)^2, \sqrt{N} \sum_{n=1}^N r_{N,n}, \sqrt{N} \sum_{n=1}^N r_{N,n}^2, \sqrt{N} \sum_{n=1}^N X_{N,n}^2, \tag{23}$$

$$\sqrt{N} \sum_{n=1}^N (Y_{N,n} - a_x \Delta t) r_{N,n}, \sqrt{N} \sum_{n=1}^N X_{N,n} r_{N,n}, \sqrt{N} \sum_{n=1}^N X_{N,n} (Y_{N,n} - a_x \Delta t) \tag{24}$$

all converge to zero in probability. Thus,  $\sqrt{N} \sum_{n=1}^N (Y_{N,n} - a_x/N + X_{N,n} + r_{N,n})$  converges in distribution to a normal distribution, and  $\sqrt{N} \sum_{n=1}^N (Y_{N,n} - a_x/N + X_{N,n} + r_{N,n})^2$  converges to zero in probability. And because for sufficiently small  $v$ ,  $\exp\{v - v^2\} \leq 1 + v \leq \exp\{v + v^2\}$ , we can show that

$$\prod_{n=1}^N \left( \sum_i w_i h_{N,n,i}(x, \sqrt{\Delta t}) \right)^{\sqrt{N}} \leq \exp \left\{ \sqrt{N} \sum_{n=1}^N (1 + Y_{N,n} + X_{N,n} + r_{N,n}) + \sqrt{N} \sum_{n=1}^N (1 + Y_{N,n} + X_{N,n} + r_{N,n})^2 \right\}$$

and

$$\prod_{n=1}^N \left( \sum_i w_i h_{N,n,i}(x, \sqrt{\Delta t}) \right)^{\sqrt{N}} \geq \exp \left\{ \sqrt{N} \sum_{n=1}^N (1 + Y_{N,n} + X_{N,n} + r_{N,n}) + \sqrt{N} \sum_{n=1}^N (1 + Y_{N,n} + X_{N,n} + r_{N,n})^2 \right\}$$

So using (21), (22) and convergence of terms in (24), we have

$$\left( \frac{\hat{V}_N}{V(T)} \right)^{\sqrt{N}} \exp\{-a_x \sqrt{N}\} |(\log V(T) = \hat{y}_N) \Rightarrow \exp\{N(b_x, \tilde{\sigma}^2)\}, \text{ for some } \tilde{\sigma} > 0.$$

By Theorem 2.19 in van der Vaart (1998), we can, through a change in probability space, replace convergence in distribution with almost sure convergence. Since the limit  $\exp\{N(b_x, \tilde{\sigma}^2)\}$  is non-negative, we can conclude that the result (3) holds.  $\square$

### A.2 Bounds for the Normal Distribution

For  $z > 0$ , any  $\delta > 0$  and choose  $\theta = z$ , for  $Z \sim N(0, 1)$ , there exists some constant  $C > 0$  such that

$$\begin{aligned} \Phi(-z) &= 1 - \Phi(z) = E_\theta [I_{Z>z} \exp\{-\theta Z + \Psi_z(\theta)\}] \\ &\geq E_\theta [I_{z < Z < z+\delta} \exp\{-\theta Z + \Psi_z(\theta)\}] \\ &\geq P_\theta (z < Z < z + \delta) \exp\{-\theta(z + \delta) + \Psi_z(\theta)\} \\ &\geq C \exp\{-\frac{z^2}{2} - z\delta\}, \end{aligned} \tag{25}$$

where  $\Phi_z(\theta) = \log E[\exp\{\theta Z\}]$ . By choosing  $\theta = x$ , we have get

$$\begin{aligned} \Phi(-z) &= 1 - \Phi(z) = E_\theta [I_{Z>z} \exp\{-\theta Z + \Psi_Z(\theta)\}] \\ &\leq \exp\{-\theta z + \Psi_Z(\theta)\} = \exp\{-z^2/2\}. \end{aligned} \tag{26}$$

By applying L'Hopital's rule, we can easily get

$$\lim_{z \rightarrow \infty} \frac{\Phi(-z)}{e^{-z^2}/z} = \frac{1}{\sqrt{2\pi}}$$

leading to a well-known asymptotic for the normal tail distribution

$$\Phi(-z) = 1 - \Phi(z) \sim \frac{\phi(z)}{z}, \text{ as } z \rightarrow \infty. \tag{27}$$

### REFERENCES

Anderson, T. 1984. *An introduction to multivariate statistical analysis*. Wiley, New York.

Asmussen, S., and P. W. Glynn. 2007. *Stochastic simulation*. Springer.

Bertsimas, D., L. Kogan, and A. Lo. 2000. When is time continuous? *Journal of Financial Economics* 55:173–204.

Boyle, P., and D. Emanuel. 1980. Discretely adjusted option hedges. *Journal of Financial Economics* 8:259–282.

Chung, K. 2000. *A course in probability theory, 2nd ed.* Academic Press.

Duffie, D., and T.-S. Sun. 1990. Transactions costs and portfolio choice in a discretecontinuous- time setting. *Journal of Economic Dynamics and Control* 14:35C51.

Glasserman, P. 2009. Risk horizon and rebalancing horizon in portfolio risk measurement. *Mathematical Finance* to appear.

Glasserman, P., P. Heidelberger, and P. Shahabuddin. 2000. Variance reduction techniques for estimating value-at-risk. *Management Science* 46:1349–1364.

Glasserman, P., and J. Li. 2003. Importance sampling for a mixed poisson model of portfolio credit risk. *Proceedings of the Winter Simulation Conference*.

Jacod, J., and P. Protter. 1998. Asymptotic error distributions for the euler method for stochastic differential equations. *Annals of Probability* 26:267–307.

Leland, H. 1985. Option pricing and replication with transactions costs. *Journal of Finance* 40:1283–1301.

- Morton, A., and S. Pliska. 1995. Optimal portfolio management with fixed transaction costs. *Mathematical Finance* 5:337C356.
- Tankov, P., and E. Voltchkova. 2009. Asymptotic analysis of hedging errors in models with jumps. *Stochastic Processes and their Applications* 119:2004–2027.
- van der Vaart, A. W. 1998. *Asymptotic statistics*. New York: Cambridge University Press.

#### **AUTHOR BIOGRAPHIES**

**PAUL GLASSERMAN** is the Jack R. Anderson Professor at Columbia Business School. His research focuses on modeling and computational problems in derivatives pricing and risk management. His book, *Monte Carlo Methods in Financial Engineering* (Springer, 2004), received the 2005 Outstanding Simulation Publication Award and the 2006 Lanchester Prize.

**XINGBO XU** is a Ph.D. candidate in the Industrial Engineering and Operations Research in Columbia University, New York, NY. His main focus of interest has been the application of simulation methods in finance. He received his B.S. in mathematics and applied mathematics from Zhejiang University in 2006 and M.S. in mathematics from Purdue University in 2008. His email address is <[xx2126@columbia.edu](mailto:xx2126@columbia.edu)>.