Asia-Pacific Journal of Financial Studies (2009) v38 n5 pp745-772

An Importance Sampling Method to Evaluate Value-at-Risk for Assets with Jump Risks*

Ren-Her Wang**

National Taiwan University, Taipei, Taiwan, R.O.C.

Shih-Kuei Lin

National University of Kaohsiung, Kaohsiung, Taiwan, R.O.C.

Cheng-Der Fuh

National Central University, Jhong-Li and Academia Sinica, Taipei, Taiwan, R.O.C.

Received 22 August 2008; Accepted 18 May 2009

Abstract

Risk management is an important issue when there is a catastrophic event that affects asset price in the market such as a sub-prime financial crisis or other financial crisis. By adding a jump term in the geometric Brownian motion, the jump diffusion model can be used to describe abnormal changes in asset prices when there is a serious event in the market. In this paper, we propose an importance sampling algorithm to compute the Value-at-Risk for linear and nonlinear assets under a multi-variate jump diffusion model. To be more precise, an efficient computational procedure is developed for estimating the portfolio loss probability for linear and nonlinear assets with jump risks. And the titling measure can be separated for the diffusion and the jump part under the assumption of independence. The simulation results show that the efficiency of importance sampling improves over the naive Monte Carlo simulation from 7 times to 285 times under various situations. We also show the robustness of the importance sampling algorithm by comparing it with the EVT-Copula method proposed by Oh and Moon (2006).

Keywords: Jump Diffusion Models; Value-at-Risk; Quick Simulation; Importance Sampling; Risk Management

JEL Classification: C15, C63, G32

---

* The research of Shih-Kuei Lin was partially supported by the National Science Council under grant NSC 94-2416-H-259-031, and partially supported by Academia Sinica. The research of Cheng-Der Fuh was partially supported by the National Science Council under grant NSC 96-2118-M-001-002.
We thank anonymous referees for providing many useful comments on an earlier draft.

** Corresponding Author. Address: Department of Finance, National Taiwan University, Taipei, No. 1, Sec. 4, Roosevelt Rd., Taipei City 106, Taiwan, R.O.C.; E-mail: d92723009@ntu.edu.tw; Tel: 886-2-2783-5611 ext 147; Fax: 886-2-2783-1523.

1. Introduction

The essence a financial institution lies in managing risks. The trader manages a normal event risk where the world operates in a fashion similar to the Black-Scholes model of random walks and dynamic hedging. Intuitively, we are interested in tail events in which large jumps are important. Value-at-Risk (VaR) is a measure of the potential losses due to movement in the underlying market. VaR usually is associated with a time frame and an estimate of the maximum sudden change in thought likely in the markets (unclear). In this paper, we propose a new method to compute a portfolio's VaR under a multi-variate jump-diffusion model.

Although Black and Scholes (1973) and Merton (1973) gave a closed formula for pricing options, and the formula has been used to approximate the price of financial products in practice, the assumptions of the model in the Black-Scholes-Merton world are not applicable to the real market situation. To capture the empirical phenomenon, Merton (1976) proposed a jump-diffusion model to describe discontinuous change of the asset price when abnormal information arrives in the market. At important events or announcements, there can be large changes in the value of financial portfolios. Events and their corresponding jumps can occur at random or scheduled times. Therefore, the amplitude of the response in either case can be unpredictable or random. While the volatility of portfolios is often modeled by continuous Brownian motion processes, discontinuous jump processes are more appropriate for modeling the response to important external events that significantly affect the prices of financial assets. Discrete jump processes are modeled by compound Poisson processes for random events or scheduled events.

The jump diffusion model for the price of an underlying asset (for example a stock or a stock index) is assumed to include two parts, a continuous part described by a geometric Brownian motion, and a discontinuous jump part described by a compound Poisson process. In the compound Poisson process, there are two components: one is jump sizes specified by the ratios of the sudden change of the underlying asset price in abnormal events, the other is jump frequencies specified by the arrival rate of abnormal events. The logarithms of the jump sizes have a special distribution such as a normal distribution (cf. Merton, 1976) or a double exponential distribution (cf. Kou, 2002), and the jump times correspond to a Poisson process. More precisely, the following stochastic differential equation is used to model the asset price, $S(t)$:

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t) + d\sum_{j=1}^{N(t)} (Y(j) - 1), \tag{1}$$

where the parameters of the drift and the volatility are denoted as μ and σ, $W(t)$ is a standard Wiener process, $N(t)$ is a Poisson process with the arrival rate λ, and $\{Y(j)\}$ is a sequence of independent identically distributed nonnegative random variables such that $\log(Y(j))$ follows a normal distribution.

Monte Carlo analysis is a powerful method to evaluate a portfolio's VaR, as it can be used to calculate the distribution of the portfolio returns. However, the biggest drawback of this method is that it is computationally demanding when there are hundreds of securities (cf. Jorion, 2000; Hull, 2000). There are variate techniques to reduce the variance of Monte Carlo simulation. The importance sampling can reduce the sample size to save computational time with some accuracy. The main idea behind the importance sampling is the change of measure. That is, we carry out a simulation with an importance tilting measure instead of the original probability measure, to obtain the accurate estimator for the portfolio VaR. Many investigations have proposed using importance sampling to improve the efficiency of evaluating the portfolio VaR or quantile. (cf. Johns, 1988; Goffinet and Wallach, 1996; Glasserman, Heidelberger and Shahabuddin, 2000, 2002; Glasserman, 2003; Fuh and Hu, 2004; Lin, Wang and Fuh, 2006). Glasserman, Heidelberger and Shahabuddin (2000, 2002) consider the idea of combining importance testing with stratified sampling for further reduction in variance. Lin, Wang and Fuh (2006) propose the bootstrap method and Laplace method to improve the tilting distribution of Glasserman, Heidelberger and Shahabuddin (2000, 2002).

The jump diffusion model can be used to describe huge changes of the asset prices in the markets due to unexpected events such as a sub-prime financial crisis or the slashing of interest rates by three-quarters of a point. Recently, Oh and Moon (2006) apply Extreme Value Theory (EVT) to fit fat-tailed marginal distributions, to which it is called the EVT-Copula method. They show that the EVT-Copula method outperforms the traditional portfolio VaR method during the period of Financial Crises. In this paper, we develop importance sampling for a multi-variate jump diffusion model for a linear asset and nonlinear asset to evaluate VaR. Efficient algorithms are proposed for estimating the portfolio loss probability for the linear asset and the nonlinear asset with jump risks. By assuming independence, we can separate the tilting measure from the diffusion and jump part. The simulation results show that the efficiency of importance sampling is improved over the naive Monte Carlo simulation from 7 times to 285 times under various situations.

The rest of this paper is organized as follows. Section 2 introduces the general case for multi-variate jump diffusion models for a linear asset and a non-linear asset. Section 3 interprets the change of measure used for importance sampling for multi-variate jump diffusion models. Numerical results are reported in Section 4. Section 5 concludes. The technical details are deferred to the appendix.

2. Multi-variate Jump Diffusion Models

Consider the process of the i th asset to be described by the following stochastic differential equation:

$$\frac{dS_i(t)}{S_i(t)} = \mu_i dt + \sigma_i dW_i(t) + d\sum_{j=1}^{N(t)} (Y_i(j) - 1), \tag{2}$$

where the parameters of the drift and the volatility are denoted as μ_i and σ_i, the jump event $N(t)$ is assumed to follow a Poisson process with the parameter λt, and the jump sizes $Y_i(j)$ are assumed to follow a lognormal distribution with the parameters of location η_i and scale ν_i^2. The discrete return of the jump diffusion model (cf. Merton, 1976; Kou, 2002) for the i th asset is

$$\frac{\Delta S_i(t)}{S_i(t)} = \frac{S_i(t+\Delta t)}{S_i(t)} - 1 \approx \mu_i \Delta t + \sigma_i W_i(\Delta t) + \sum_{j=1}^{N(\Delta t)} \log(Y_i(j)). \tag{3}$$

By the normal distributions of Brownian motion and jump sizes, the return $r_i(t) \equiv \frac{\Delta S_i(t)}{S_i(t)}$ of the i th asset for a period $[t,\ t+\Delta t]$ is denoted as

$$r_i(t) = \mu_i \Delta t + \sigma_i \sqrt{\Delta t} Z_i + \sum_{j=1}^{N(\Delta t)} V_i(j), \tag{4}$$

where Z_i is the standard normal distribution with mean 0, variance 1 for the i th asset, and correlation coefficient ρ_{ik} for the i th asset and the k th asset. The j th jump size of the i th asset $V_i(j)=\log(Y_i(j))$ follows a normal distribution with the mean η_i and the variance v_i^2. The correlation coefficient of the jump sizes of the i th and j th asset are assumed as ρ_{ij}^J. For simplicity, the notations $\mu_i\Delta t$, $\sigma_i\sqrt{\Delta t}$, and $N(\Delta t)$ with the parameter $\lambda\Delta t$ are rewritten as μ_i, σ_i, and N with the arrival rate λ. The return can be presented as

$$r_i(t)=\mu_i+\sigma_i Z_i+\sum_{j=1}^{N} V_i(j). \tag{5}$$

2.1 Return of a linear asset

Assume we have a linear asset, which is the linear combination of d investment assets. The vector of the returns can be denoted as

$$\tilde{r}(t)=\begin{bmatrix} r_1(t) \\ \vdots \\ r_d(t) \end{bmatrix}=\begin{bmatrix} \mu_1+\sigma_1 Z_1+\sum_{j=1}^{N} V_1(j) \\ \vdots \\ \mu_d+\sigma_d Z_d+\sum_{j=1}^{N} V_d(j) \end{bmatrix}=\tilde{\mu}+\Sigma_\sigma\tilde{Z}+\tilde{J},$$

where $\tilde{r}(t)=[r_1(t),\cdots,r_d(t)]^T$ is a vector of the discrete return of the assets, $\tilde{\mu}=[\mu_1(t),\cdots,$ $\mu_d(t)]^T$ denotes a vector of the discrete drift, $\tilde{Z}$ presents the multi-variate normal distribution with zero mean and the correlation matrix:

$$\Sigma_B=\begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1d} \\ \rho_{21} & 1 & \cdots & \rho_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{d1} & \cdots & \cdots & 1 \end{bmatrix},$$

in which Σ_σ is the diagonal matrix with the diagonal elements of standard deviations as follows

$$\Sigma_\sigma=\begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \sigma_d \end{bmatrix},$$

$\tilde{J}$ denotes a vector of the jump sizes $\tilde{J} = [\sum_{j=1}^{N} V_1(j), \cdots, \sum_{j=1}^{N} V_d(j)]^T = \sum_{j=1}^{N} \tilde{V}(j)$ where $\tilde{V}(j)$ denotes the jump size vector of the assets when the abnormal jump j th happens, and the arrival rate of the assets are the same with λ, that is to say that the abnormal event, such as the catastrophic fall of stock prices, makes all the assets have abnormal jumps, and the jump sizes have a multi-normal distribution with the mean vector $\tilde{\eta} = [\eta_1, \cdots, \eta_d]^T$ and the covariance matrix

$$\Sigma_J = \begin{bmatrix} \nu_1^2 & \nu_1\nu_2\rho_{12}^J & \cdots & \nu_1\nu_d\rho_{1d}^J \\ \nu_1\nu_2\rho_{12}^J & \nu_2^2 & \cdots & \nu_2\nu_d\rho_{2d}^J \\ \vdots & \vdots & \ddots & \vdots \\ \nu_1\nu_d\rho_{1d}^J & \cdots & \cdots & \nu_d^2 \end{bmatrix}.$$

Here, Σ_J is the matrix of the correlation coefficient for the jump sizes.

Let $\tilde{w}(t) = [w_1(t), \cdots, w_d(t)]^T$ denote a portfolio weight vector of the investment assets for the portfolio value, in which $w_i(t)$ is an adapted process, i.e., F_t-measurable. Then the return of the portfolio at time t is the linear combination of the asset returns multiplied by the portfolio weight vector denoted as $\tilde{w}^T(t)\tilde{r}(t)$, and we are interested in the event $A = \{\tilde{r}(t) : f(\tilde{r}(t)) = \tilde{w}^T\tilde{r}(t) - r_p > 0\}$, where r_p is the biggest risk with which we are concerned for the linear asset (also called the quantile). Hence, what we are interested in is the probability of the return of the linear assets $\tilde{r}(t)$ being greater than the quantile of the return r_p when the return vector of the assets is assumed as a multi-variate jump diffusion model. The linear portfolio is satisfactory for small movements in the underlying asset. A better approximation may be achieved by using higher order terms and incorporating the gamma or convexity effect. The event, which we are interested in, can be rewritten as

$$f(\tilde{r}(t)) = \tilde{w}^T\tilde{r}(t) - r_p = \tilde{w}^T\tilde{\mu} + \tilde{w}_\sigma^T\tilde{Z} + \tilde{w}^T\tilde{J} - r_p.$$

According to the Cholesky decomposition for Σ_B, we can find the equal distribution as follows:

$$f(\tilde{r}(t)) = \tilde{w}^T\tilde{\mu} + \tilde{w}_\sigma^T C\tilde{U} + \tilde{w}^T\tilde{J} - r_p,$$

where C is used by the Cholesky decomposition for the covariance of the assets Σ_B such that $CC^T = \Sigma_B$, $\tilde{U} = [U_1, U_2, \cdots, U_d]^T$ is a vector of d-variate independent standard normal distribution, and $\tilde{w}_\sigma = [w_1\sigma_1, \cdots, w_d\sigma_d]^T$.

2.2 Return of a nonlinear asset

Assume that there is a nonlinear asset $P(t, \tilde{S}(t))$ such as derivatives. In a multi-variate jump diffusion model, the quadratic approximation of the nonlinear asset in the risk factors (cf. Glasserman, Heidelberger, and Shahabuddin, 2000, 2002; Lin, Wang, and Fuh, 2006) can be changed into the function of the return of the stocks in the nonlinear asset, $a(t) + \tilde{\delta}_N^T \tilde{r}(t) + \tilde{r}^T(t)\Gamma_N \tilde{r}(t)$, (proved in Appendix A), and the interesting event is $A^N = \{\tilde{r}(t) : f_N(\tilde{r}(t)) = a(t) + \tilde{\delta}_N^T \tilde{r}(t) + \tilde{r}^T(t)\Gamma_N \tilde{r}(t) - r_p^N > 0\}$. Hence, we consider the function of the non-linear asset return denoted as

$$f_N(\tilde{r}(t)) = a(t) + \tilde{\delta}_N^T \tilde{r}(t) + \tilde{r}^T(t)\Gamma_N \tilde{r}(t) - r_p^N,$$

where $a(t) = \frac{\frac{\partial P}{\partial t}\Delta t}{P(t,\tilde{S}(t))}$ is the change of the nonlinear value relative to the nonlinear asset by the time change, $\tilde{\delta}_N = [\frac{S_1\delta_1}{P(t,\tilde{S}(t))}, \cdots, \frac{S_n\delta_n}{P(t,\tilde{S}(t))}]^T$ denotes the delta approximation vector relative to the asset weight of the nonlinear value, and $(\Gamma_N)_{ij} = \frac{1}{2}(\Gamma)_{ij}\frac{S_iS_j}{P(t,\tilde{S}(t))}$ presents the gamma approximation matrix of the weight of the product of the ith and jth assets relative to the value of the nonlinear asset at time t.

By the Cholesky decomposition, $\tilde{Z} \overset{D}{=} C\tilde{U}$, where $\tilde{U} = [U_1, \cdots, U_d]^T$ has an independent multi-variate normal distribution with zero mean vector and the identity covariance matrix I, and D denotes "equal in distribution." We can find C such that $C^T\Gamma_N C = \Lambda$ and $CC^T = \Sigma_B$, where Λ is a diagonal matrix with λ_i, where $i = 1, \cdots, d$.

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_d \end{bmatrix}.$$

Therefore, the return of the nonlinear asset by the quadratic approximation and the

Cholesky decomposition method is equal to the distribution (proven in Appendix B):

$$f_N(\tilde{r}(t)) \overset{D}{=} a_N + \sum_{i=1}^{d}(b_i U_i + \lambda_i \sigma_i^2 U_i^2) + \tilde{\delta}_N^T \tilde{J}, \tag{6}$$

where $\tilde{b} = [b_1, b_2, \cdots, b_d]^T = \tilde{\delta}_N^T \Sigma_\sigma C,$ and $a_N = a(t) + \tilde{\delta}_N^T \tilde{\mu} - r_p^N.$

3. Importance Sampling Algorithms

We consider the problem of estimating small probabilities by Monte Carlo simulations, where such problems appear in the construction of confidence regions for asymptotically normal statistics concerning VaR for risk management (cf. Beran, 1987; Beran and Millar, 1986; Hall, 1987, 1992; Jorion, 2001; Fuh and Hu, 2004). It is well-known that the importance sampling, in which one uses observations from an alternative distribution to estimate the target distribution, is a useful tool for efficient simulation of events with small probabilities. Efficient Monte Carlo simulation of such events has been obtained by Sadowsky and Bucklew (1990) based on the large deviations theory. For events of large deviations, previous authors have showed that the asymptotically optimal alternative distribution is obtained through exponential tilting to determine the parameter of tilting. The optimal parameter of tilting is such that the mean of estimation for tail events is unbiased and the second moment of the estimation is asymptotically efficient. The subsections below focus on the importance sampling algorithms for linear and nonlinear assets for Monte Carlo simulations.

3.1 An algorithm for linear assets

When searching for a convenient importance distribution, particularly if we wish to increase or decrease the frequency of observations in the tails, it is quite common to embed a given density in an exponential family. From this probability density function, we can now produce a whole (exponential) family of densities:

$$dP^\theta = e^{\theta f(\tilde{r}(t)) - \Psi(\theta)} dP, \tag{7}$$

where dP is the original probability measure for the linear asset, and dP^θ is the tilting measure for the linear asset. Suppose $\Psi(\cdot)$ denotes the cumulate generating function (the logarithm of the moment generating function) of the function $f(\tilde{r}(t))$. The cumulate generating function is a useful summary of the moments of a distribution since the mean can be determined as $\Psi'(0)$ and the variance as $\Psi''(0)$. Therefore, we compute the moment generating function of the interesting event:

$$\begin{aligned} &E(\exp\{\theta f(\tilde{r}(t))\}) \\ &= \underbrace{\exp\{\theta\tilde{w}^T\tilde{\mu}+\frac{1}{2}\theta^2\tilde{w}_\sigma^T\Sigma_B\tilde{w}_\sigma-\theta r_p\}}_{\text{(diffusion part)}}\underbrace{\exp\{\lambda(\exp\{\theta\tilde{w}^T\tilde{\eta}+\frac{1}{2}\theta^2\tilde{w}^T\Sigma_J\tilde{w}\}-1)\}}_{\text{(jump part)}}, \end{aligned}$$

which is separated into a diffusion part and a jump part, both coming from the event of interest. Then, we obtain $\Psi(\theta)$ as

$$\begin{aligned} \Psi(\theta) &= \log E(\exp\{\theta f(\tilde{r}(t))\}) \\ &= \underbrace{(\theta\tilde{w}^T\tilde{\mu}+\frac{1}{2}\theta^2\tilde{w}_\sigma^T\Sigma_B\tilde{w}_\sigma-\theta r_p)}_{\text{(diffusion part)}}+\underbrace{(\lambda(\exp\{\theta\tilde{w}^T\tilde{\eta}+\frac{1}{2}\theta^2\tilde{w}^T\Sigma_J\tilde{w}\}-1))}_{\text{(jump part)}}. \end{aligned}$$

By using the same idea as that in Glasserman, Heidelberger and Shahabuddin (1999), we also obtain the tilting parameter θ_p. Let $\tilde{r}_\theta(t)$ be drawn from the tilting measure dP^θ, and $A_\theta=\{\tilde{r}_\theta(t): f(\tilde{r}_\theta(t))=\tilde{w}'(t)\tilde{r}_\theta(t)-r_p>0\}$ be the event of interest. The estimation of the tail probability is denoted as

$$\hat{p}_\theta = 1_{A_\theta}\exp\{-\theta f(\tilde{r}_\theta(t))+\Psi(\theta)\}, \tag{8}$$

then the estimator

$$E(1_A)=E^\theta(1_{A_\theta}\exp\{-\theta f(\tilde{r}_\theta(t))+\Psi(\theta)\})=E^\theta(\hat{p}_\theta)=p$$

is unbiased. Therefore, the second moment of the estimator for the tail probability of the interesting event is approximated as

$$\begin{aligned} M_2(\theta) &= E^\theta(1_{A_\theta}\exp\{2\Psi(\theta)-2\theta f(\tilde{r}_\theta(t))\}) \\ &= E(1_{A_\theta}\exp\{\Psi(\theta)-\theta f(\tilde{r}_\theta(t))\}) \\ &\le \exp\{\Psi(\theta)\}, \end{aligned} \tag{9}$$

because $\exp\{-\theta f(\tilde{r}_\theta(t))\}\le 1$ ($f(\tilde{r}_\theta(t))>0$ and $\theta>0$), and $1_{A_\theta(t)}\le 1$. Taking the *log* of the bound equation and differentiating θ, we can minimize the upper bound of the second moment of the estimator for the tail probability of the interesting event.

Using a numerical method to find θ_p such that $\Psi'(\theta_p)=0$, we observe that $\Psi(\theta)$ can be divided into two parts. The first part is from the diffusion process, the other part is from the jump process. If we set $\lambda=0$, the reflection of the jump part will vanish. The case will degenerate to a portfolio of multi-variate normal distribution, which can be found in Lin, Wang and Fuh (2006).

Let $dP^{\theta}=dP_{\alpha,\beta,\gamma}$ denote the new measure for the given n jump time with a multi-variate normal distribution in new parameter α, a jump size in new parameter β, and a jump rate in new parameter γ, and the original measure with given jump time n:

$$dP=dP_{0,0,0}=\phi_d(\tilde{U};\tilde{0},I)\times\prod_{j=1}^{n}\phi_d(\tilde{V}(j);\tilde{\eta},\Sigma_J)\times \mathrm{P}(\lambda) \tag{10}$$

in which $\phi_d(\tilde{U};\tilde{0},I)$ is a probability density function of the d-variate independent normal distribution with mean vector $\tilde{0}$ and covariance matrix I, $\phi_d(\tilde{V}(j);\tilde{\eta},\Sigma_J)$ is a probability density function of the multi-variate independent normal distribution with mean vector $\tilde{0}$ and covariance matrix Σ_J for the j th time. $\mathrm{P}(\lambda)$ is the probability density function of a Poisson distribution with arrival rate λ.

First, based on equation (7), the new measure of the diffusion part with parameter α is as follows (derived in Appendix C):

$$\begin{aligned} dP_{\alpha,0,0} &= \exp\{\theta f(\tilde{r}(t))-\Psi(\theta)\}\times\phi_d(\tilde{U};\tilde{0},I)\times\prod_{j=1}^{n}\phi_d(\tilde{V}(j);\tilde{\eta},\Sigma_J)\times \mathrm{P}(\lambda) \\ &= \exp\{\theta_p\tilde{w}^T\tilde{\mu}+\frac{1}{2}\theta_p^2\tilde{w}_\sigma^T\Sigma_B\tilde{w}_\sigma+\theta_p\tilde{w}^T\tilde{J}-\theta_p r_p-\Psi(\theta_p)\} \\ &\quad\times\ \phi_d(\tilde{U}^*;\tilde{u},I)\times\prod_{j=1}^{n}\phi_d(\tilde{V}(j);\tilde{\eta},\Sigma_J)\times \mathrm{P}(\lambda), \end{aligned}$$

where $dP_{\alpha,0,0}$ denotes the new measure of the multi-variate normal distribution with the diffusion part and the original measures of jump frequencies and jump sizes with the jump part. $\phi_d(\tilde{U}^*;\tilde{u},I)$ is a multi-variate independent normal with mean vector $\tilde{u}=[\theta_p w_1\sigma_1\sum_i c_{1i},\cdots,\theta_p w_d\sigma_d\sum_i c_{di}]^T$ and covariance I. Second, the new measures of multi-variate normal distributions with the diffusion part and the jump sizes part are given by

$$
\begin{aligned}
dP_{\alpha,\beta,0} = {} & \exp\{\theta_p \tilde{w}^T\tilde{\mu} + \frac{1}{2}\theta_p^2 \tilde{w}_\sigma^T \Sigma_B \tilde{w}_\sigma + \theta_p \tilde{w}^T (\sum_{j=1}^{n} \tilde{V}(j) - \tilde{\eta}) + \theta_p \tilde{w}^T \tilde{\eta} n - \theta_p r_p - \Psi(\theta_p)\} \\
& \times \phi_d(\tilde{U}^*;\ \tilde{u},\ I) \times \prod_{j=1}^{n} \phi_d(\tilde{V}(j);\ \tilde{\eta},\ \Sigma_J) \times \mathrm{P}(\lambda) \\
= {} & \exp\{\theta_p \tilde{w}^T\tilde{\mu} + \frac{1}{2}\theta_p^2 \tilde{w}_\sigma^T \Sigma_B \tilde{w}_\sigma + \theta_p \tilde{w}^T \tilde{\eta} n + \frac{1}{2}\theta_p^2 \tilde{w}^T \Sigma_J \tilde{w} n - \theta_p r_p - \Psi(\theta_p)\} \\
& \times \phi_d(\tilde{U}^*;\ \tilde{u},\ I) \times \prod_{j=1}^{n} \phi_d(\tilde{V}^*(j);\ \tilde{\eta}^*,\ \Sigma_J) \times \mathrm{P}(\lambda),
\end{aligned}
$$

where $dP_{\alpha,\beta,0}$ denotes the new measures of the d-variate normal distribution conditional with the diffusion part, the multi-variate normal distribution with the jump sizes and the origin measures with the jump rate. $\phi_d(\tilde{V}^*;\ \tilde{\eta}^*, \Sigma_J)$ is a d-variate independent normal with mean vector $\tilde{\eta}^* = \tilde{\eta} + \theta_p$ $\tilde{w}^T\Sigma_J$ and covariance Σ_J. Finally, the new measures of the diffusion and jump parts are presented as follow:

$$
\begin{aligned}
dP_{\alpha,\beta,\gamma} &= \phi_d(\tilde{U}^*; \tilde{u}, I) \times \prod_{j=1}^{n} \phi_d(\tilde{V}^*(j); \tilde{\eta}^*, \Sigma_J) \\
& \quad \times \frac{\exp\{-\lambda \exp\{\theta_p \tilde{w}^T \tilde{\eta} + \frac{1}{2}\theta_p^2 \tilde{w}^T \Sigma_J \tilde{w}\}\}(\lambda \exp\{\theta_p \tilde{w}^T \tilde{\eta} + \frac{1}{2}\theta_p^2 \tilde{w}^T \Sigma_J \tilde{w}\})^n}{n!}, \\
&= \phi_d(\tilde{U}^*;\ \tilde{u}, I) \times \prod_{j=1}^{n} \phi_d(\tilde{V}^*(j);\ \tilde{\eta}^*, \Sigma_J) \times \mathrm{P}(\lambda^*)
\end{aligned}
$$

where $dP_{\alpha,\beta,\gamma}$ is the new measure with multivariate independent normal distribution of the random variable $\tilde{U}^*$, multivariate normal distribution of the random variable $\tilde{V}^*(j)$ for given j th jump size, and Poisson distribution of the random variable n with notation $\mathrm{P}(\lambda^*)$ for the parameter $\lambda^* = \lambda \exp\{\theta_p \tilde{w}^T \tilde{\eta} + \frac{1}{2}\theta_p^2 \tilde{w}^T \Sigma_J \tilde{w}\}$ (derived in Appendix C).

If λ decreases, the jump effect will decrease and θ will increase. The density, equation (7), is often referred to as an exponential tilt of the original density function and increase the weight in the right tail for $\theta > 0$, and decreases it for $\theta < 0$. The events of interest is that the return of the portfolio is greater than the quantile. Therefore, when $\theta_p \tilde{w}^T \tilde{\mu} + \frac{1}{2}\theta_p^2 \tilde{w}_\sigma^T \Sigma_B \tilde{w}_\sigma > 0$, jump frequency is also increased to improve the jump effect in rare events. This is because $\theta_p \tilde{w}^T \tilde{\mu} + \frac{1}{2}\theta_p^2 \tilde{w}_\sigma^T \Sigma_B \tilde{w}_\sigma > 0,$ then

$$\lambda \exp\{\theta_p \tilde{w}^T \tilde{\eta} + \frac{1}{2}\theta_p^2 \tilde{w}^T \Sigma_J \tilde{w}\} > \lambda.$$

Therefore, the simulation algorithm for estimating the tail probability $\hat{p}(\theta_p),$ and the standard error of $se(\theta_p)$ from the observed data with the return vector of the assets $\tilde{r}_\theta$ can be described as follows:

Algorithm 1

(1) Compute the tilting point θ_p such that $\Psi'(\theta_p) = 0.$

(2) Generate the return of all d assets with jump risks from three new samples of the tilting distribution.

1. Generate a sample for the d-variate independent normal distribution

$$\tilde{U}^* \sim N_d([\theta_p w_1 \sigma_1 \sum_i c_{1i}, \cdots, \theta_p w_d \sigma_d \sum_i c_{di}]^T, I),$$

where N_d is the d-dimension normal distribution.

2. Generate $N^* \sim \mathrm{P}(\lambda^*),$ where $\lambda^* = \lambda \exp\{\theta_p \tilde{w}^T \tilde{\eta} + \frac{1}{2}\theta_p^2 \tilde{w}^T \Sigma_J \tilde{w}\}.$

3. In the jump frequency $N = n$ of sample 2, simulate the jump sizes of the assets from

$$[V_1^*(j), V_2^*(j), \cdots, V_d^*(j)]^T \sim N_d(\tilde{\eta} + \theta_p \tilde{w}^T \Sigma_J, \Sigma_J),\ j = 1, \cdots, n,$$

where N_d denotes the d multi-variate normal distribution.

(3) Repeat step (2) k times to compute

$$\hat{p}_m(\theta_p) = \frac{1}{k}\sum_{i=1}^{k} 1_{\{f(\tilde{r}_{\theta_p}^{(i)}(t))>0\}} \exp\{-\theta_p f(\tilde{r}_{\theta_p}^{(i)}(t)) + \Psi(\theta_p)\},$$

where $\tilde{r}_{\theta_p}^{(i)}(t)$ is the return of the i th with jump risks in the importance sampling.

(4) Repeat steps (2) and (3) with the sizes of the importance sampling $M,$ and compute $\hat{p}(\theta_p)$ and standard error $se(\theta_p)$ as follows

$$\hat{p}(\theta_p) = \frac{1}{M}\sum_{m=1}^{M} \hat{p}_m(\theta_p),$$

$$se(\theta_p) = \sqrt{\frac{\sum_{m=1}^{M} (\hat{p}_m(\theta_p) - \hat{p}(\theta_p))^2}{M-1}}.$$

3.2 An algorithm for nonlinear assets

Similarly, we also obtain an algorithm of simulation for nonlinear assets based on an algorithm of simulation for linear assets. From the change measure of the exponential family, the transformation from the original measure dP_N to the new measure dP_N^θ can be written as

$$dP_N^\theta = e^{\theta f_N(\tilde{r}(t)) - \Psi_N(\theta)} dP_N. \tag{11}$$

The same process to compute the moment generating function of the density is

$$\begin{aligned} &E(\exp\{\theta f_N(\tilde{r}(t)) - \theta r_p^N\}) \\ &\quad = \exp\{\theta a_N + \frac{1}{2}\sum_{i=1}^{d}(\frac{(\theta b_i)^2}{1-2\theta\lambda_i\sigma_i^2} - \log(1-2\theta\lambda_i\sigma_i^2)) + \lambda(\exp\{\theta\tilde{\delta}_N^T\tilde{\eta} + \frac{1}{2}\theta^2\tilde{\delta}_N^T\Sigma_J\tilde{\delta}_N\} - 1)\}. \end{aligned}$$

Then, we obtain $\Psi_N(\theta)$ as

$$\Psi_N(\theta) = a_N + \frac{1}{2}\sum_{i=1}^{d}(\frac{(\theta b_i)^2}{1-2\theta\lambda_i\sigma_i^2} - \log(1-2\theta\lambda_i\sigma_i^2)) + \lambda(\exp\{\theta\tilde{\delta}_N^T\tilde{\eta} + \frac{1}{2}\theta^2\tilde{\delta}_N^T\Sigma_J\tilde{\delta}_N\} - 1).$$

Next, we use the numerical method to find θ_p^N such that

$$\begin{aligned} \Psi_N(\theta) &= a_N + \sum_{i=1}^{d}(\frac{\theta b_i^2(1-\theta\lambda_i\sigma_i^2)}{(1-2\theta\lambda_i\sigma_i^2)^2} + \frac{\lambda_i\sigma_i^2}{1-2\theta\lambda_i\sigma_i^2}) + \lambda\exp\{\theta\tilde{\delta}_N^T\tilde{\eta} + \frac{1}{2}\theta^2\tilde{\delta}_N^T\Sigma_J\tilde{\delta}_N\}(\tilde{\delta}_N^T\tilde{\eta} + \theta\tilde{\delta}_N^T\Sigma_J\tilde{\delta}_N) \\ &= 0. \end{aligned}$$

Then, we can also find the tilting distribution as described in the subsection above. By assuming independence, we can divide the tilting measure into two parts, the diffusion and jump parts. With the computed parameters, U_i becomes normal with mean and variance

$$\mu_i(\theta) = \frac{\theta b_i}{1-2\theta\lambda_i\sigma_i^2}, \quad \sigma_i^2(\theta) = \frac{1}{1-2\theta\lambda_i\sigma_i^2},$$

and the U_i remain independent of each other. The tilting distribution of the jump part is the same as in Section 3.1 (derived in Appendix D). Therefore, we can present the algorithm with importance resampling for the multi-variate jump diffusion distribution as follows:

Algorithm 2

(1) Compute the tilting point θ_p^N such that $\Psi'_N(\theta_p^N) = 0$.

(2) Generate the returns of d assets from three new samples of the tilting distribution.

1. Generate d samples for the independent identical distribution

$$U_{Ni}^* \sim N_1\left(\frac{\theta_p^N b_i}{1-2\theta_p^N \lambda_i \sigma_i^2}, \frac{1}{1-2\theta_p^N \lambda_i \sigma_i^2}\right), \ i = 1, \cdots, d.$$

2. Generate $N_N^* \sim \mathrm{P}(\lambda_N^*)$, where $\lambda_N^* = \lambda \exp\{\theta_p^N \tilde{\delta}_N^T \tilde{\eta} + \frac{1}{2}\theta_p^{N2} \tilde{\delta}_N^T \Sigma_J \tilde{\delta}_N\}$.

3. Given the abnormal event $N = n$, generate the jump sizes of the assets

$$[V_1^{N*}(j), V_2^{N*}(j), \cdots, V_d^{N*}(j)]^T \sim N_d(\tilde{\eta} + \theta_p \tilde{\delta}_N^T \Sigma_J, \Sigma_J), \ j = 1, \cdots, n.$$

(3) Repeat step (2) k times to compute

$$\hat{p}_m(\theta_p^N) = \frac{1}{k}\sum_{i=1}^{k} 1_{\{f_N(\tilde{r}_{\theta_p^N}^{(i)}(t))>0\}} \exp\{-\theta_p^N f_N(\tilde{r}_{\theta_p^N}^{(i)}(t)) + \Psi(\theta_p^N)\},$$

where $\tilde{r}_{\theta_p^N}^{(i)}(t)$ is the return of the ith simulation.

(4) Repeat steps (2) and (3) with the sizes of the importance sampling M, and compute $\hat{p}(\theta_p^N)$ and standard error $se(\theta_p^N)$ as follows

$$\hat{p}(\theta_p^N) = \frac{1}{M}\sum_{m=1}^{M} \hat{p}_m(\theta_p^N),$$

$$se(\theta_p^N) = \sqrt{\frac{\sum_{m=1}^{M}(\hat{p}_m(\theta_p^N) - \hat{p}(\theta_p^N))^2}{M-1}}.$$

Using quadratic approximation only effects the diffusion part. By the independence assumption, we can see that the diffusion part is the same as the case of quadratic approximation for multi-variate normal distribution in Glasserman, Heidelberger and Shahabuddin (2000). The jump part with the nonlinear portfolio is the same as the jump part with the linear portfolio, because the quadratic term of the jump part is small enough to be negligible.

4. Numerical Results

We consider events with $p = 0.05$, 0.01 and 0.001 to compare the efficiency of importance sampling relative to the Monte Carlo naive simulation. We study the estimated probability under various circumstance such as a linear portfolio, a nonlinear portfolio and pure jump portfolio. The comparison of efficiency is measured by the relative efficiency of the estimate $\hat{p}(\theta)$ relative to the estimate $\hat{p}$, in which it is defined (cf. Hall, 1991) as

$$\text{eff}(\hat{p}(\theta), \hat{p}) = \frac{Var(\hat{p})}{Var(\hat{p}(\theta))}, \tag{12}$$

where $Var(\hat{p}(\theta))$ is the variance of the probability estimator $\hat{p}(\theta)$ with parameter of importance sampling θ.

4.1 A linear asset

In this subsection, the relative efficiencies of simulating the tail probability with a naive Monte Carlo simulation and with the importance sampling simulation in a single jump diffusion model and a multi-variate jump diffusion are compared. In a single

Table 1. The relative efficiency of the probability estimation for a linear asset when "$d = 1$" vs. "$d = 2$"

This table reports relative efficiency of the probability estimation for a linear asset under single and multi-variate jump diffusion models. p denotes the true tail probability, r_p denotes the quantile of p, $\hat{p}$ and $\widehat{se}_p$ are the mean and standard error of the probability estimator with $10,000$ Monte Carlo simulations, $\hat{p}(\theta_p)$ and $\widehat{se}_p(\theta_p)$ are the mean and the standard error of the tail probability estimator with importance sampling, $\text{eff}(\hat{p}(\theta_p), \hat{p})$ is the relative efficiency of $\hat{p}(\theta_p)$ relative to $\hat{p}$ in single and multi-variate jump diffusion models.

	$d=1$			$d=2$		
r_p	0.0211	0.0298	0.0400	0.0429	0.0608	0.0816
p	0.0500	0.0100	0.0010	0.0500	0.0100	0.0010
$\hat{p}$	0.0499	0.0100	0.0010	0.0501	0.0099	0.0010
$\widehat{se}_p$	6.9E-03	3.1E-03	9.9E-04	6.8E-03	3.1E-03	1.0E-03
$\hat{p}(\theta_p)$	0.0501	0.0100	0.0010	0.0499	0.0100	0.0010
$\widehat{se}_p(\theta_p)$	2.4E-03	6.8E-04	1.0E-04	2.5E-03	7.0E-04	1.0E-04
$\text{eff}(\hat{p}(\theta_p), \hat{p})$	8.25	21.2	82.1	7.01	20.5	85.9

jump diffusion model, we consider parameters $\mu_1 = 0.06,\ \sigma_1 = 0.2,\ \lambda = 1,\ \eta_1 = 0,$ and $\nu_1 = 0.02,$ sample size $k = 1{,}000$ and Monte Carlo replications $M = 10{,}000.$ In a multi-variate jump diffusion model, we consider the parameters of a single jump diffusion model and let $d = 1,\ \mu_2 = 0.05,\ \sigma_2 = 0.3,\ \rho_{12} = 0.3,\ \eta_2 = 0,\ \nu_2 = 0.03,\ \rho_{12}^J = 0.5,$ sample size $k = 1{,}000$ and Monte Carlo replications $M = 10{,}000.$ For simplicity, we set $w_i(t) = 1$ for all i and $t.$ Table 1 reports that the estimate of the probability with importance sampling is more efficient than the naive Monte Carlo simulation for a large deviation.

4.2 A nonlinear asset

In Section 3.2, we propose a method to compute the tilting point for quadratic approximation for portfolios of nonlinear assets. In Table 2, we set the specific parameters to compute the tilting points for the case of a single asset and two assets, respectively. We report the estimator and standard deviation by using naive simulation and tilting point importance sampling. The parameters are $\lambda = 1,\ \eta_1 = \eta_2 = 0,\ \nu_1 = 0.02,\ \nu_2 = 0.03,\ \rho_{12}^J = 0.5,$ sample size $k = 1{,}000$ and Monte Carlo replications $M = 10{,}000.$ In Table 2,

Table 2. The relative efficiency of the probability estimation for a nonlinear asset when "d =1"vs. "d =2"

This table reports relative efficiency of the probability estimation for a nonlinear asset under single and multi-variate jump diffusion models. p denotes the true tail probability, r_p^N denotes the quantile of p, $\hat{p}$ and $\widehat{se}_p$ are the mean and standard error of the probability estimator with Monte Carlo, $\hat{p}(\theta_p^N)$ and $\widehat{se}_p(\theta_p^N)$ are the mean and the standard error of the tail probability estimator with importance sampling, $\mathrm{eff}(\hat{p}(\theta_p^N),\ \hat{p})$ is the relative efficiency of $\hat{p}(\theta_p^N)$ relative to $\hat{p}$ in single and multi-variate jump diffusion models.

	$d = 1$			$d = 2$		
r_p^N	1.6451	2.3249	3.0901	1.9454	2.7501	3.6573
p	0.0500	0.0100	0.0010	0.0500	0.0100	0.0010
$\hat{p}$	0.0500	0.0100	0.0010	0.0500	0.0099	0.0010
$\widehat{se}_p$	6.9E-03	3.1E-03	9.9E-04	6.8E-03	3.1E-03	9.8E-04
$\hat{p}(\theta_p^N)$	0.0500	0.0100	0.0010	0.0500	0.0100	0.0010
$\widehat{se}_p(\theta_p^N)$	2.2E-03	5.2E-04	1.0E-04	2.2E-03	5.2E-04	5.8E-05
$\mathrm{eff}(\widehat{p}(\theta_p^N),\ \widehat{p})$	9.39	36.5	282.3	9.75	37.5	285.7

we set $b_1 = -1$ and $\lambda_1 = 0.5$ with the quadratic approximation function for a single jump diffusion model. In Table 2, we set $b_1 = 0$, $b_2 = -1.183$, $\lambda_1 = 0.247$, and $\lambda_2 = 0.147$ with the quadratic approximation function for a two-variate jump diffusion model. The estimate of the probability with importance sampling is more efficient than naive Monte Carlo simulation when p is smaller. The efficiency of the importance sampling method is not influenced by increasing assets. The efficiency improvement in the quadratic approximation function seems to be better than in linear approximation.

4.3 An asset with pure jump risks

In this subsection, the relative efficiency of estimating the quantile with a naive Monte Carlo simulation and the importance sampling method of the pure jump diffusion are computed. For the case without a jump process, multi-variate normal distribution was proposed in Glasserman, Heidelberger and Shahabuddin (2000) and the Laplace method in Lin, Wang and Fuh (2006). Therefore, we focus on discussing the efficiency of importance sampling with a pure jump process.

We remove the diffusion part of the jump diffusion model by setting all parameters of the diffusion part to zero, that is

Table 3. The relative efficiency of the probability estimation for a linear asset with pure jump risks when "$d = 1$"vs. "$d = 2$"

This table reports relative efficiency of the probability estimation for a linear asset with pure jump risks under single and multi-variate jump diffusion models. p denotes the true tail probability, r_p denotes the quantile of p, $\hat{p}$ and $\widehat{se}_p$ are the mean and standard error of the probability estimator with Monte Carlo, $\hat{p}(\theta_p)$ and $\widehat{se}_p(\theta_p)$ are the mean and the standard error of the tail probability estimator with importance sampling, $\text{eff}(\hat{p}(\theta_p), \hat{p})$ is the relative efficiency of $\hat{p}(\theta_p)$ relative to $\hat{p}$ in single and multi-variate pure jump diffusion models.

	$d = 1$			$d = 2$		
r_p	0.0220	0.0413	0.0650	0.0481	0.0901	0.1415
p	0.0500	0.0100	0.0010	0.0500	0.0100	0.0010
$\hat{p}$	0.0500	0.0100	0.0010	0.0498	0.0100	0.0010
$\widehat{se}_p$	6.8E-03	3.1E-03	1.0E-03	6.8E-03	3.1E-03	1.0E-03
$\hat{p}(\theta_p)$	0.0499	0.0100	0.0010	0.0499	0.0100	0.00099
$\widehat{se}_p(\theta_p)$	2.9E-03	6.8E-04	8.1E-05	2.9E-03	6.8E-04	8.2E-05
$\text{eff}(\hat{p}(\theta_p), \hat{p})$	5.21	21.1	150.1	5.42	21.3	148.6

$$r_i(t) \;=\; \sum_{j=1}^{N} V_j^{(i)},\; i=1,\cdots,d.$$

Consider parameters $\lambda=100,\ \eta_1=\eta_2=0,\ \nu_1=0.02,\ \nu_2=0.03,\ \rho_{12}^J=0.5,$ sample size $k=1{,}000$ and Monte Carlo replications $M=10{,}000$. For simplicity, we set $w_i(t)=1$ for all i and t. In Table 3, the estimate of the quantile with importance sampling is more efficient than naive Monte Carlo simulation when p is smaller. Under the same p, the relative efficiency is similar for a single asset or two assets.

To summarize Table 1-Table 3, we find the following outcomes. First, the standard deviations of naive simulations are almost equal under the same probability event. Therefore, the relative efficiency is dependent on improvements to the efficiency of the importance sampling method. Second, under the three portfolio cases, the efficiency of importance sampling is different with the same p: however, the efficiency increases for large deviations and is not effected by increasing assets. Third, the pure jump process can be improved to reduce the variance in Monte Carlo simulations. Finally, there are some tools to improve the measure of the tilting point. One is to use the Laplace method (cf. Lin, Wang, and Fuh, 2006) for large deviations, and the other is to use the method of Fuh and Hu (2004) for moderate deviations.

4.4 Comparing with Extreme value theory

Oh and Moon (2006) apply the EVT-Copula method to fit fat-tailed marginal distributions, showing that the EVT-Copula method outperforms the traditional portfolio VaR method during the period of financial crises. In this subsection, we will show the robustness of our algorithm by comparing it with the EVT-Copula method under 95% VaR. Note that Oh and Moon (2006) use generalized Pareto distribution (GPD) and some copula methods to estimate portfolio VaRs, to which the model setting is different from our method. Therefore, we generate the data set from these two models, and compare the portfolio VaRs in two methods. In the EVT-Copula method, the copula functions are chosen from the GPD model with Normal and Gumbel copulas. The MLEs is computed by the pot package in **R**. The samples of returns $r_1,\cdots,r_n$ and $(r_i-\beta)/\alpha,$ with β and α are the location and scale parameters. The function $G_{c,\phi}$ is the distribution of the GPD with

$$G_{c,\phi}(x)=\begin{cases}1-(1-cx/\phi)^{1/c} & \text{if } c\neq 0,\\ 1-\exp(-x/\phi) & \text{if } c=0,\end{cases}$$

where $\phi := \alpha - c(\eta-\beta) > 0$, $x > 0$ when $c \le 0$ and $0 < x \le \phi / c$ when $c > 0$. The c and ϕ can be estimated from sample of returns r_i that exceed η by the method of maximum likelihood. In the jump diffusion model, we can estimate the parameters by the method of moment given in Ball and Torous (1983). In Table 4, we set $d = 2$, the estimated parameters are $c_1 = -0.281$, $\phi_1 = 0.0056$, $c_2 = -0.385$, $\phi_2 = 0.0062$ in the GPD model, and $\mu_1 = 0.000412$, $\sigma_1 = 0.007$, $\lambda = 0.588$, $\eta_1 = 0.001$, $\nu_1 = 0.0141$, $\mu_2 = 0.000767$, $\sigma_2 = 0.0145$, $\rho_{12} = 0.322$, $\eta_2 = -0.002$, $\nu_2 = 0.0474$, and $\rho^J_{12} = 0.261$ in the jump diffusion model. We use sample size of $k = 1{,}000$, and Monte Carlo replications $M = 10{,}000$. Here, we approximate the true VaR by using large sample size $k = 100{,}000$ in two models. It is noted from Table 4 that when the data set is from the jump diffusion model, our method is unbiased and the EVT-Copula method covers the true VaR; while when data set is from the GPD model, the estimated portfolio VaR by our method is also in between of the two chosen copula functions. Hence from this simulation study, we show that when the model is from the jump diffusion model or from the GPD model, our method is robust and efficient to estimate the portfolio VaR.

Table 4. The comparison of the 95% VaR estimation for jump diffusion model and GPD model

This table reports the 95% VaR estimation for jump diffusion model and GPD model. Under the jump diffusion model, our method is unbiased and the EVT-Copula method covers the true VaR. Under the GPD model, the estimated portfolio VaR by our method is between of the two chosen copula functions. It shows that when the model is from the jump diffusion model or from the GPD model, our method is robust and efficient to estimate the portfolio VaR.

	Jump diffusion model	GPD model
True VaR	-0.00188	-0.00194
GPD+Normal copula	-0.00205	-0.00198
GPD+Gumbel copula	-0.00185	-0.00193
Importance sampling	-0.00188	-0.00196

5. Conclusion

In this paper, algorithms for the multi-variate jump diffusion model with importance sampling are proposed to evaluate the VaR of portfolios with linear assets and nonlinear assets. To this end, we first develop an efficient computational procedure for estimating the portfolio loss probability with the jump process. The jump diffusion model is assumed to have heavy tails in the joint distribution of market risk factors. By the assumption of independence, we can separate the tilting measure from the diffusion and jump part. Importance sampling tilting measures for linear and nonlinear portfolios with a multi-variate jump diffusion model are proposed when there is a serious event in the market such as a sub-prime financial crisis or other financial crisis. For a wide range of portfolios, this method is efficient in the sense of obtaining variance reductions. The variance ratios of rare events are relatively more effective than those obtained using the naive Monte Carlo method. This method is robust for comparison with the ETV-Copula method.

There are some limitations that can be further delved into in the future studies. First, jump diffusion models can be extended to a Markov switching jump diffusion model, in which the arrival rates of jump events follow a Markov chain for multiple states in order to develop an importance sampling for evaluating VaR. Second, the Laplace method can be used to obtain more accurate tilting measures (cf. Lin, Wang, and Fuh, 2006) for large deviations or the method of Fuh and Hu (2004) for moderate deviations. Third, nonlinear financial derivatives can be used for the empirical studies. In other words, we can use a nonparametric bootstrap method with the importance resampling to evaluate VaR.

References

Ball, C. A. and W. N. Torous, 1983, A simplified jump process for common stock returns, *The Journal of Financial and Quantitative Analysis* 18, pp. 53-65.

Beran, R., 1987, Prepivoting to reduce level error of confidence sets, *Biometrika* 74, pp. 457-468.

Beran, R. and P. W. Millar, 1986, Confidence sets for a multivariate distribution, *Annals of Statistics* 14, pp. 431-443.

Black, F. and M. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, pp. 637-654.

Bucklew, J. A., 1990, *Large Deviation Techniques in Decision, Simulation, and Estimation*, New York: Wiley.

Fuh, C. D. and I. Hu, 2004, Efficient importance sampling for events of moderate deviations with applications, *Biometrika* 91, pp. 471-490.

Glasserman, P., P. Heidelberger, and P. Shahabuddin, 1999, Asymptotically optimal importance sampling and stratification for pricing path-dependent options, *Mathematical Finance* 9, pp. 117-152.

Glasserman, P., P. Heidelberger, and P. Shahabuddin, 2000, Variance reduction techniques for estimating Value-at-Risk, *Management Science* 46, pp. 1349-1364.

Glasserman, P., P. Heidelberger, and P. Shahabuddin, 2002, Portfolio Value-at-Risk with heavy-tailed risk factors, *Mathematical Finance* 12 (3), pp. 239-269.

Glasserman, P., 2003, *Monte Carlo Methods in Financial Engineering*, Springer-Verlag: New York.

Goffinet, B. and D. Wallach, 1996, Optimized importance sampling quantile estimation, *Biometrika* 83 (4), pp. 791-800.

Hall, P., 1987, On the bootstrap and likelihood-based confidence regions, *Biometrika* 74, pp. 481-493.

Hall, P., 1992, *The Bootstrap and Edgeworth Expansion*, Springer-Verlag: New York.

Hull, J., 2000, *Options, Futures, and Other Derivatives*, Prentice-Hall International, Inc. Fourth Edition.

Johns, M. V., 1988, Importance sampling for bootstrap conference intervals, *Journal of the American Statistical Association* 83 (403), pp. 709-714.

Jorion, P., 2000, *Value-at-Risk*, McGraw-Hill: New York.

Kou, S. G., 2002, A jump diffusion model for option pricing, *Management Science* 48, pp. 1086-1101.

Lin, S. K., R. H. Wang, and C. D. Fuh, 2006, Risk management for linear and non-linear assets-a bootstrap method with importance resampling to evaluate Value-at-Risk, *Asia-Pacific Financial Markets* 13, pp. 261-295.

Merton, R. C., 1973, Theory of rational option pricing, *Bell Journal of Economics and Management Science* 4, pp. 141-183.

Merton, R. C., 1976, Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics* 3, pp. 125-144.

Oh, S. and S. J. Moon, 2006, Comparative analysis of portfolio risk measures based on EVT-copula approach during financial crises, *Asia-Pacific Journal of Financial Studies* 35, pp. 175-205.

Sadowsky, J. S. and J. A. Bucklew, 1990, On large deviations theory and asymptotically efficient monte carlo estimation, *IEEE Transactions on Information Theory* 36 (3), pp. 579-588.

<Appendix A> The quadratic approximation of the return with a nonlinear asset

Assume the change in the nonlinear asset is a quadratic function of the change in the risk factors. By the delta-gamma approximation (quadratic approximation), the change in portfolio value for the nonlinear portfolio can be denoted as

$$P(t+\Delta t,\, \tilde{S}(t+\Delta t)) - P(t,\, \tilde{S}(t)) \approx \frac{\partial P}{\partial t}\Delta t + \tilde{\delta}^T \Delta\tilde{S}(t) + \frac{1}{2}\Delta\tilde{S}(t)^T \Gamma \Delta\tilde{S}(t),$$

where $\frac{\partial P}{\partial t}$ is the change of the portfolio from t to $t+\Delta t$, $\delta_i = \frac{\partial P}{\partial S_i}$ denotes the delta approximation of the portfolio for the change of asset i and $\tilde{\delta} = [\delta_1, \cdots, \delta_n]^T$ is the vector of the delta approximation, $\Gamma_{ij} = \frac{\partial^2 P}{\partial S_i \partial S_j}$ is the gamma approximation of the portfolio for the change of asset i and asset j, Γ is the matrix of the gamma approximation, and $\Delta\tilde{S}(t) = [\Delta S_1(t), \cdots, \Delta S_n(t)]^T$ denotes the change of the assets. Hence, the return of the quadratic approximation in the nonlinear asset can be rewritten as

$$\begin{aligned} f_N(\tilde{r}(t)) &= \frac{P(t+\Delta t,\, \tilde{S}(t+\Delta t)) - P(t,\, \tilde{S}(t))}{P(t,\, \tilde{S}(t))} - r_p^N \\ &\approx \frac{\frac{\partial P}{\partial t}\Delta t + \tilde{\delta}^T \Delta\tilde{S}(t) + \frac{1}{2}\Delta\tilde{S}(t)^T \Gamma \Delta\tilde{S}(t)}{P(t,\, \tilde{S}(t))} - r_p^N \\ &= \frac{\frac{\partial P}{\partial t}\Delta t}{P(t,\, \tilde{S}(t))} + \frac{\tilde{\delta}^T}{P(t,\, \tilde{S}(t))}\Delta\tilde{S}(t) + \Delta\tilde{S}(t)^T \frac{\frac{1}{2}\Gamma}{P(t,\, \tilde{S}(t))}\Delta\tilde{S}(t) - r_p^N \\ &= a(t) + \tilde{\delta}_N^T \tilde{r}(t) + \tilde{r}^T(t)\Gamma_N \tilde{r}(t) - r_p^N \end{aligned}$$

where $a(t) = \frac{\frac{\partial P}{\partial t}\Delta t}{P(t,\, \tilde{S}(t))}$ is the change of the nonlinear value relative to the nonlinear asset by the time change, $\delta_N(t) = [\frac{S_1(t)\delta_1}{P(t,\, \tilde{S}(t))}, \cdots, \frac{S_n(t)\delta_n}{P(t,\, \tilde{S}(t))}]^T$ denotes the delta approximation vector relative to the asset weight of the nonlinear value at time t, and $(\Gamma_N)_{ij} = \frac{1}{2}(\Gamma)_{ij}\frac{S_i(t)S_j(t)}{P(t,\, \tilde{S}(t))}$ presents the gamma approximation matrix of the i th asset and j th asset relative to the value of the nonlinear asset at time t.

<Appendix B> The linear form by the Cholesky decomposition

The return of the nonlinear asset by the quadratic approximation and the Cholesky decomposition method is equal to

$$
\begin{aligned}
f_N(\tilde{r}(t)) &\approx a(t)+\delta_N^T\tilde{\mu}+\delta_N^T\Sigma_\sigma\tilde{Z}+\delta_N^T\tilde{J}+\Sigma_\sigma\tilde{Z}^T\Gamma_N\tilde{Z}\Sigma_\sigma-r_p^N \\
&\overset{D}{=} a_N+\tilde{\delta}_N^T\Sigma_\sigma C\tilde{U}+\Sigma_\sigma\tilde{U}^TC^T\Gamma_N C\tilde{U}\Sigma_\sigma+\delta_N^T\tilde{J} \\
&= a_N+b^T\tilde{U}+\Sigma_\sigma\tilde{U}^T\Lambda\tilde{U}\Sigma_\sigma+\tilde{\delta}_N^T\tilde{J} \\
&= a_N+\sum_{i=1}^{d}(b_iU_i+\lambda_j\sigma_i^2U_i^2)+\tilde{\delta}_N^T\tilde{J}
\end{aligned}
$$

where $\tilde{b}^T=[b_1,b_2,\cdots,b_d]=\tilde{\delta}_N^T\Sigma_\sigma C$ and $a_N=a(t)+\tilde{\delta}_N^T\tilde{\mu}-r_p^N$.

<Appendix C> The tilting distribution of importance sampling in the linear asset

In this appendix, we find the tilting distribution as follows. By the assumption of independence, we divide the tilting measure into two parts, the diffusion and jump part. For simplicity, the measure of the diffusion part is changed into the new measure, but the measure of the jump part does not change.

$$
\begin{aligned}
dP_{\alpha,0,0} &= \exp\{\theta f(\tilde{r}(t))-\Psi(\theta)\}dP \\
&= \exp\{\theta f(\tilde{r}(t))-\Psi(\theta)\}\times\phi_d(\tilde{U};\,\tilde{0},I)\times\prod_{j=1}^{n}(\phi_d(\tilde{V}(j);\,\tilde{\eta},\Sigma_J))\times\mathrm{P}(\lambda) \\
&= \exp\{\theta_p\tilde{w}^T\tilde{\mu}+\theta\tilde{w}_\sigma^TC\tilde{U}+\theta_p\tilde{w}^T\tilde{J}-\theta_pr_p-\Psi(\theta_p)\}\times \\
&\quad \frac{1}{(2\pi)^{d/2}}\exp\{-\frac{1}{2}\tilde{U}^TI\tilde{U}\}\times\prod_{j=1}^{n}(\phi_d(\tilde{V}(j);\,\tilde{\eta},\Sigma_J))\times\mathrm{P}(\lambda) \\
&= \exp\{\theta_p\tilde{w}^T\tilde{\mu}+\frac{1}{2}\theta_p^2\tilde{w}_\sigma^T\Sigma_B\tilde{w}_\sigma+\theta_p\tilde{w}^T\tilde{J}-\theta_pr_p-\Psi(\theta_p)\}\times \\
&\quad \phi_d(\tilde{U}^*;\,\tilde{u},I)\times\prod_{j=1}^{n}(\phi_d(\tilde{V}(j);\,\tilde{\eta},\Sigma_J))\times\mathrm{P}(\lambda)
\end{aligned}
$$

where $\tilde{U}^*=[U_1-\theta_pw_1\sigma_1\sum_i c_{1i},\cdots,U_d-\theta_pw_d\sigma_d\sum_i c_{di}]^T$. Let $N=n$, then find the tilting measure of the jump size as follows:

$$
\begin{aligned}
dP_{\alpha,\beta,0} &= \exp\{\theta_p \tilde{w}^T\tilde{\mu} + \frac{1}{2}\theta_p^2 \tilde{w}_\sigma^T\Sigma_B \tilde{w}_\sigma + \theta_p \tilde{w}^T\tilde{J} - \theta_p r_p - \Psi(\theta_p)\} \\
&\quad \times \phi_d(\tilde{U}^*;\, \tilde{u}, I)\times\prod_{j=1}^{n}(\phi_d(\tilde{V}(j);\tilde{\eta},\Sigma_J))\times \mathrm{P}(\lambda) \\
&= \exp\{\theta_p \tilde{w}^T\tilde{\mu} + \frac{1}{2}\theta_p^2 \tilde{w}_\sigma^T\Sigma_B \tilde{w}_\sigma + \theta_p \tilde{w}^T\sum_{j=1}^{n}(\tilde{V}(j)-\tilde{\eta}) + \theta_p \tilde{w}^T\tilde{\eta}n - \theta_p r_p - \Psi(\theta_p)\} \\
&\quad \times\left(\frac{1}{(2\pi)^{d/2}|\Sigma_J|^{1/2}}\right)^n \exp\{-\frac{1}{2}\sum_{j=1}^{n}(\tilde{V}(j)-\tilde{\eta})^T(\Sigma_J)^{-1}(\tilde{V}(j)-\tilde{\eta})\}\times\phi_d(\tilde{U}^*;0,I)\times \mathrm{P}(\lambda) \\
&= \exp\{\theta_p \tilde{w}^T\tilde{\mu} + \frac{1}{2}\theta_p^2 \tilde{w}_\sigma^T\Sigma_B \tilde{w}_\sigma + \theta_p \tilde{w}^T\tilde{\eta}n + \frac{1}{2}\theta_p^2\tilde{w}^T\Sigma_J\tilde{w}n - \theta_p r_p - \Psi(\theta_p)\}\times\phi_d(\tilde{U}^*;\tilde{u},I)\times \mathrm{P}(\lambda) \\
&\quad \left(\frac{1}{(2\pi)^{d/2}|\Sigma_J|^{1/2}}\right)^n \exp\{-\frac{1}{2}\sum_{j=1}^{n}(\tilde{V}(j)-(\tilde{\eta}+\theta_p\tilde{w}^T\Sigma_J))^T\Sigma_J^{-1}(\tilde{V}(j)-(\tilde{\eta}+\theta_p\tilde{w}^T\Sigma_J))\} \\
&= \exp\{\theta_p \tilde{w}^T\tilde{\mu} + \frac{1}{2}\theta_p^2 \tilde{w}_\sigma^T\Sigma_B \tilde{w}_\sigma + \theta_p \tilde{w}^T\tilde{\eta}n + \frac{1}{2}\theta_p^2\tilde{w}^T\Sigma_J\tilde{w}n - \theta_p r_p - \Psi(\theta_p)\} \\
&\quad \phi_d(\tilde{U}^*;\, \tilde{u},\, I)\times\prod_{j=1}^{n}(\phi_d(\tilde{V}(j);\, \tilde{\eta}^*,\, \Sigma_J))\times \mathrm{P}(\lambda)
\end{aligned}
$$

where $\phi_d(\tilde{V}(j);\, \tilde{\eta}^*, \Sigma_J)$ is the d dimensional multi-variate normal distribution with mean vector $\tilde{\eta}+\theta_p\tilde{w}^T\Sigma_J$ and covariance Σ_J. Finally, the tilting distribution of the jump frequency is changed into

$$
\begin{aligned}
dP_{\alpha,\beta,\gamma} &= \exp\{\theta_p \tilde{w}^T\tilde{\mu} + \frac{1}{2}\theta_p^2 \tilde{w}_\sigma^T\Sigma_B \tilde{w}_\sigma + \theta_p \tilde{w}^T\tilde{\eta}n + \frac{1}{2}\theta_p^2\tilde{w}^T\Sigma_J\tilde{w}n - \theta_p r_p - \Psi(\theta_p)\} \\
&\quad \times \frac{\exp\{-\lambda\}\lambda^n}{n!}\times\phi_d(\tilde{U}^*;\tilde{u},I)\times\prod_{j=1}^{n}(\phi_d(\tilde{V}(j);\tilde{\eta}^*,\Sigma_J)) \\
&= \exp\{\theta_p \tilde{w}^T\tilde{\mu} + \frac{1}{2}\theta_p^2 \tilde{w}_\sigma^T\Sigma_B \tilde{w}_\sigma + \lambda\exp\{\theta_p\tilde{w}^T\tilde{\eta} + \frac{1}{2}\theta_p^2\tilde{w}^T\Sigma_J\tilde{w}\} - \lambda - \theta_p r_p - \Psi(\theta_p)\} \\
&\quad \times \frac{\exp\{-\lambda\exp\{\theta_p\tilde{w}^T\tilde{\eta}+\frac{1}{2}\theta_p^2\tilde{w}^T\Sigma_J\tilde{w}\}\}(\lambda\exp\{\theta_p\tilde{w}^T\tilde{\eta}+\frac{1}{2}\theta_p^2\tilde{w}^T\Sigma_J\tilde{w}\})^n}{n!} \\
&\quad \times \phi_d(\tilde{U}^*;\tilde{u},I)\times\prod_{j=1}^{n}(\phi_d(\tilde{V}(j);\tilde{\eta}^*,\Sigma_J)) \\
&= \phi_d(\tilde{U}^*;\tilde{u},I)\times\prod_{j=1}^{n}(\phi_d(\tilde{V}(j);\tilde{\eta}^*,\Sigma_J))\times \mathrm{P}(\lambda^*)
\end{aligned}
$$

where $\lambda^* = \lambda\exp\{\theta_p\tilde{w}^T\tilde{\eta}+\frac{1}{2}\theta_p^2\tilde{w}^T\Sigma_J\tilde{w}\}$, and $\theta_p \tilde{w}^T\tilde{\mu} + \frac{1}{2}\theta_p^2 \tilde{w}_\sigma^T\Sigma_B \tilde{w}_\sigma + \lambda\exp\{\theta_p\tilde{w}^T\tilde{\eta} + \frac{1}{2}\theta_p^2\tilde{w}^T\Sigma_J\tilde{w}\} - \lambda$ $-\theta_p r_p - \Psi(\theta_p) = 0.$

<Appendix D> The tilting distribution of importance sampling for the nonlinear asset

We can find the tilting distribution as follows. By the assumption of independence, we can divide the tilting measure into two parts, the diffusion and jump part. For simplicity, the measure of the diffusion part is changed into the new measure, but the measure of the jump part does not change.

$$\begin{aligned} dP_{\alpha,0,0} &= \exp\{\theta_p^N f_N(\tilde{r}(t)) - \Psi_N(\theta_p^N)\} \times \phi_d(\tilde{U};0,I) \times \prod_{j=1}^{n}(\phi_d(\tilde{V}(j);\tilde{\eta},\Sigma_J)) \times \mathrm{P}(\lambda) \\ &= \exp\{\theta_p^N(\sum_{i=1}^{d}(b_i U_i + \lambda_i \sigma_i^2 U_i^2) + \tilde{\delta}_N^T \tilde{J}) + \theta_p^N a_N - \Psi_N(\theta_p^N)\} \\ &\quad \times \frac{1}{(2\pi)^{d/2}} \exp\{-\frac{1}{2}\tilde{U}^T I \tilde{U}\} \times \prod_{j=1}^{n}(\phi_d(\tilde{V}(j);\tilde{\eta},\Sigma_J)) \times \mathrm{P}(\lambda) \\ &= \exp\{\frac{1}{2}\sum_{i=1}^{d}(\frac{(\theta_p^N b_i)^2}{1-2\theta_p^N \lambda_i \sigma_i^2} - \log(1-2\theta_p^N \lambda_i \sigma_i^2)) + \theta_p^N \tilde{\delta}_N^T \tilde{J} + \theta_p^N a_N - \Psi_N(\theta_p^N)\} \\ &\quad \times \phi_d(\tilde{U}_N^*;\tilde{0},\Sigma_N) \times \prod_{j=1}^{n}(\phi_d(\tilde{V}(j);\tilde{\eta},\Sigma_J)) \times \mathrm{P}(\lambda) \end{aligned}$$

where the new measure of the multi-variate independent normal distribution with

$$\tilde{U}_N^* = [U_1^* - \frac{\theta_p^N b_1}{1-2\theta_p^N \lambda_1 \sigma_1^2}, \cdots, U_d^* - \frac{\theta_p^N b_d}{1-2\theta_p^N \lambda_d \sigma_d^2}]^T$$

and covariance

$$\Sigma_N = \begin{bmatrix} \frac{1}{1-2\theta_p^N \lambda_1 \sigma_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{1-2\theta_p^N \lambda_2 \sigma_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{1-2\theta_p^N \lambda_d \sigma_d^2} \end{bmatrix},$$

and $\mathrm{P}(\lambda)$ is the probability density function of a Poisson distribution with parameter λ. Conditional on $N = n$, we first compute the tilting measure of the jump size,

$$
\begin{aligned}
dP_{\alpha,\beta,0} &= \exp\{\frac{1}{2}\sum_{i=1}^{d}(\frac{(\theta_p^N b_i)^2}{1-2\theta_p^N\lambda_i\sigma_i^2}-\log(1-2\theta_p^N\lambda_i\sigma_i^2))+\theta_p^N\tilde{\delta}_N^T\tilde{J}+\theta_p^N a_N-\Psi_N(\theta_p^N)\} \\
&\quad\times\phi_d(\tilde{U}_N^*;\tilde{0},\Sigma_N)\times\prod_{j=1}^{n}\phi_d(\tilde{V}(j);\tilde{\eta},\Sigma_J)\times\mathrm{P}(\lambda) \\
&= \exp\{\frac{1}{2}\sum_{i=1}^{d}(\frac{(\theta_p^N b_i)^2}{1-2\theta_p^N\lambda_i\sigma_i^2}-\log(1-2\theta_p^N\lambda_i\sigma_i^2))+\theta_p^N\tilde{\delta}_N^T\tilde{J}+\theta_p^N a_N-\Psi_N(\theta_p^N)\} \\
&\quad\times\phi_d(\tilde{U}_N^*;\tilde{0},\Sigma_N)\times\mathrm{P}(\lambda)\times\left(\frac{1}{(2\pi)^{d/2}|\Sigma_J|^{1/2}}\right)^n\exp\{-\frac{1}{2}\sum_{j=1}^{n}(\tilde{V}(j)-\tilde{\eta})^T\Sigma_J^{-1}(\tilde{V}(j)-\tilde{\eta})\} \\
&= \exp\{\frac{1}{2}\sum_{i=1}^{d}(\frac{(\theta_p^N b_i)^2}{1-2\theta_p^N\lambda_i\sigma_i^2}-\log(1-2\theta_p^N\lambda_i\sigma_i^2))+\theta_p^N\tilde{\delta}_N^T\tilde{\eta}n+\frac{1}{2}\theta_p^{N2}\tilde{\delta}_N^T\Sigma_J\tilde{\delta}_N n+\theta_p^N a_N-\Psi_N(\theta_p^N)\} \\
&\quad\times\phi_d(\tilde{U}_N^*;\tilde{0},\Sigma_N)\times\mathrm{P}(\lambda)\times\left(\frac{1}{(2\pi)^{d/2}|\Sigma_J|^{1/2}}\right)^n \\
&\quad\times\exp\{-\frac{1}{2}\sum_{j=1}^{n}(\tilde{V}(j)-(\tilde{\eta}+\theta_p^N\tilde{\delta}_N^T\Sigma_J))^T\Sigma_J^{-1}(\tilde{V}(j)-(\tilde{\eta}+\theta_p^N\tilde{\delta}_N^T\Sigma_J))\} \\
&= \exp\{\frac{1}{2}\sum_{i=1}^{d}(\frac{(\theta_p^N b_i)^2}{1-2\theta_p^N\lambda_i\sigma_i^2}-\log(1-2\theta_p^N\lambda_i\sigma_i^2))+\theta_p^N\tilde{\delta}_N^T\tilde{\eta}n+\frac{1}{2}\theta_p^{N2}\tilde{\delta}_N^T\Sigma_J\tilde{\delta}_N n+\theta_p^N a_N-\Psi_N(\theta_p^N)\} \\
&\quad\times\phi_d(\tilde{U}_N^*;\tilde{0},\Sigma_N)\times\prod_{j=1}^{n}(\phi_d(\tilde{V}_N^*(j);\tilde{\eta}+\theta_p^N\tilde{\delta}_N^T\Sigma_J,\Sigma_J)\times\mathrm{P}(\lambda),
\end{aligned}
$$

where $\phi_d(\tilde{V}_N^*(j);\tilde{\eta}+\theta_p^N\tilde{\delta}_N^T\Sigma_J,\Sigma_J)$ is the d-dimensional normal density function of the j th jump size with mean vector $\tilde{\eta}+\theta_p^N\tilde{\delta}_N^T\Sigma_J$ and variance matrix Σ_J. Finally, the tilting distribution of jump frequency is

$$
\begin{aligned}
dP_{\alpha,\beta,\gamma} &= \exp\{\frac{1}{2}\sum_{i=1}^{d}(\frac{(\theta_p^N b_i)^2}{1-2\theta_p^N\lambda_i\sigma_i^2}-\log(1-2\theta_p^N\lambda_i\sigma_i^2))+\theta_p^N\tilde{\delta}_N^T\tilde{\eta}n+\frac{1}{2}\theta_p^{N2}\tilde{\delta}_N^T\Sigma_J\tilde{\delta}_N n+\theta_p^N a_N-\Psi_N(\theta_p^N)\} \\
&\quad\times\phi_d(\tilde{U}_N^*;\tilde{0},\Sigma_N)\times\prod_{j=1}^{n}(\phi_d(\tilde{V}_N^*(j);\tilde{\eta}+\theta_p^N\tilde{\delta}_N^T\Sigma_J,\Sigma_J)\times\frac{e^{-\lambda}\lambda^n}{n!} \\
&= \exp\{\frac{1}{2}\sum_{i=1}^{d}(\frac{(\theta_p^N b_i)^2}{1-2\theta_p^N\lambda_i\sigma_i^2}-\log(1-2\theta_p^N\lambda_i\sigma_i^2))+\lambda(\exp\{\theta_p^N\tilde{\delta}_N^T\tilde{\eta}+\frac{1}{2}\theta_p^{N2}\tilde{\delta}_N^T\Sigma_J\tilde{\delta}_N\}-1) \\
&\quad+\theta_p^N a_N-\Psi_N(\theta_p^N)\}\times\phi_d(\tilde{U}_N^*;\tilde{0},\Sigma_N)\times\prod_{j=1}^{n}(\phi_d(\tilde{V}_N^*(j);\tilde{\eta}+\theta_p^N\tilde{\delta}_N^T\Sigma_J,\Sigma_J) \\
&\quad\times\frac{\exp\{-\lambda\exp\{\theta_p^N\tilde{\delta}_N^T\tilde{\eta}+\frac{1}{2}\theta_p^{N2}\tilde{\delta}_N^T\Sigma_J\tilde{\delta}_N\}\}(\lambda\exp\{\theta_p^N\tilde{\delta}_N^T\tilde{\eta}+\frac{1}{2}\theta_p^{N2}\tilde{\delta}_N^T\Sigma_J\tilde{\delta}_N\})^n}{n!} \\
&= \phi_d(\tilde{U}_N^*;\tilde{0},\Sigma_N)\times\prod_{j=1}^{n}(\phi_d(\tilde{V}_N^*(j);\tilde{\eta}+\theta_p^N\tilde{\delta}_N^T\Sigma_J,\Sigma_J)\times\mathrm{P}(\lambda_N^*),
\end{aligned}
$$

where $P(\lambda_N^*)$ is the probability density function of Poisson distribution with $\lambda_N^* = \lambda \exp\{\theta_p^N \tilde{\delta}_N^T \tilde{\eta} + \frac{1}{2}\theta_p^{N2}\, \tilde{\delta}_N^T \Sigma_J \tilde{\delta}_N)\}$ and $\frac{1}{2}\sum_{i=1}^{d}(\frac{(\theta_p^N b_i)^2}{1-2\theta_p^N \lambda_i \sigma_i^2} - \log(1-2\theta_p^N \lambda_i \sigma_i^2)) + \lambda(\exp\{\theta_p^N \tilde{\delta}_N^T \tilde{\eta} + \frac{1}{2}\theta_p^{N2}\, \tilde{\delta}_N^T \Sigma_J \tilde{\delta}_N\} - 1) + \theta_p^N a_N - \Psi_N(\theta_p^N) = 0.$