QUASI-MONTE CARLO METHOD FOR INFINITELY DIVISIBLE RANDOM VECTORS VIA SERIES REPRESENTATIONS*

JUNICHI IMAI[†] AND REIICHIRO KAWAI[‡]

Abstract. An infinitely divisible random vector without Gaussian component admits representations of shot noise series. Due to possible slow convergence of the series, they have not been investigated as a device for Monte Carlo simulation. In this paper, we investigate the structure of shot noise series representations from a simulation point of view and discuss the effectiveness of quasi-Monte Carlo methods applied to series representations. The structure of series representations in nature tends to decrease their effective dimension and thus increase the efficiency of quasi-Monte Carlo methods, thanks to the greater uniformity of low-discrepancy sequence in the lower dimension. We illustrate the effectiveness of our approach through numerical results of moment and tail probability estimations for stable and gamma random variables.

Key words. quasi-Monte Carlo method, effective dimension, gamma process, moment estimation, tail probability estimation, Poisson process, shot noise

AMS subject classifications. 60E07, 65C05, 65D30, 65B10

DOI. 10.1137/090752365

1. Introduction. An infinitely divisible random vector without Gaussian component admits representations of shot noise series. Such series representations have played an important role in theories such as the tail probability of stable random vectors and have also been studied in the applied literature, known as “shot noise.” (See, for example, Vervaat [28] and the references therein.) Series representations involving Poisson arrival times are given by Ferguson and Klass [6] for real independent increment processes without Gaussian component and with positive jumps. The theory of stable processes and their applications are expanded, due to LePage [16], on series representation of stable random vectors. The simulation of nonnegative infinitely divisible random variables is considered, and their series representations as a special form of generalized shot noise is developed in Bondesson [3]. The same approach is used in Rosiński [21] as a general pattern for series representations of Banach space valued infinitely divisible random vectors.

Infinitely divisible random vectors consist of three independent components: a constant, a centered normal random vector, and a Poisson component. The generation of the Gaussian component is standard and will not be discussed in this paper. The Poisson component is governed in full by the so-called Lévy measure. If the Lévy measure is infinite, then the Poisson component consists of infinitely many jumps. It is obviously impossible to generate infinitely many jumps on a computer. Possible simulation methods can be summarized as follows.

(i) Direct generation: When an explicit density function, or an alternative exact simulation method, is available, a simulation should be straightforward in principle.

*Received by the editors March 11, 2009; accepted for publication (in revised form) April 12, 2010; published electronically June 24, 2010. This work was supported by the Japan Society for the Promotion of Science through Grants-in-Aid for Scientific Research (18710126 and 20740059).

<http://www.siam.org/journals/sisc/32-4/75236.html>

[†]Faculty of Science and Technology, Keio University, Yokohama 223-8522, Japan (jimai@ae.keio.ac.jp).

[‡]Department of Mathematics, University of Leicester, Leicester LE1 7RH, UK (reiichiro.kawai@gmail.com).

Stable, exponential, gamma, geometric, and negative binomials are in this category. To simulate random variables from more complicated distributions, a numerical inversion method is proposed in Hörmann and Leydold [7] to numerically approximate the density functions with piecewise Hermite functions. Moreover, this method is applied in Imai and Tan [13, 14] to the simulation of variates from infinitely divisible random distributions of practical interest, such as generalized hyperbolic, normal inverse Gaussian, and Meixner distributions.

(ii) Generation with series representation: Series representation provides perfect and often easy simulation of infinitely divisible random vectors. A disadvantage of this method is that some series may converge at an extremely slow rate. Then, one might need a huge number of terms to reach a desired accuracy of the approximation. On the other hand, with ever increasing computational speed, a slow convergence may no longer cause a serious practical issue. (See, for example, Kawai [15] and Madan and Yor [17] for simulation use of series representations.)

(iii) Generation from compound Poisson: If a neighborhood of the origin of the Lévy measure is discarded or replaced by its mean value, then the remainder corresponds to a compound Poisson random vector with a constant shift. It converges to its true random vector as the intensity of the discarded part of the Lévy measure decreases. The compound Poisson component can be simulated in a precise manner. The aforementioned series representation provides a consistent way to construct such approximations. However, when the discarded part of the Lévy measure is too intense, the corresponding component may produce a substantial error. In such cases, one may often approximate the discarded component by a Gaussian random vector with an appropriate small variance, as investigated in Asmussen and Rosiński [2]. This approximation complements a method through series representations when the series converges slowly.

The main purpose of this paper is to discuss the effectiveness of the quasi-Monte Carlo (QMC) method applied to shot noise series representations. We aim at moment estimations as well as tail probability estimations, rather than a pointwise approximation. It is well known that the QMC method could produce more accurate estimations than the Monte Carlo (MC) method since a low-discrepancy sequence in the QMC method has an asymptotically superior convergence rate compared to random numbers used in the MC method. The practical advantage of the QMC method, on the other hand, can be justified by a notion of effective dimension of the problem, first introduced in Caflisch, Morokoff, and Owen [5]. It is demonstrated that the QMC method works better than the MC method for nominally high-dimensional problems as long as their effective dimensions are small. In this paper, we demonstrate that the structure of series representations tends to decrease their effective dimension in nature and thus increase the efficiency of the QMC method. Since the QMC method makes the best use of a greater uniformity of low-discrepancy sequence in the lower dimension, it is natural to expect the applicability of the QMC method to series representations of infinitely divisible random vectors. We illustrate the effectiveness of our approach through numerical results of moment and tail probability estimations for stable and gamma random variables.

The rest of this paper is organized as follows. Section 2 briefly summarizes background material on series representations of infinitely divisible random vectors with a view towards simulation. In section 3, we discuss the effectiveness of the QMC method applied to Poisson interarrival times in series representations and review the concept of cumulative explanatory ratio. In section 4, we provide numerical results

to illustrate the effectiveness of the QMC method applied to series representations of stable and of gamma random variables. Finally, section 5 concludes.

2. Series representation of infinitely divisible random vectors. Let us begin this section with the notation which will be used throughout this paper. We denote by $\mathbb{N}$ the collection of positive integers, by $\mathbb{R}^d$ the d -dimensional Euclidean space with the norm $\|\cdot\|$, $\mathbb{R}_0^d := \mathbb{R}^d \setminus \{0\}$, $\mathbb{R}_+ := (0, +\infty)$, and by $\mathcal{B}(\mathbb{R}_0^d)$ the Borel σ -field of $\mathbb{R}_0^d$. We denote by $\stackrel{\mathcal{L}}{=}$ the identity in law. We say that an infinitely divisible random vector X in $\mathbb{R}^d$ is generated by the triplet (γ, A, ν) if its characteristic function is given by

$$\mathbb{E} \left[e^{i\langle y, X \rangle} \right] = \exp \left[\left(i\langle y, \gamma \rangle - \frac{1}{2} \langle y, Ay \rangle + \int_{\mathbb{R}_0^d} (e^{i\langle y, z \rangle} - 1 - i\langle y, z \rangle \mathbb{1}_{(0,1]}(\|z\|)) \nu(dz) \right) \right],$$

where $\gamma \in \mathbb{R}^d$, A is a symmetric nonnegative-definite $d \times d$ matrix, and ν is a Lévy measure on $\mathbb{R}_0^d$, that is, a σ -finite measure satisfying $\int_{\mathbb{R}_0^d} (\|z\|^2 \wedge 1) \nu(dz) < +\infty$. Throughout this paper, we consider only infinitely divisible laws without Gaussian component, that is, $A \equiv 0$. Moreover, we denote by $\{\Gamma_k\}_{k \in \mathbb{N}}$ the arrival times of a standard Poisson process and by $\{E_k\}_{k \in \mathbb{N}}$ a sequence of independent and identically distributed (i.i.d.) exponential random variables with unit mean.

Let us review generalities on the series representation of infinitely divisible random vectors with a view towards simulation. Our discussion is essentially parallel to the so-called inverse Lévy measure method [6, 16]. Notice first that the random variable $\sum_{k=1}^{+\infty} \Gamma_k \mathbb{1}(\Gamma_k \in [0, T])$ is infinitely divisible with Lévy measure $\nu(dz) = dz$ defined on $(0, T]$. Recall also that the epochs of an inhomogeneous Poisson process on $[0, T]$ with intensity $h(t)$ can be generated by $H(\Gamma_1), H(\Gamma_2), \dots$, where $H(t) := \inf\{u \in [0, T] : \int_0^u h(s)ds < t\}$, provided that $\int_0^T h(s)ds < +\infty$. Therefore, by regarding the intensity $h(t)$ as a Lévy measure (“on state space” rather than on time), we deduce that $\sum_{k=1}^{+\infty} H(\Gamma_k) \mathbb{1}(\Gamma_k \in [0, T])$ is an infinitely divisible random variable with Lévy measure $\nu(dz) = h(z)dz$ defined on $(0, T]$. Notice here that the definition of $H(t)$ implicitly assumes that the Lévy measure ν has a compact support. Moreover, the condition $\int_0^T h(s)ds < +\infty$ implies that the Lévy measure is finite. We can extend this formulation to an infinite Lévy measure on $\mathbb{R}_+$, simply by redefining the inverse function H as running down from the infinity rather than up the other way, that is,

$$(2.1) \quad H(r) := \inf \left\{ u > 0 : \int_u^{+\infty} h(s)ds < r \right\},$$

and compute $\sum_{k=1}^{+\infty} H(\Gamma_k)$, where $\{\Gamma_k\}_{k \in \mathbb{N}}$ is no longer restricted to lie in $[0, T]$. (See Asmussen and Glynn [1].) A series representation of stable random vectors can be derived in closed form through this inverse Lévy measure method as follows.

Example 2.1 (stable random vector). An infinitely divisible random vector X in $\mathbb{R}^d$, without Gaussian component, is called *stable* if its Lévy measure is given by

$$\nu_\alpha(B) = \int_{S^{d-1}} \int_{\mathbb{R}_+} \mathbb{1}_B(r\xi) \frac{dr}{r^{1+\alpha}} \sigma(d\xi), \quad B \in \mathcal{B}(\mathbb{R}_0^d),$$

where $\alpha \in (0, 2)$ is the stability index and where σ is a finite positive measure on

S^{d-1} . Its characteristic function is given by

(2.2)

$$\begin{aligned} \mathbb{E} \left[e^{i\langle y, X \rangle} \right] &= \exp \left[i\langle y, \eta \rangle + \int_{\mathbb{R}_0^d} (e^{i\langle y, z \rangle} - 1 - i\langle y, z \rangle \mathbb{1}_{(0,1]}(\|z\|)) \nu_\alpha(dz) \right] \\ &= \begin{cases} \exp \left[i\langle y, \tau_\alpha \rangle - c_\alpha \int_{S^{d-1}} |\langle y, \xi \rangle|^\alpha \left(1 - i \tan \frac{\pi\alpha}{2} \operatorname{sgn} \langle y, \xi \rangle \right) \sigma(d\xi) \right] & \text{if } \alpha \neq 1, \\ \exp \left[i\langle y, \tau_1 \rangle - \frac{\pi}{2} \int_{S^{d-1}} \left(|\langle y, \xi \rangle| + i \frac{2}{\pi} \langle y, \xi \rangle \ln |\langle y, \xi \rangle| \right) \sigma(d\xi) \right] & \text{if } \alpha = 1, \end{cases} \end{aligned}$$

where $c_\alpha = |\Gamma(-\alpha) \cos \frac{\pi\alpha}{2}|$ and where $\tau_\alpha = \eta - \frac{1}{1-\alpha} \int_{S^{d-1}} \xi \sigma(d\xi)$ when $\alpha \neq 1$ and $\tau_1 = \eta - (1-\gamma) \int_{S^{d-1}} \xi \sigma(d\xi)$, with $\gamma (= 0.5772 \dots)$ being the Euler constant. Let $\{U_k\}_{k \in \mathbb{N}}$ be a sequence of i.i.d. random vectors in S^{d-1} with common distribution $\sigma(d\xi)/\sigma(S^{d-1})$, and set

$$\eta = \begin{cases} \frac{1}{1-\alpha} \int_{S^{d-1}} \xi \sigma(d\xi) & \text{if } \alpha \neq 1, \\ 0 & \text{if } \alpha = 1, \end{cases} \quad z_0 = \begin{cases} 0 & \text{if } \alpha \in (0, 1), \\ \int_{S^{d-1}} \xi \sigma(d\xi)/\sigma(S^{d-1}) & \text{if } \alpha \in [1, 2) \end{cases}$$

and

$$\lambda = \begin{cases} 0 & \text{if } \alpha \in (0, 1), \\ \sigma(S^{d-1}) (\gamma + \ln(\sigma(S^{d-1}))) & \text{if } \alpha = 1, \\ \left(\frac{\alpha}{\sigma(S^{d-1})} \right)^{-1/\alpha} \zeta\left(\frac{1}{\alpha}\right) & \text{if } \alpha \in (1, 2), \end{cases}$$

where ζ denotes the Riemann zeta function. Then, it holds that

$$(2.3) \quad \left(\frac{\alpha}{\sigma(S^{d-1})} \right)^{-1/\alpha} \sum_{k=1}^{+\infty} \left(\Gamma_k^{-1/\alpha} U_k - k^{-1/\alpha} z_0 \right) + \lambda z_0 \stackrel{\mathcal{L}}{=} X.$$

The above centering constants are obtained in Proposition 5.5 of Rosiński [22].

Every infinitely divisible random vector admits a series representation through the inverse Lévy measure method. In most cases, however, the tail inverse of Lévy measure, that is, the function $H(r)$ defined by (2.1), is not available in closed form, with exceptions such as the stable random vector of Example 2.1 and the layered stable random vector of Houdré and Kawai [9]. To obtain a closed form, some alternative methods have been proposed, for example, the thinning method and the rejection method of [21]. Each of those methods can be considered as a special case of the so-called generalized shot noise method of [3, 21], which we describe as follows. Assume that Lévy measure ν can be decomposed as

$$(2.4) \quad \nu(B) = \int_{\mathbb{R}_+} \mathbb{P}(H(r, U) \in B) dr, \quad B \in \mathcal{B}(\mathbb{R}_0^d),$$

where U is a random variable taking values in a suitable space $\mathcal{U}$ and where $H : \mathbb{R}_+ \times \mathcal{U} \mapsto \mathbb{R}_0^d$ is such that, for each $u \in \mathcal{U}$, $r \mapsto \|H(r, u)\|$ is nonincreasing. Then, an infinitely divisible random vector X generated by the triplet $(0, 0, \nu)$ is identical in law to the shot noise series representation

$$(2.5) \quad X \stackrel{\mathcal{L}}{=} \sum_{k=1}^{+\infty} [H(\Gamma_k, U_k) - c_k],$$

where $\{U_k\}_{k \in \mathbb{N}}$ is a sequence of i.i.d. copies of the random variable U , and $\{c_k\}_{k \in \mathbb{N}}$ is a sequence of constants defined by $c_k := \mathbb{E}[H(\Gamma_k, U) \mathbb{1}(\|H(\Gamma_k, U)\| \leq 1)]$. Here, the sequence $\{U_k\}_{k \in \mathbb{N}}$ is independent of $\{\Gamma_k\}_{k \in \mathbb{N}}$. Let us present two concrete examples for which the inverse Lévy measure method provides no explicit representations and the generalized shot noise series method based upon (2.4) is required.

Example 2.2 (gamma random variable). An infinitely divisible random variable X is *gamma* if it is characterized by the characteristic function

$$\mathbb{E}[e^{iyX}] = \exp \left[\int_{\mathbb{R}_+} (e^{iyz} - 1) a \frac{e^{-bz}}{z} dz \right],$$

where $a > 0$ and $b > 0$. The inverse Lévy measure representation provides a shot noise series representation

$$\frac{1}{b} \sum_{k=1}^{+\infty} \text{Ei}^{-1} \left(\frac{\Gamma_k}{a} \right) \stackrel{\mathcal{L}}{=} X,$$

where $\text{Ei}(x) := \int_x^{+\infty} u^{-1} e^{-u} du$ is the exponential integral function. However, this representation is not directly useful since neither the function $\text{Ei}(x)$ nor its inverse can be given in closed form. On the other hand, an explicit representation is given in [3] by

$$(2.6) \quad \sum_{k=1}^{+\infty} e^{-\Gamma_k/a} \frac{V_k}{b} \stackrel{\mathcal{L}}{=} X,$$

where $\{V_k\}_{k \in \mathbb{N}}$ is a sequence of i.i.d. exponential random variables with unit mean. There are yet other explicit representations available through the thinning method and the rejection method of [21].

The availability of different representations implies nonuniqueness of the decomposition (2.4). For example, the representation (2.6) is based on the decomposition

$$\int_{\mathbb{R}_+} \mathbb{1}_B(z) a \frac{e^{-bz}}{z} dz = \int_{\mathbb{R}_+} \left(\int_{\mathbb{R}_+} \mathbb{1} \left(e^{-\frac{r}{a}} \frac{v}{b} \in B \right) e^{-v} dv \right) dr,$$

where $\int_{\mathbb{R}_+} \mathbb{1}(\cdot \in B) e^{-v} dv$ acts as the standard exponential law. Observe that for each $v \in \mathbb{R}_+$ the random sequence $\{e^{-\Gamma_k/a} v/b\}_{k \in \mathbb{N}}$ is almost surely decreasing in k , while $\{e^{-\Gamma_k/a} V_k/b\}_{k \in \mathbb{N}}$ is not because of the additional random sequence $\{V_k\}_{k \in \mathbb{N}}$.

The next example presents a series representation of more intricate structure.

Example 2.3 (CGMY random variable). The CGMY random variable is characterized by a Lévy measure in the form

$$\nu(dz) = C|z|^{-1-Y} \left[e^{-G|z|} \mathbb{1}(z < 0) + e^{-M|z|} \mathbb{1}(z > 0) \right] dz,$$

where $C > 0$, $G > 0$, $M > 0$, and $Y < 2$. (See Carr et al. [4] for more details.) If $Y \in (0, 2)$, then it is in a subclass of the tempered stable random vector of Rosiński [22]. An explicit representation is given in [22] by

$$\sum_{k=1}^{+\infty} \left[\left(\frac{Y\Gamma_k}{2C} \right)^{-1/Y} \wedge V_k J_k^{1/Y} |Z_k| \right] \frac{Z_k}{|Z_k|} + \gamma,$$

where $\{V_k\}_{k \in \mathbb{N}}$ is a sequence of i.i.d. exponential random variables with unit mean, $\{J_k\}_{k \in \mathbb{N}}$ is a sequence of i.i.d. uniform random variables on $[0, 1]$, $\{Z_k\}_{k \in \mathbb{N}}$ is a sequence of i.i.d. random variables taking values $-G^{-1}$ and M^{-1} with equal probability, and γ is a suitable constant. Here, all the random sequences $\{\Gamma_k\}_{k \in \mathbb{N}}$, $\{V_k\}_{k \in \mathbb{N}}$, $\{J_k\}_{k \in \mathbb{N}}$, and $\{Z_k\}_{k \in \mathbb{N}}$ are mutually independent. This representation involves three additional sequences of i.i.d. random variables. Thus, four integrals are involved in the decomposition (2.4). (Three other representations are available in closed form and will be discussed in [11]. See also, for example, [15, 17], where the above series representation is used for simulation.)

As can be seen in the above examples, series representations for an infinite Lévy measure necessarily have a form of infinite sum as well. To deal with such infinite series towards simulation, we need to truncate the infinite series up to a certain finite point. A straightforward approach is the truncation to a finite number of terms of the series

$$(2.7) \quad X_n := \sum_{k=1}^n [H(\Gamma_k, U_k) - c_k].$$

An alternative approach to the infinite series is the truncation to a finite time span of the Poisson process

$$(2.8) \quad X_{(n)} := \sum_{\{k \in \mathbb{N}: \Gamma_k \leq n\}} [H(\Gamma_k, U_k) - c_k].$$

In connection with the fact that the function $H(r, v)$ is nonincreasing in r , when the truncation level n is large enough, it is fairly reasonable to expect that the above two truncations (2.7) and (2.8) behave in a similar manner.

The common and key building block for simulation of series representations is sampling of the epochs of a standard Poisson process. This can be generated iteratively as a successive summation of i.i.d. exponential random variables:

$$(2.9) \quad \{\Gamma_1, \Gamma_2, \Gamma_3, \dots\} \stackrel{\mathcal{L}}{=} \left\{ \sum_{k=1}^1 E_k, \sum_{k=1}^2 E_k, \sum_{k=1}^3 E_k, \dots \right\}.$$

The exponential random variables $\{E_k\}_{k \in \mathbb{N}}$ act as interarrival times of a standard Poisson process and can easily be generated as $E_k = -\ln J_k$, where $\{J_k\}_{k \in \mathbb{N}}$ is a sequence of i.i.d. uniform random variables on $[0, 1]$. The truncation (2.8) provides an interesting interpretation of discarded small jumps. Assume, for simplicity, that the function H has the form $\|H(r, u)\| = h(r)$, where h is a nonincreasing function from $\mathbb{R}_+$ to $\mathbb{R}_+$. Then, for each $n \in \mathbb{N}$ such that $h(n) < 1$, the Itô–Wiener isometry yields the $L^2(\Omega)$ -estimate of the discarded small jump component

$$\mathbb{E} [\|X - X_{(n)}\|^2] = \mathbb{E} \left[\left(\int_0^1 \int_{\|z\| \leq h(n)} z(\mu - \nu)(dz, ds) \right)^2 \right] = \int_{\|z\| \leq h(n)} \|z\|^2 \nu(dz),$$

where μ is a Poisson random measure on $\mathbb{R}_0^d \times [0, 1]$ whose compensator is the Lévy measure ν . This discarded small jump component may be replaced with a Gaussian random vector under some mild condition. (For details, see [2].)

3. Variance contribution of the lower dimension. As mentioned in the introduction, the advantage of the QMC method is attributed to its potentially superior convergence rate of $O(N^{-1} \ln^d N)$, while that of the MC method is $O(N^{1/2})$, where N and d denote, respectively, the number of iterations and the dimension. Nevertheless, several studies point out that this theoretical higher asymptotic convergence rate of the QMC is not achievable in practice for high-dimensional problems. To resolve the above puzzle and to explain a practical advantage of the QMC method, especially when we deal with shot noise series representations, let us review in this section the analysis of variance (ANOVA) decomposition and the concept of effective dimension. We refer the reader to Wang and Fang [27] for more details.

Let the set $\mathcal{A} := \{1, 2, \dots, d\}$ denote the coordinate axes of the d -dimensional unit cube $[0, 1]^d$. Then, for each subset $\mathcal{S} \subseteq \mathcal{A}$, we define by $|\mathcal{S}|$ its cardinality and by $\mathcal{A} \setminus \mathcal{S}$ its complementary set. Each point in $[0, 1]^d$ is denoted by $x := (x_1, \dots, x_d)'$, and $x_{\mathcal{S}}$ denotes the $|\mathcal{S}|$ -vector of components x_k for $k \in \mathcal{S}$. Let f be a function mapping from $[0, 1]^d$ to $\mathbb{R}$, and consider a problem of computing the following integral:

$$I(f) := \int_{[0,1]^d} f(x) dx.$$

We can employ the MC and QMC methods for approximating the integral by taking a sample mean, that is,

$$Q_n(f) := \frac{1}{n} \sum_{k=1}^n f(x^{(k)}),$$

where the sequence $\{x^{(k)}\}_{k \in \mathbb{N}}$ corresponds to a sequence of i.i.d. replications in the MC method, while it is generated from a low-discrepancy sequence in the QMC method. It is well known that the MC error is of order $O(n^{1/2})$ for square integrable functions and is independent of the dimension d . On the other hand, the Koksma–Hlawka inequality provides the QMC error bound by

$$|I(f) - Q_n(f)| \leq V_{HK} D^*(\{x^{(k)}\}_{k=1, \dots, n}),$$

where $D^*(\{x^{(k)}\}_{k=1, \dots, n})$ denotes the star discrepancy of the sequence $\{x^{(k)}\}_{k=1, \dots, n}$ and V_{HK} is the variation in the sense of Hardy and Krause.

Under the mild condition that f is a square integrable function, the ANOVA decomposition expresses the integrand f as a sum of 2^d additive functions:

$$f(x) = \sum_{\mathcal{S} \subseteq \mathcal{A}} f_{\mathcal{S}}(x),$$

where the function $f_{\mathcal{S}}$, which depends only on the component of x in the dimension $\mathcal{S}$, is defined recursively by

$$f_{\mathcal{S}}(x) := \int_{[0,1]^{\mathcal{A} \setminus \mathcal{S}}} f(y) dy_{\mathcal{A} \setminus \mathcal{S}} - \sum_{\mathcal{S}' \subset \mathcal{S}} f_{\mathcal{S}'}(x),$$

with the usual convention $f_{\emptyset}(x) := \int_{[0,1]^d} f(y) dy = I(f)$. Note that the ANOVA decomposition is orthogonal, namely, $\int f_{\mathcal{S}}(x) f_{\mathcal{S}'}(x) dx = 0$, for $\mathcal{S} \neq \mathcal{S}'$. Let $\sigma^2(f)$ and

$\sigma_S^2(f)$ denote, respectively, the variance of f and f_S , that is, $\sigma^2(f) := \int_{[0,1]^d} (f(x) - I(f))^2 dx$, and $\sigma_S^2(f) := \int_{[0,1]^d} [f_S(x)]^2 dx$, whenever $|\mathcal{S}| > 0$. By using the above variance decomposition, we define the total variance for the subset $\mathcal{S}$ by

$$(3.1) \quad D_{\mathcal{S}}(f) := \sum_{\mathcal{S}' \subseteq \mathcal{S}} \sigma_{\mathcal{S}'}^2(f).$$

We are now in a position to define the effective dimension in the truncation sense by the smallest integer d' satisfying

$$(3.2) \quad D_{\{1,2,\dots,d'\}}(f) \geq p\sigma^2(f),$$

where p is a quantile of the variance. In other words, the first d' dimensions capture more than $p = 95\%$, say, of the total variations, even though the nominal dimension of f may be very large.

The efficiency of the QMC is intricately related to the effective dimension. Fix $N \in \mathbb{N}$, and consider a low-discrepancy point set $P := \{x^{(k)}\}_{k=1,\dots,N}$, where each point $x^{(k)} \in [0,1]^d$ is used for approximation of the integral $I(f)$. Then, the QMC error satisfies the bound

$$|I(f) - Q_N(f)| \leq \sum_{|\mathcal{S}| > 0} D_{\mathcal{S},N}(P_{\mathcal{S}}) \|f_{\mathcal{S}}\|,$$

where $P_{\mathcal{S}}$ indicates the projection of the point set P on $[0,1]^{|\mathcal{S}|}$, where $D_{\mathcal{S},N}(P_{\mathcal{S}})$ is the discrepancy of $P_{\mathcal{S}}$ consisting of N points, and where $\|f\|$ denotes the variation of the function f . The above inequality indicates that the QMC error bound associates with the uniformity of all the projections $P_{\mathcal{S}}$ and all the low-dimensional structures of $f_{\mathcal{S}}$ in an explicit manner. Although the success of the QMC method relies on a greater uniformity of low-discrepancy sequences than random sequences, the uniformity is not always preserved for every component and projections with a finite number of points. In fact, it is well known that the uniformity of low-discrepancy sequences decreases as the dimension increases. Nevertheless, as claimed in Wang and Fang [27], the QMC method can still be more effective than the MC method, in particular, for functions with a low effective dimension in the truncation sense. This observation can be justified by decomposing the bound as

$$|I(f) - Q_N(f)| \leq \sum_{\mathcal{S} \subseteq \{1,\dots,d'\}} D_{\mathcal{S},N}(P_{\mathcal{S}}) \|f_{\mathcal{S}}\| + \sum_{\{\mathcal{S}: \mathcal{S} \cap (\mathcal{A} \setminus \{1,\dots,d'\}) \neq \emptyset\}} D_{\mathcal{S},N}(P_{\mathcal{S}}) \|f_{\mathcal{S}}\|,$$

provided that the effective dimension of the function f is given by d' . This inequality indicates that when d' is small the discrepancy of all the low-dimensional projections of low-discrepancy point sets $P_{\mathcal{S}}$ are much smaller relative to those of the random point sets. This means that the first summation is much smaller for the QMC method than for the MC method. As the dimension increases further, the uniformity of low-discrepancy point sets may deteriorate, resulting in a higher value of $D_{\mathcal{S},N}(P_{\mathcal{S}})$. Yet the second sum may not be significant in the sense that the quantity $\|f_{\mathcal{S}}\|$ is often small. Consequently, the QMC method can be more effective than the MC method when the function f has a low effective dimension.

Let us next illustrate why the structure of series representations suits this effective dimension argument. There are two primal aspects: the domination of $\{\Gamma_k\}_{k \in \mathbb{N}}$ by a lower dimension of the interarrival times $\{E_k\}_{k \in \mathbb{N}}$, and the tails of Lévy measure.

For simplicity, we focus on the function $H(r)$ defined by (2.1). In view of (2.9), the first few terms of $\{H(\Gamma_k)\}_{k \in \mathbb{N}}$ are expressed solely by the first few terms of $\{E_k\}_{k \in \mathbb{N}}$ as well. For a simple illustration, fix $\lambda > 1$ and consider the infinite sum of a decreasing sequence,

$$(3.3) \quad \sum_{k=1}^{+\infty} \lambda^{-k+1} \Gamma_k \stackrel{\mathcal{L}}{=} \sum_{k=1}^{+\infty} \lambda^{-k+1} \sum_{l=1}^k E_l = \sum_{l=1}^{+\infty} E_l \sum_{k=l}^{+\infty} \lambda^{-k+1} = \sum_{l=1}^{+\infty} \frac{\lambda^{-l+2}}{\lambda-1} E_l,$$

where the interchange of two summations can be justified by the Fubini theorem due to the almost sure positivity of the summands, including the case where both sides are infinite. To see the impact of the first exponential E_1 , we compare (3.3) to its counterpart, with all the rest of $\{E_k\}_{k \geq 2}$ being set to be degenerate, in terms of the first and the second moments, that is,

$$(3.4) \quad \frac{\mathbb{E} \left[\sum_{k=1}^1 \frac{\lambda^{-k+2}}{\lambda-1} E_k \right]}{\mathbb{E} \left[\sum_{k=1}^{+\infty} \frac{\lambda^{-k+2}}{\lambda-1} E_k \right]} = 1 - \frac{1}{\lambda} \quad \text{and} \quad \frac{\text{Var} \left(\sum_{k=1}^1 \frac{\lambda^{-k+2}}{\lambda-1} E_k \right)}{\text{Var} \left(\sum_{k=1}^{+\infty} \frac{\lambda^{-k+2}}{\lambda-1} E_k \right)} = 1 - \frac{1}{\lambda^2}.$$

Here, a greater λ assigns a more significant role to the lower dimension of the interarrival times $\{E_k\}_{k \in \mathbb{N}}$. Roughly speaking, this tends to happen with a faster decay of the function $H(r)$ in r , that is, less activity near the origin of Lévy measure.

Next, concerning the Lévy measure tails, recall that most of the Lévy measure tails tends to be expressed by the first few terms of the sequence $\{H(\Gamma_k)\}_{k \in \mathbb{N}}$. The Lévy measure tails are closely related to moments of its associated infinitely divisible random vector in such a way that for a submultiplicative, locally bounded, measurable function $f : \mathbb{R}^d \mapsto [0, +\infty)$ the expectation $\mathbb{E}[f(X)]$ is well defined if and only if $\int_{\|z\| > 1} f(z) \nu(dz)$ is finite. (For example, the functions $\|x\|^k \vee 1$ and $e^{\langle c, x \rangle}$ are submultiplicative. See Theorem 25.3 and its surrounding arguments of Sato [24] for details.) In this respect, the simulation of the largest jumps through a systematic generation of the lower dimension of the interarrival times $\{E_k\}_{k \in \mathbb{N}}$ is expected to contribute to improvements in both precision and convergence when computing expectations.

The effectiveness of the QMC method may be increased by combining this structure with the superior uniformity of a low-discrepancy sequence in earlier dimensions. Accordingly, it is expected that the QMC method may work better than the MC method. The efficiency of the QMC method is often measured through the ratio

$$\frac{D_S(f)}{\sigma^2(f)},$$

which we refer to as the cumulative explanatory ratio (CER), where $D_S(f)$ is defined by (3.1). In other words, this represents the proportion of the variance captured by the first d dimensions. In view of the inequality (3.2), it is possible that the effective dimension is given by d when the CER becomes larger than p . In general, it is difficult to provide an explicit expression for $D_S(f)$. However, as shown in Sobol' [26] and [27], this can be computed via the formula

$$D_S(f) = \int_{[0,1]^{2d-|S|}} f(x) f(y_{A \setminus S}) dx dy - [I(f)]^2,$$

where $x := (x_S, x_{A \setminus S})$ and $y := (y_S, y_{A \setminus S})$. This allows us to estimate $D_{\{1,2,\dots,d\}}$ by the MC method. Let us now translate this into our framework. For simplicity, we

restrict ourselves to an infinitely divisible random variable; that is, the function f here maps $\mathbb{R}$ to $\mathbb{R}$. Denote by $\{E'_k\}_{k \in \mathbb{N}}$ an i.i.d. copy of the interarrival times $\{E_k\}_{k \in \mathbb{N}}$, and write $\mathfrak{U} := \{U_k\}_{k \in \mathbb{N}}$ and, for each $n \in \mathbb{N}$,

$$\mathfrak{E}_n := \{E_1, \dots, E_n, E'_{n+1}, E'_{n+2}, \dots\},$$

where $\mathfrak{E}_0 := \{E_k\}_{k \in \mathbb{N}}$. Then, for natural numbers n and N such that $n \leq N$, define, in view of (2.9),

$$\begin{aligned} X_f(\mathfrak{E}_n, \mathfrak{U}) &:= f\left(\sum_{k=1}^N \left[H\left(\sum_{l=1}^{k \wedge n} E_l + \sum_{l=(k \wedge n)+1}^k E'_l, U_k\right) - c_k\right]\right) \\ &\stackrel{\mathcal{L}}{=} f\left(\sum_{k=1}^N [H(\Gamma_k, U_k) - c_k]\right). \end{aligned}$$

Then, the CER with respect to the first n interarrival times is defined by

$$(3.5) \quad (\text{CER})_n := \frac{\text{Cov}(X_f(\mathfrak{E}_n, \mathfrak{U}), X_f(\mathfrak{E}_0, \mathfrak{U}))}{\text{Var}(X_f(\mathfrak{E}_0, \mathfrak{U}))}.$$

We note that the sequence $\{U_k\}_{k \in \mathbb{N}}$ appears in common on both sides of the covariance operator. As a toy example, the CER of the random variable (3.3) is given by $(\text{CER})_n = 1 - \lambda^{-2n}$, since

$$\text{Cov}\left(\sum_{k=1}^n \frac{\lambda^{-k+2}}{\lambda-1} E_k + \sum_{k=n+1}^{+\infty} \frac{\lambda^{-k+2}}{\lambda-1} E'_k, \sum_{k=1}^{+\infty} \frac{\lambda^{-k+2}}{\lambda-1} E_k\right) = \text{Var}\left(\sum_{k=1}^n \frac{\lambda^{-k+2}}{\lambda-1} E_k\right).$$

4. Numerical illustrations. In this section, we provide numerical results of estimations of moments and tail probabilities for the stable and gamma random variables presented in Examples 2.1 and 2.2. Let us begin by formulating five statistics we will report in the following numerical examples for discussion. Ideally, we are able to estimate $\mu := \mathbb{E}[f(X)]$, where $X = \sum_{k=1}^{+\infty} [H(\Gamma_k, U_k) - c_k]$ and f is a function of interest; for example, $f(x) = x^2$ for estimation of the second moment. In reality, as discussed, we need to truncate infinite shot noise series up to a finite number of terms; that is, we instead estimate $\mu_N := \mathbb{E}[f(X_N)]$, where $X_N = \sum_{k=1}^N [H(\Gamma_k, U_k) - c_k]$. Clearly, $\lim_{N \uparrow +\infty} \mu_N = \mu$. Now, let $\{X_N^{(k_1, k_2)}\}_{(k_1, k_2) \in \mathbb{N}^2}$ be an array of i.i.d. random variables such that $X_N^{(1,1)} \stackrel{\mathcal{L}}{=} X_N$. For $k_2 = 1, \dots, N_2$, we compute the sample mean $\mu_N^{(k_2)} := N_1^{-1} \sum_{k_1=1}^{N_1} f(X_N^{(k_1, k_2)})$. Then, we compute average, standard error, root mean squared error (RMSE), and relative error defined, respectively, by

$$\begin{aligned} (4.1) \quad (\text{average}) &= \frac{1}{N_2} \sum_{k_2=1}^{N_2} \mu_N^{(k_2)} =: \tilde{\mu}_N, \\ (\text{standard error}) &= \left(\frac{1}{N_2(N_2-1)} \sum_{k_2=1}^{N_2} \left(\mu_N^{(k_2)} - \tilde{\mu}_N \right)^2 \right)^{1/2}, \\ (\text{RMSE}) &:= \left(\frac{1}{N_2} \sum_{k_2=1}^{N_2} \left(\mu_N^{(k_2)} - \mu \right)^2 \right)^{1/2}, \\ (\text{relative error}) &:= \left| \frac{\tilde{\mu}_N - \mu}{\mu} \right|. \end{aligned}$$

Note that we consider only examples for which the limiting value μ is available. Note also that the RMSE and the relative error are affected by both the estimator accuracy and the finite truncation of infinite shot noise series. The fifth statistic is the effective dimension by estimating the CER (3.5).

In our numerical experiments, we fix $N_1 = 4096$ and $N_2 = 30$ and use a scrambled version of Sobol' sequences [25], proposed in Owen [20]. The reason for $N_1 = 4096 = 2^{12}$ is because the Sobol' sequence, which is a (t, m, s) -net in base 2, attains a better uniformity when sample size is in the power of 2 (see Niederreiter [18]). Furthermore, we employ the Latin supercube sampling method of Owen [19] to avoid generating more than 50-dimensional Sobol' sequences. For example, 500-dimensional low-discrepancy sequences are constructed as 10 sets of 50-dimensional (randomized) low-discrepancy sequences. This randomization scheme makes full sense of the standard error (4.1) and has often been employed, for example, in [12, 13] for error analysis under the uniformity of low-discrepancy sequence in the randomized QMC (RQMC) framework.

For comparison, we employ the same procedure for estimating those reference numbers with the MC method, except that low-discrepancy sequences are replaced by pseudorandom sequences. All the experiments are implemented on the computer platform Intel Xeon(R) CPU 3.00GHz with 3.25GB memory. It takes around 437 seconds to generate and store 4096×1000 random variables in the MC simulation. Meanwhile, in the RQMC method, it takes around 3593 seconds to implement 4096 sets of 1000-dimensional Latin supercube sampling sequences. Therefore, relative to the plain MC experiments, implementation of the RQMC method increases computing time by roughly a factor of 8.

4.1. One-sided stable random variable. Let X be a stable random variable in $\mathbb{R}$ with series representation

$$(4.2) \quad X := \sum_{k=1}^{+\infty} \left(\frac{\alpha \Gamma_k}{b} \right)^{-1/\alpha},$$

where $\alpha \in (0, 1)$. This corresponds to $d = 1$ and $\sigma(d\xi) = b\delta_{\{1\}}(d\xi)$ in the formulation (2.3). In our numerical experiments, we use the truncation (2.7) up to a finite number of terms

$$X_N := \sum_{k=1}^N \left(\frac{\alpha \Gamma_k}{b} \right)^{-1/\alpha},$$

where the truncation level N is supposed to be set sufficiently large so that X_N is close enough to the infinite sum X . As a reference of the closeness, consider a different truncation $X_{(N)}$ up to a finite time span of a standard Poisson process in the spirit of (2.8). As discussed earlier, the $L^2(\Omega)$ -distance is given by

$$(4.3) \quad \mathbb{E} \left[|X - X_{(N)}|^2 \right] = \frac{b}{2-\alpha} \left(\frac{\alpha N}{b} \right)^{1-2/\alpha} + \left(\frac{b}{1-\alpha} \right)^2 \left(\frac{\alpha N}{b} \right)^{2-2/\alpha},$$

whose decay is of $O(N^{2-2/\alpha})$ as $N \uparrow +\infty$. This implies that the truncation $X_{(N)}$ converges in $L^2(\Omega)$ to the infinite sum X at a slower rate for a greater α .

First, we examine the effectiveness of the RQMC method applied to Poisson interarrival times in moment estimation, where moments are given in closed form

(Example 25.10 and Example 24.12 of Sato [24]); for $\eta \in (-\infty, \alpha)$,

$$(4.4) \quad \mathbb{E}[X^\eta] = \left(\frac{b\Gamma(1-\alpha)}{\alpha} \right)^{\eta/\alpha} \frac{\Gamma(1-\eta/\alpha)}{\Gamma(1-\eta)}.$$

As can be seen in (4.2) and (4.4), the parameter b serves only as a constant scale. Hence, without loss of generality, we fix $b = 1$. Our simulation results are based upon parameter sets $\alpha = \{0.1, 0.4, 0.7\}$ and $\eta = \{-0.2, +\alpha/4\}$. Note that X^η is square integrable under those parameter sets and thus the variance $\text{Var}(X^\eta) = \mathbb{E}[X^{2\eta}] - (\mathbb{E}[X^\eta])^2$ is well defined.

TABLE 1
Moment estimation for one-sided stable random variables.

α	$\mathbb{E}[X^{-0.2}]$	N	Average		Standard error		RMSE		Relative error	
			MC	RQMC	MC	RQMC	MC	RQMC	MC	RQMC
0.1	0.019075	10	0.018974	0.019078	1.21E-04	7.35E-06	5.62E-05	3.40E-06	0.53%	0.02%
		100	0.019037	0.019077	1.26E-04	7.34E-06	5.80E-05	3.39E-06	0.20%	0.01%
		1000	0.019003	0.019077	1.13E-04	7.34E-06	5.23E-05	3.39E-06	0.38%	0.01%
0.4	0.500238	10	0.503474	0.503948	6.34E-04	3.33E-06	4.02E-04	3.18E-04	0.65%	0.74%
		100	0.501221	0.500367	6.93E-04	2.72E-06	3.30E-04	1.11E-05	0.20%	0.03%
		1000	0.499978	0.500237	6.54E-04	3.24E-06	3.02E-04	1.50E-06	0.05%	0.00%
0.7	0.647097	10	0.690219	0.690498	4.66E-04	2.27E-06	3.70E-03	3.71E-03	6.66%	6.71%
		100	0.661708	0.660954	4.26E-04	2.84E-06	1.27E-03	1.19E-03	2.26%	2.14%
		1000	0.651962	0.651980	3.98E-04	2.42E-06	4.55E-04	4.18E-04	0.75%	0.75%
α	$\mathbb{E}[X^{\alpha/4}]$	N	Average		Standard error		RMSE		Relative error	
			MC	RQMC	MC	RQMC	MC	RQMC	MC	RQMC
0.1	2.182717	10	2.183678	2.182594	2.36E-03	2.31E-04	1.09E-03	1.07E-04	0.04%	0.01%
		100	2.178075	2.182595	2.48E-03	2.31E-04	1.21E-03	1.07E-04	0.21%	0.01%
		1000	2.181600	2.182595	2.26E-03	2.31E-04	1.04E-03	1.07E-04	0.05%	0.01%
0.4	1.592868	10	1.590911	1.590189	1.65E-03	1.64E-04	7.80E-04	2.41E-04	0.12%	0.17%
		100	1.589506	1.592699	1.67E-03	1.64E-04	8.24E-04	7.70E-05	0.21%	0.01%
		1000	1.592077	1.592785	1.54E-03	1.64E-04	7.15E-04	7.60E-05	0.05%	0.01%
0.7	1.549229	10	1.487895	1.487339	1.38E-03	1.44E-04	5.29E-03	5.30E-03	3.96%	3.99%
		100	1.525699	1.528252	1.30E-03	1.44E-04	2.10E-03	1.80E-03	1.52%	1.35%
		1000	1.540968	1.541631	1.19E-03	1.44E-04	8.96E-04	6.54E-04	0.53%	0.49%

Numerical results of moment estimations are presented in Table 1. To illustrate the differences between average and analytical value, both RMSEs and relative errors are also reported. Each relative error is given by a percentile of the difference between the estimated average and the analytical value. The speed of convergence depends on the parameter value α . The relative errors indicate that each average value tends to be very close to its limiting value even when N is very small for smaller α , while it is slower for larger α . This observation is consistent with the speed of the $L^2(\Omega)$ -convergence presented in (4.3). Most importantly, we can assert that the RQMC method produces a smaller standard error than the MC method by a factor in the range of 10–200. Recall that, in plain MC simulation, 100-times more replications are required, in principle, to obtain a 10-times more accurate estimate. In comparing the RMSE values, it can also be observed that each moment estimate via the RQMC method tends to converge to its true limiting value $\mathbb{E}[X^\eta]$ much faster than the MC method. The difference in relative error is mainly attributed to an inefficiency of the MC method; that is, the sample size 4096×30 is not sufficient for estimating the average. In fact, this also asserts the superiority of the RQMC method.

Note that the standard errors of the RQMC method are almost invariant under an increase of the numerical dimension N . We conjecture that $\text{Var}(X_{10})$, $\text{Var}(X_{100})$, and $\text{Var}(X_{1000})$ are close to each other. If so, the effective dimension remains small even in the high nominal-dimensional problem such as $N = 1000$.

Let us next examine the cumulative explanatory ratio (CER) $_n$ to analyze the impact of the RQMC method in terms of effective dimension. Each CER is estimated

through (3.5) by a plain MC method with 1000000 replications. By computing this quantity, we examine the variance contribution of the first n dimensions of $\{E_k\}_{k \in \mathbb{N}}$. We report numerical results in Table 2, where all the numbers are rounded off to the nearest integer. (Hence, “100%” may be any number greater than 99.49%, not necessarily exactly 100%.) Here, we present only results of the case where the truncation level is set at $N = 1000$ to avoid overloading this section with nonessential details. (Let us note that we have obtained very similar results even for smaller N ’s.) Let us emphasize that the CERs achieve almost 100% at the first dimension. In other words, the effective dimension with $p = 0.99$ is equal to one without additional variance reduction techniques, even when dealing with the high truncation level $N = 1000$. This indicates that the RQMC method remains very effective even with a very high truncation level at which the RQMC method often ruins its effectiveness against a plain MC method. Those CER results are also consistent with the superiority of the RQMC method in terms of standard error and support our claim that the RQMC method has an advantage over the MC method for shot noise series representations of infinitely divisible laws.

TABLE 2
(CER) $_n$ of moment estimation for one-sided stable random variables.

η	α	(CER) $_n$				
		1	2	3	4	5
−0.2	0.1	100%	100%	100%	100%	100%
	0.4	98%	100%	100%	100%	100%
	0.7	96%	99%	99%	100%	100%
+ $\alpha/4$	0.1	100%	100%	100%	100%	100%
	0.4	100%	100%	100%	100%	100%
	0.7	99%	100%	100%	100%	100%

Next, we examine the effectiveness of the RQMC method in the probability tail asymptotics (Property 1.2.15 of Samorodnitsky and Taqqu [23]):

$$(4.5) \quad \lim_{\lambda \uparrow +\infty} \lambda^\alpha \mathbb{P}(X > \lambda) = \frac{b}{\alpha}.$$

We again fix $b = 1$, and the parameter sets tested are $(\alpha, \lambda) = (0.1, 10^{35})$, $(0.4, 10^{10})$, and $(0.7, 10^5)$. We first present results of the CER in Table 3. Let us emphasize that they are all exactly 100% without rounding operations, unlike the results presented in Table 2. This is not surprising at all, since, as discussed in [23], the first term of the series (4.2) dominates the tail probability behavior in such a way that

$$\lambda^\alpha \mathbb{P} \left(\left(\frac{\alpha \Gamma_1}{b} \right)^{-1/\alpha} > \lambda \right) = \lambda^\alpha \mathbb{P} \left(V_1 < \frac{b}{\alpha} \lambda^{-\alpha} \right) = \lambda^\alpha \left(1 - e^{-\frac{b}{\alpha} \lambda^{-\alpha}} \right) \rightarrow \frac{b}{\alpha}$$

as $\lambda \uparrow +\infty$. In addition, recall that the first exponential E_1 also appears in all the remaining terms of the series, not only in the first term. This fact provides a stronger reason for the result $(\text{CER})_1 = 100\%$.

Numerical results of tail probability estimation are presented in Table 4. Note that b/α is the limiting value in terms of both the truncation level N and the threshold λ . On the one hand, we can observe through the standard errors and the RMSEs that the RQMC method yields estimates with very small variance, regardless of the truncation level N . Clearly, this fact is attributed to the great uniformity of 4096 samples

TABLE 3

 $(\text{CER})_n$ of probability tail estimation for one-sided stable random variables.

(α, λ)	$(\text{CER})_n$				
	1	2	3	4	5
$(0.1, 10^{35})$	100%	100%	100%	100%	100%
$(0.4, 10^{10})$	100%	100%	100%	100%	100%
$(0.7, 10^5)$	100%	100%	100%	100%	100%

TABLE 4

Tail probability estimation for one-sided stable random variables.

$(\alpha, \lambda) = (0.1, 10^{35})$		Average		Standard error		RMSE		Relative error	
b/α	N	MC	RQMC	MC	RQMC	MC	RQMC	MC	RQMC
10	1	10.47402	10.01079	4.54E-01	2.57E-02	2.13E-01	1.19E-02	4.74%	0.11%
	10	10.01079	10.01079	5.29E-01	2.57E-02	2.44E-01	1.19E-02	0.11%	0.11%
	100	10.52548	10.01079	4.16E-01	2.57E-02	1.97E-01	1.19E-02	5.25%	0.11%
$(\alpha, \lambda) = (0.4, 10^{10})$		Average		Standard error		RMSE		Relative error	
b/α	N	MC	RQMC	MC	RQMC	MC	RQMC	MC	RQMC
2.5	1	2.929687	2.522786	5.16E-01	8.14E-02	2.40E-01	3.76E-02	17.19%	0.91%
	10	2.278646	2.522786	5.85E-01	8.14E-02	2.70E-01	3.76E-02	8.85%	0.91%
	100	2.522786	2.522786	3.79E-01	8.14E-02	1.75E-01	5.01E-03	0.91%	0.91%
$(\alpha, \lambda) = (0.7, 10^5)$		Average		Standard error		RMSE		Relative error	
b/α	N	MC	RQMC	MC	RQMC	MC	RQMC	MC	RQMC
1.428571	1	1.749958	1.415407	2.03E-01	5.34E-02	9.73E-02	2.46E-02	22.50%	0.92%
	10	1.492612	1.415407	2.48E-01	5.34E-02	1.14E-01	2.46E-02	4.48%	0.92%
	100	1.569816	1.415407	1.55E-01	5.34E-02	7.23E-02	2.46E-02	9.89%	0.92%

in the first dimension, although they are scrambled. It seems that the uniformity contributes to stabilizing the number of samples falling beyond the threshold λ . In the MC experiments, on the other hand, different (independent) 4096 replications are sampled for different simulation rounds. Hence, as seen in Table 4, estimates tend to fluctuate more through the MC method.

4.2. Gamma random variable. Let $\{V_k\}_{k \in \mathbb{N}}$ be a sequence of i.i.d. exponential random variables with unit mean, independent of $\{\Gamma_k\}_{k \in \mathbb{N}}$. We have seen in Example 2.2 that for $a > 0$ and $b > 0$ the random variable

$$(4.6) \quad X := \sum_{k=1}^{+\infty} e^{-\Gamma_k/a} \frac{V_k}{b}$$

has a gamma distribution. It is straightforward that, for $\eta \in (-a, +\infty)$,

$$\mathbb{E}[X^\eta] = b^{-\eta} \frac{\Gamma(a + \eta)}{\Gamma(a)}$$

and that, for $\lambda > 0$,

$$\mathbb{P}(X > \lambda) = \frac{1}{\Gamma(a)} \int_{b\lambda}^{+\infty} z^{a-1} e^{-z} dz,$$

which can be computed by using an incomplete gamma function in standard math packages. As is the case with the one-sided stable random variable, the parameter b serves only as a constant scale. Hence, without essential loss of generality, we fix $b = 1$.

4.2.1. RQMC method only for Poisson interarrival times. First, we apply the RQMC method to the Poisson interarrival times $\{E_k\}_{k \in \mathbb{N}}$, while the independent

exponential random variables $\{V_k\}_{k \in \mathbb{N}}$ are still governed by the plain MC method. Numerical results of moment estimation are presented in Table 5. The truncation level is set at $N = 1000$ throughout the experiments. In short, those results indicate that the RQMC method improves the plain MC method in standard error only by a factor in the range of 2–4. (We have confirmed that even much smaller truncation levels, say $N = 50$, provide very similar results.) The improvements here are not as appealing as in the case of one-sided stable random variables examined in section 4.1.

TABLE 5

Moment estimation for gamma random variables; RQMC method for $\{E_k\}_{k \in \mathbb{N}}$ and MC method for $\{V_k\}_{k \in \mathbb{N}}$.

a	$\mathbb{E}[X^{-0.2}]$	Average		Standard error		RMSE		Relative error	
		MC	RQMC	MC	RQMC	MC	RQMC	MC	RQMC
0.5	1.687812	1.683850	1.688163	4.16E-03	1.63E-03	1.95E-03	7.53E-04	0.23%	0.02%
1.0	1.164230	1.163710	1.164274	1.03E-03	4.04E-04	4.75E-04	1.86E-04	0.04%	0.00%
1.5	1.012687	1.012363	1.013300	6.01E-04	3.24E-04	2.78E-04	1.58E-04	0.03%	0.06%

a	$\mathbb{E}[X^1]$	Average		Standard error		RMSE		Relative error	
		MC	RQMC	MC	RQMC	MC	RQMC	MC	RQMC
0.5	0.5	0.497839	0.499543	2.61E-03	1.48E-03	1.22E-03	6.84E-04	0.43%	0.09%
1.0	1.0	0.998212	0.998301	3.44E-03	1.83E-03	1.59E-03	8.54E-04	0.18%	0.17%
1.5	1.5	1.498708	1.496747	3.98E-03	2.55E-03	1.84E-03	1.21E-03	0.09%	0.22%

The effectiveness of the RQMC method is somewhat “watered down” by the purely random element $\{V_k\}_{k \in \mathbb{N}}$. To confirm this conjecture, consider the random variable

$$(4.7) \quad Y := \sum_{k=1}^{+\infty} e^{-\Gamma_k/a} \frac{1}{b};$$

that is, all the $\{V_k\}_{k \in \mathbb{N}}$ in (4.6) are frozen at 1. We can prove that the random variable Y of (4.7) is an infinitely divisible random variable with Lévy measure $\nu(dz) = a/zdz$ defined only on $(0, 1/b)$ and $\mathbb{E}[Y] = a/b$. As a matter of course, this random variable is not directly comparable with gamma random variables in terms of estimator efficiency since they have different distributions. For example, their variances are not identical, while their means coincide. Nevertheless, the series representation (4.7) provides a useful intuition as to how much the exponential random variables $\{V_k\}_{k \in \mathbb{N}}$ in (4.6) have ruined the effectiveness of the RQMC method applied to $\{E_k\}_{k \in \mathbb{N}}$. To avoid overloading this section with nonessential details, we give in Table 6 only numerical results of moment estimation of $\mathbb{E}[Y] = a/b$. For a better comparison with the results presented in Table 5, we also fix $b = 1$ and the truncation level $N = 1000$. Those results indicate a sufficient efficiency of the RQMC method against the MC method. Based upon the standard errors, the RQMC method provides more accurate values by a factor in the range of 37–67, and similar results are observed in the RMSE comparison.

TABLE 6

Moment estimation for random variable (4.7); RQMC method for $\{E_k\}_{k \in \mathbb{N}}$.

a	$\mathbb{E}[Y]$	Average		Standard error		RMSE		Relative error	
		MC	RQMC	MC	RQMC	MC	RQMC	MC	RQMC
0.5	0.5	0.499489	0.500043	1.22E-03	1.83E-05	5.65E-04	9.19E-06	0.10%	0.01%
1.0	1.0	0.999498	1.000058	1.78E-03	4.51E-05	8.19E-04	2.14E-05	0.05%	0.01%
1.5	1.5	1.499562	1.500042	2.24E-03	5.95E-05	1.03E-03	2.77E-05	0.03%	0.00%

4.2.2. Improvement through application of RQMC method to every random sequence. In generating gamma variates in section 4.2.1, we generated $\{E_k\}_{k \in \mathbb{N}}$ using a low-discrepancy sequence, while a random sequence is used to generate $\{V_k\}_{k \in \mathbb{N}}$. The previous example of (4.7) indicates that the additional i.i.d. exponential variates $\{V_k\}_{k \in \mathbb{N}}$ in (4.6) have blurred the effectiveness of the RQMC method applied to $\{E_k\}_{k \in \mathbb{N}}$. In this respect, we might be able to enhance the accuracy of simulation by appropriately assigning earlier dimensions of low-discrepancy sequences to $\{E_k\}_{k \in \mathbb{N}}$ and $\{V_k\}_{k \in \mathbb{N}}$ so that the effective dimension is decreased. From the ANOVA decomposition, it is evident that the performance of the RQMC method is affected by how to assign each dimension of low-discrepancy sequence. To see this, let us compare the following four schemes to examine estimator efficiency of the MC and RQMC methods.

Scheme I. Plain MC method: Both sequences $\{E_k\}_{k \in \mathbb{N}}$ and $\{V_k\}_{k \in \mathbb{N}}$ are generated by i.i.d. random variables.

Scheme II. Hybrid: We use the RQMC method only for $\{E_k\}_{k \in \mathbb{N}}$, while we use the MC method for $\{V_k\}_{k \in \mathbb{N}}$.

Scheme III. RQMC($k, k + N$): Prepare a $2N$ -dimensional (randomized) low-discrepancy sequence. The first N -dimensional points are assigned to $\{E_k\}_{k \in \mathbb{N}}$, while the second N -dimensional points are assigned to $\{V_k\}_{k \in \mathbb{N}}$. More precisely, we set $(E_k, V_k) \leftarrow (LD_k, LD_{k+N})$ for $k = 1, \dots, N$, where LD_k indicates the k th component of low-discrepancy sequence.

Scheme IV. RQMC($2k - 1, 2k$): Every odd component of the $2N$ -dimensional sequence is used for $\{E_k\}_{k \in \mathbb{N}}$, while every even component is used for $\{V_k\}_{k \in \mathbb{N}}$. More precisely, we set $(E_k, V_k) \leftarrow (LD_{2k-1}, LD_{2k})$ for $k = 1, \dots, N$.

Numerical results of moment estimation for gamma variates are reported in Table 7. As in the previous examples, the truncation level is fixed at $N = 1000$. All the numerical results support the effectiveness of the RQMC method (Schemes II, III, and IV) over the plain MC method (Scheme I). The effectiveness of the RQMC method is also observed by comparing Schemes II and III, where $\{V_k\}_{k \in \mathbb{N}}$ is generated, respectively, by the MC and the RQMC methods. Amongst the four schemes, Scheme IV yields the most accurate estimates. The superiority of Scheme IV can be explained in terms of effective dimension. We present in Table 8 a CER of moment estimation for gamma random variables. To distinguish from $(\text{CER})_n$ of (3.5), we define an appropriate CER for Scheme IV as follows. Let $\{V'_k\}_{k \in \mathbb{N}}$ be an i.i.d. copy of $\{V_k\}_{k \in \mathbb{N}}$ and also be independent of $\{E_k\}_{k \in \mathbb{N}}$. Define, for $n \in \mathbb{N}$,

$$\mathfrak{U}_n := \{V_1, \dots, V_n, V'_{n+1}, V'_{n+2}, \dots\},$$

where $\mathfrak{U}_0 := \{V_k\}_{k \in \mathbb{N}}$. Then, define a CER, slightly different from (3.5), by

$$(4.8) \quad (\text{CER})_{(m,n)} := \frac{\text{Cov}(X_f(\mathfrak{E}_m, \mathfrak{U}_n), X_f(\mathfrak{E}_0, \mathfrak{U}_0))}{\text{Var}(X_f(\mathfrak{E}_0, \mathfrak{U}_0))}.$$

Recall that in computing $(\text{CER})_n$ via (3.5) a common i.i.d. random sequence $\mathfrak{U}_0$ is plugged into both sides of the covariance operator.

We present in Table 8 numerical results of $(\text{CER})_n$ and $(\text{CER})_{(m,n)}$, again with rounding operations. The results provide clear evidence that Scheme IV significantly increases the CER with the first few dimensions, relative to Scheme III. Surprisingly, a simple use of the RQMC method without additional variance reduction techniques can capture almost 90% of the total variance only with the first five dimensions out of as many as 1000 nominal dimensions if the sequence is assigned in a smart way.

TABLE 7
Moment estimation for gamma random variables through four schemes.

a	$\mathbb{E}[X^{-0.2}]$	Scheme	Average	Standard error	RMSE	Relative error
0.5	1.687812	I	1.683850	4.16E-03	1.95E-03	0.23%
		II	1.688163	1.63E-03	7.53E-04	0.02%
		III	1.688572	1.56E-03	7.22E-04	0.05%
		IV	1.685838	5.68E-04	3.12E-04	0.12%
1.0	1.164230	I	1.163710	1.03E-03	4.75E-04	0.04%
		II	1.164274	4.04E-04	1.86E-04	0.00%
		III	1.164073	2.74E-04	1.27E-04	0.01%
		IV	1.164230	8.90E-05	4.10E-05	0.00%
1.5	1.012687	I	1.012363	6.01E-04	2.78E-04	0.03%
		II	1.013300	3.24E-04	1.58E-04	0.06%
		III	1.012563	1.57E-04	7.31E-05	0.01%
		IV	1.012661	3.16E-05	1.47E-05	0.00%
a	$\mathbb{E}[X^1]$	Scheme	Average	Standard error	RMSE	Relative error
0.5	0.5	I	0.497839	2.61E-03	1.22E-03	0.43%
		II	0.499543	1.48E-03	6.84E-04	0.09%
		III	0.502280	8.86E-04	4.52E-04	0.46%
		IV	0.499953	7.07E-05	3.28E-05	0.01%
1.0	1.0	I	1.163710	1.03E-03	4.75E-04	0.04%
		II	1.164274	4.04E-04	1.86E-04	0.00%
		III	1.164073	2.74E-04	1.27E-04	0.01%
		IV	1.164230	8.90E-05	4.10E-05	0.00%
1.5	1.5	I	1.498708	3.98E-03	1.84E-03	0.09%
		II	1.496747	2.55E-03	1.21E-03	0.22%
		III	1.502389	1.07E-03	5.33E-04	0.16%
		IV	1.499717	1.90E-04	9.10E-05	0.02%

TABLE 8
Estimated CER for the moment estimations presented in percentage.

		$(\text{CER})_n$					$(\text{CER})_{(m,n)}$				
η	a	1	2	3	4	5	(1,0)	(1,1)	(2,1)	(2,2)	(3,2)
-0.2	0.5	90%	90%	91%	95%	95%	90%	95%	99%	100%	100%
	1.0	67%	75%	76%	76%	77%	67%	80%	91%	95%	98%
	1.5	51%	63%	66%	67%	68%	51%	66%	81%	89%	93%
1	0.5	39%	48%	50%	50%	50%	39%	79%	88%	96%	97%
	1.0	33%	45%	48%	49%	50%	33%	66%	77%	89%	92%
	1.5	28%	41%	46%	49%	49%	28%	57%	69%	82%	87%

5. Concluding remarks. In this paper, we have proposed an effective application of QMC methods to shot noise series representations of infinitely divisible random vectors. We have observed that the largest jumps consisting of an infinitely divisible random vector are expressed only by the first few terms of the series and that this structure tends to decrease the effective dimension and thus increase the efficiency of the QMC method, thanks to the greater uniformity of low-discrepancy sequence in the lower dimension. We have illustrated the effectiveness of our approach through numerical results of moment and tail probability estimations for stable and gamma random variables. We expect that our results support a simulation use of shot noise

series representations of various infinitely divisible laws with ever increasing computational speed.

Let us close by describing some related research topics. First, we have discussed that an explicit representation is rarely available through the inverse Lévy measure method, while some other methods may provide a representation in closed form, as seen in Examples 2.2 and 2.3. Nevertheless, additional random sequences are required for such representations and ruin the effectiveness of the QMC method applied to Poisson interarrival times, as demonstrated in sections 4.2.1 and 4.2.2. In particular, the representation for CGMY random variables in Example 2.3 contains so many additional random sequences that our approach in section 4.2.2 would no longer be effectively applicable. To address this issue, the authors have recently developed in [10] a numerical method for implementing the inverse Lévy measure method on a computer. This is worthwhile for two reasons: (i) the effectiveness of the QMC method is not watered down by other random sequences, and (ii) in principle, any infinitely divisible random vector can then be simulated with shot noise series representation.

Second, it is well known that the series representation can be extended to Lévy processes on a compact time interval simply by scattering all the jumps uniformly on the time interval, that is,

$$\{X_t : t \in [0, T]\} \stackrel{\mathcal{L}}{=} \left\{ \sum_{k=1}^{+\infty} [H(\Gamma_k, U_k) \mathbf{1}(T_k \in [0, t]) - tc_k] : t \in [0, T] \right\},$$

where $\{T_k\}_{k \in \mathbb{N}}$ is a sequence of i.i.d. uniform random variables on $[0, T]$. This representation is known to be useful in, for example, simulation of trajectories of Lévy processes, weak approximation of Lévy-driven stochastic differential equations, and simulation of fractional Lévy motions (see, for example, Houdré and Kawai [8]). Here, QMC techniques can also be applied to the jump timings $\{T_k\}_{k \in \mathbb{N}}$, not only to the jump size driving $\{\Gamma_k\}_{k \in \mathbb{N}}$, just as in section 4.2.2.

Acknowledgment. Part of this work was carried out while RK was based at the Center for the Study of Finance and Insurance, Osaka University, Japan.

REFERENCES

- [1] S. ASMUSSEN AND P. W. GLYNN, *Stochastic Simulation*, Springer-Verlag, New York, 2007.
- [2] S. ASMUSSEN AND J. ROSIŃSKI, *Approximations of small jumps of Lévy processes with a view towards simulation*, J. Appl. Probab., 38 (2001), pp. 482–493.
- [3] L. BONDESSON, *On simulation from infinitely divisible distributions*, Adv. in Appl. Probab., 4 (1982), pp. 855–869.
- [4] P. CARR, H. GEMAN, D. B. MADAN, AND M. YOR, *The fine structure of asset returns: An empirical investigation*, J. Business, 75 (2002), pp. 303–325.
- [5] R. CAFLISCH, W. MOROKOFF, AND A. OWEN, *Valuation of mortgaged-backed securities using Brownian bridges to reduce effective dimension*, J. Comput. Finance, 1 (1997), pp. 27–46.
- [6] T. S. FERGUSON AND M. J. KLASS, *A representation of independent increment processes with Gaussian components*, Ann. Math. Statist., 43 (1972), pp. 1634–1643.
- [7] W. HÖRMANN AND J. LEYDOLD, *Continuous random variate generation by fast numerical inversion*, ACM Trans. Model. Comput. Simul., 13 (2003), pp. 347–362.
- [8] C. HOUDRÉ AND R. KAWAI, *On fractional tempered stable motion*, Stochastic Process. Appl., 116 (2006), pp. 1161–1184.
- [9] C. HOUDRÉ AND R. KAWAI, *On layered stable processes*, Bernoulli, 13 (2007), pp. 252–278.
- [10] J. IMAI AND R. KAWAI, *Numerical Inverse Lévy Measure Method for Infinite Shot Noise Series Representation*, <http://sites.google.com/site/reiichirokawai/> (2010).
- [11] J. IMAI AND R. KAWAI, *On Finite Truncation of Infinite Shot Noise Series Representation of Tempered Stable Laws*, <http://sites.google.com/site/reiichirokawai/> (2010).

- [12] J. IMAI AND K. S. TAN, *A general dimension reduction technique for derivative pricing*, J. Comput. Finance, 10 (2006), pp. 129–155.
- [13] J. IMAI AND K. S. TAN, *An accelerating quasi-Monte Carlo method for option pricing under the generalized hyperbolic Lévy process*, SIAM J. Sci. Comput., 31 (2009), pp. 2282–2302.
- [14] J. IMAI AND K. TAN, *Dimension reduction approach to simulating exotic options in a Meixner Lévy market*, IAENG Int. J. Appl. Math., 39 (2009), pp. 265–275.
- [15] R. KAWAI, *An importance sampling method based on the density transformation of Lévy processes*, Monte Carlo Methods Appl., 12 (2006), pp. 171–186.
- [16] R. LEPAGE, *Multidimensional infinitely divisible variables and processes. II*, in Probability in Banach Spaces, III, Lecture Notes in Math. 860, Springer-Verlag, Berlin, New York, Heidelberg, 1981, pp. 279–284.
- [17] D. MADAN AND M. YOR, *Representing the CGMY and Meixner Lévy processes as time changed Brownian motions*, J. Comput. Finance, 12 (2008), pp. 27–47.
- [18] H. NIEDERREITER, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.
- [19] A. OWEN, *Latin supercube sampling for very high-dimensional simulations*, ACM Trans. Model. Comput. Simul., 8 (1998), pp. 71–102.
- [20] A. OWEN, *Randomly permuted (t, m, s) -nets and (t, s) -sequences*, in Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, H. Niederreiter and P. J. S. Shiue, eds., Springer-Verlag, New York, 1995, pp. 299–317.
- [21] J. ROSIŃSKI, *Series representations of Lévy processes from the perspective of point processes*, in Lévy Processes—Theory and Applications, O.-E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick, eds., Birkhäuser, Boston, 2001, pp. 401–415.
- [22] J. ROSIŃSKI, *Tempering stable processes*, Stochastic Process. Appl., 117 (2007), pp. 677–707.
- [23] G. SAMORODNITSKY AND M. S. TAQQU, *Stable Non-Gaussian Random Processes*, Chapman & Hall, New York, 1994.
- [24] K. SATO, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, UK, 1999.
- [25] I. M. SOBOLOV, *Distribution of points in a cube and integration nets*, Uspehi Mat. Nauk., 21 (1966), pp. 271–272.
- [26] I. M. SOBOLOV, *Global sensitivity indices for nonlinear mathematical methods and their Monte Carlo estimates*, Math. Comput. Simul., 55 (2001), pp. 271–280.
- [27] X. WANG AND K. T. FANG, *The effective dimension and quasi-Monte Carlo integration*, J. Complexity, 19 (2003), pp. 101–124.
- [28] W. VERVAAT, *On a stochastic difference equation and representation of nonnegative infinitely divisible random variables*, Adv. in Appl. Probab., 11 (1979), pp. 750–783.