An Optimization Approach to Weak Approximation of Stochastic Differential Equations with Jumps

KENJI KASHIMA* AND REIICHIRO KAWAI†

Abstract

We propose an optimization approach to weak approximation of stochastic differential equations with jumps. A mathematical programming technique is employed to obtain numerically upper and lower bound estimates of the expectation of interest, where the optimization procedure ends up with a polynomial programming. A major advantage of our approach is that we do not need to simulate sample paths of jump processes, for which few practical simulation technique exist. We provide numerical results of moment estimations for Doléans-Dade stochastic exponential, truncated stable Lévy processes and Ornstein-Uhlenbeck-type processes to illustrate that our method is able to capture very well the distributional characteristics of stochastic differential equations with jumps.

MSC: 60H10, 65C30, 60G51, 90C22.

Keywords: Doléans-Dade stochastic exponential, Lévy processes, stochastic differential equations, truncated stable process, Ornstein-Uhlenbeck-type process, polynomial programming, weak approximation.

1 Introduction

Stochastic differential equations have long been used to build realistic models in economics, finance, biology, the social sciences, chemistry, physics and other fields. In most active fields of application, dynamics with possible sudden shift have become more and more important. To model such shifts, one would like to employ stochastic differential equations where the underlying randomness contains jumps. Regardless of its practical importance, the theory and the computational techniques of the stochastic differential equations with jumps have not been developed as thoroughly as in the diffusion case. As nice references on the subject, we refer to Applebaum [1], Bass [2] and Situ [21].

From a practical point of view, the sample paths approximation of stochastic differential equations has been a central issue for the purpose of numerical evaluation and simulation on the computer. There are two notions of the approximation; strong and weak approximations. On one hand, strong approximation schemes provide pathwise approximations $\mathbb{E}[\|X_T - X_T^\Delta\|^2] \leq C\Delta^\beta$, where $\Delta \in (0,1)$ is the maximum stepsize of a time discretization $\{X_t^\Delta : t \in [0,T]\}$, where C is a positive constant and β is the order of the approximation. The strong approximation is used for scenario analysis, filtering or hedge simulation. For applications such as derivative pricing, on the other hand, the computation of moments or expected utilities, we need to estimate the expected value of a function of marginals, or a functional of sample paths. In those cases, the so-called weak approximations are sufficient, that is, $|\mathbb{E}[V(X_T)] - \mathbb{E}[V(X_T^\Delta)]| \leq C\Delta^\beta$, where V is a suitable smooth function satisfying some polynomial growth condition at infinity. Other applications of the weak approximation include the computation of functional integrals, invariant measures, and Lyapunov exponents.

Published in *Applied Numerical Mathematics* (2011) **61**(5) 641-650.

*Email address: kashima@mei.titech.ac.jp. Postal address: Graduate School of Information Science and Engineering, Tokyo Institute of Technology, Tokyo, 152-8552, Japan.

†Corresponding author. Email address: reiichiro.kawai@gmail.com. Postal address: Department of Mathematics, University of Leicester, Leicester LE1 7RH, UK.

Theoretical properties of time discretization schemes are mostly studied for the diffusion case. See Kloeden and Platen [9] for detailed investigations. For stochastic differential equations with jumps, the weak rate of convergence of Euler-Maruyama schemes is studied, for example, in Liu and Li [12], Kubilius and Platen [10], Protter and Talay [17], and Jacod et al. [5]. Jump-adapted discretization is investigated, for example, in Bruti-Liberati and Platen [4], Higham and Kloeden [6] and Mikulevičius and Platen [13], while jump adaptation is only valid in the compound Poisson framework.

The main purpose of this paper is to propose a new approach to weak approximation of stochastic differential equations with jumps. Unlike Monte Carlo simulations with time discretization approximation of sample paths, we employ a mathematical programming technique to obtain numerically upper and lower bounds of the expectation of interest, where the optimization procedure ends up with a polynomial programming. To this end, we follow the idea of Primbs [16] that reduces computation of bounds to an optimization problem.[1] Note that the framework of [16] only deals with the pure diffusion setting, that is, without jump component, for which standard Monte Carlo methods are often sufficient. In contrast, few efficient simulation techniques exist for jump processes. Due to the complexity of the Ito formula for general jump processes, we need to carefully examine whether resulting optimization problems are practically solvable. Fortunately, as we will demonstrate in what follows, our approach covers various jump processes of practical interest.

The rest of this paper is organized as follows. Section 2 discusses our motivation by demonstrating difficulties and limitations of time discretization approximation, which our approach may get around. Section 3 introduces and studies our optimization approach to the weak approximation. Section 4 provides three numerical examples to illustrate that our method is able to capture very well the marginal distributions of stochastic differential equations with jumps via moment estimation. Finally, Section 5 concludes.

2 Motivation

Let us begin this section with general notations which will be used throughout the paper. Let $\mathbb{N}$ be the collection of natural numbers with $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. We also use the notations $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$ and $\mathbb{R}_+ := (0, +\infty)$. For $k \in \mathbb{N}$, ∂_k indicates the partial derivative with respect to k-th argument. We denote by C^{k_t,k_x} the class of continuous functions with continuous differentiability of k_t-time for the first argument and of k_x-time for the second argument. We denote by $\stackrel{\mathscr{L}}{\to}$ the weak convergence of random processes in the space $\mathbb{D}([0,+\infty);\mathbb{R})$ of càdlàg functions from $[0,+\infty)$ into $\mathbb{R}$ equipped with the Skorohod topology. We henceforth fix $(\Omega, \mathscr{F}, \mathbb{P})$ as our underlying probability space.

Let X_0 be given in $\mathbb{R}$ and let $T > 0$. Consider a one-dimensional stochastic differential equation

$$dX_t = a_0(t,X_t)\,dt + a_1(t,X_t)\,dW_t + \int_{\mathbb{R}_0} b(t,X_{t-},z)\,(\mu - \nu)(dz,dt), \quad t \in [0,T], \tag{2.1}$$

where $\{W_t : t \geq 0\}$ is a standard Brownian motion and where μ is a Poisson random measure on $\mathbb{R}_0$ whose compensator is given by the Lévy measure ν, that is, a σ-finite measure defined on $\mathbb{R}_0$ satisfying $\int_{\mathbb{R}_0}(|z|^2 \wedge 1)\nu(dz) < +\infty$. Here, we assume that for each $t \in [0,T]$, the functions $a_0(t,x)$, $a_1(t,x)$ and $b(t,x,z)$ in (2.1) satisfy the usual conditions such as at most linear growth and Lipschitz so that the solution of (2.1) is well defined. We henceforth equip our underlying probability space with the natural filtration $(\mathscr{F}_t)_{t\in[0,T]}$ generated by $\{X_t : t \in [0,T]\}$. Let $\mathscr{X}$ be the support of the stochastic process $\{X_t : t \in [0,T]\}$ defined by (2.1), that is,

$$\mathscr{X} := \inf\{B \in \mathscr{B}(\mathbb{R}) : \mathbb{P}(X_t \in B, t \in [0,T]) = 1\}.$$

Our interest throughout this study is in computing the expectation

$$\mathbb{E}[V(\tau, X_\tau)],$$

[1] See also Lasserre et al. [11] for the dual formulation called generalized moment problems.

for $V : [0,+\infty) \times \mathbb{R} \mapsto \mathbb{R}$ such that $\mathbb{E}[|V(\tau, X_\tau)|] < +\infty$, where τ is an $(\mathscr{F}_t)_{t\in[0,T]}$-stopping time taking its values in $[0,T]$. To demonstrate some difficulties and limitations of time discretization approximation of stochastic differential equations with jumps, let us begin with a simple setting with $X_0 > 0$, $\tau = T$, $a_0(t,x) = a_1(t,x) \equiv 0$, and $b(t,x,z) = xz$, that is, (2.1) reduces to a Doléans-Dade stochastic exponential

$$dX_t = X_{t-} \int_{\mathbb{R}_0} z(\mu - \nu)(dz, dt), \quad X_0 > 0, \tag{2.2}$$

which is a martingale with respect to its natural filtration. Assume further that the Lévy measure ν is defined only on $(-1, +\infty)$. It holds by the Ito formula that

$$d \ln X_t = - \int_{(-1,+\infty)} z \nu (dz)\, dt + \int_{(-1,+\infty)} \ln(1+z)\, \mu (dz, dt),$$

or equivalently, in the canonical form,

$$X_t = X_0 \exp\left[-t \int_{(-1,+\infty)} z \nu (dz) + \int_0^t \int_{(-1,+\infty)} \ln(1+z)\, \mu (dz, ds) \right]. \tag{2.3}$$

It follows from the expression (2.3) that $X_t > 0$, *a.s.* For the computation of $\mathbb{E}[V(T, X_T)]$, a standard technique is the Monte Carlo simulation with sample generations of the marginal X_T. Let us discuss typical difficulties in simulation of X_T.

(i) In most cases, the marginal X_T is simulated via the time discretization of the sample paths of $\{X_t : t \in [0,T]\}$. It is however rare to know how to simulate the increments $\int_{t_1}^{t_2} \int_{\mathbb{R}_+} z(\mu - \nu)(dz, dt)$, with few exceptions such as gamma processes and stable processes. Moreover, the use of Euler-Maruyama schemes may ruin intrinsic properties of $\{X_t : t \geq 0\}$, such as the non-negativity of sample paths.

(ii) The simulation knowledge of increments $\int_{t_1}^{t_2} \int_{\mathbb{R}_+} z(\mu - \nu)(dz, dt)$ is in general not equivalent to that of the increments with an arbitrary stepsize. In contrast, this problem never appear in the diffusion case due to the Gaussian scaling property.

(iii) We may alternatively simulate the sample paths based upon (2.3) using a shot noise representation of Poisson random measure $\mu(dz, ds)$. (See Rosiński [18] for details.) However, as in the case of increments, it is difficult to find a shot noise series representation in a convenient form. In addition, it is not sensible to generate random sequences for each sample path, in particular when the series is infinite.

Even the issues (i) and (iii) may not arise in the pure diffusion case. For example, a simple linear form $dX_t = X_t \sigma dW_t$ can be rewritten in the canonical form $X_t = X_0 \exp[-\sigma^2/2 + \sigma W_t]$, which is easy to simulate. We will give concrete illustrations for those issues in Section 4.1.

3 Optimization approach to estimation of upper and lower bounds

In this section, we formulate a mathematical programming problem, which yields upper and lower bounds of the expectation $\mathbb{E}[V(\tau, X_\tau)]$ of the model (2.2). Our framework is completely different from the aforementioned Monte Carlo simulations, in the sense that we generate no random variates.

We are now in a position to introduce our optimization approach to the weak approximation. It holds by the Ito formula that for $f \in C^{1,2}([0,T] \times \mathscr{X}; \mathbb{R})$,

$$df(t, X_t) = \mathscr{A} f(t, X_t) dt + \partial_2 f(t, X_t) a_1(t, X_t) dW_t + \int_{\mathbb{R}_0} B_z f(t, X_{t-})\, (\mu - \nu)(dz, dt), \quad a.s., \tag{3.1}$$

where

$$\mathscr{A}f(t,x) := \partial_1 f(t,x) + \partial_2 f(t,x)a_0(t,x) + \frac{1}{2}\partial_2^2 f(t,x)a_1(t,x)^2 + \int_{\mathbb{R}_0} \left(B_z f(t,x) - \partial_2 f(t,x) b(t,x,z)\right) \nu(dz),$$

and for $z \in \mathbb{R}_0$,

$$B_z f(t,x) := f\left(t, x + b\left(t,x,z\right)\right) - f\left(t,x\right).$$

If

$$\mathbb{E}\left[\int_0^\tau \left(\partial_2 f(t,X_t)a_1(t,X_t)\right)^2 dt\right] < +\infty,$$

and if

$$\mathbb{E}\left[\int_0^\tau \int_{\mathbb{R}_0} \left(B_z f(t,X_t)\right)^2 \nu(dz)dt\right] < +\infty, \tag{3.2}$$

then the stochastic process $\{f(t,X_t) - f(0,X_0) - \int_0^t \mathscr{A}f(s,X_s)ds : t \in [0,T]\}$ is a square-integrable martingale with respect to the filtration $(\mathscr{F}_t)_{t\in[0,T]}$. We can then derive the associated Dynkin formula, for any $(\mathscr{F}_t)_{t\in[0,T]}$-stopping time τ taking values in $[0,T]$,

$$\mathbb{E}\left[f(\tau,X_\tau)\right] - f(0,X_0) = \mathbb{E}\left[\int_0^\tau \mathscr{A}f(s,X_s)ds\right].$$

Remark 3.1. The formula (3.1) is not in a standard form given in the literature, while this can be obtained in a straightforward manner. (See, for example, Applebaum [1].) Note also that we have not imposed any conditions on the jump component for the validity of (3.1). In fact, we will shortly need to impose conditions for optimization problems of our interest to be well defined, which are, much stronger than, sufficient to guarantee that (3.1) is well defined. □

Throughout this paper, we freeze the stopping time $\tau = T$. As soon as one finds an $f \in C^{1,2}([0,T] \times \mathscr{X};\mathbb{R})$ such that $\mathscr{A}f(t,x) \leq 0$ on $\mathscr{X}$ and $f(T,x) \geq V(T,x)$ on $\mathscr{X}$, we get

$$\mathbb{E}\left[V(T,X_T)\right] \leq \mathbb{E}\left[f(T,X_T)\right] \leq f(0,X_0).$$

Clearly, the deterministic value $f(0,X_0)$ serves as an upper bound of $\mathbb{E}[V(T,X_T)]$. To minimize this upper bound $f(0,X_0)$, we now turn to the optimization problem

$$\left|\begin{array}{ll} \min & f(0,X_0) \\ \text{s.t.} & f(T,x) \geq V(T,x) \quad \text{on } \mathscr{X}, \\ & \mathscr{A}f(t,x) \leq 0 \quad \text{on } [0,T]\times\mathscr{X}, \\ & f \in C^{1,2}([0,T]\times\mathscr{X};\mathbb{R}). \end{array}\right.$$

This optimization problem is very difficult to solve since the function class of f and V are too broad. To simplify the above optimization problem, we restrict the function f to be polynomial in both t and x, that is, in the form

$$f(t,x) = \sum_{B(0,0)} c_{k_t,k_x} t^{k_t} x^{k_x}, \tag{3.3}$$

where

$$B(l,m) := \left\{(k_t,k_x) \in \mathbb{N}^2 : l \leq k_t \leq K_t, m \leq k_x \leq K_x\right\},$$

for some fixed natural numbers K_t and K_x and for a sequence $\{c_{k_t,k_x}\}_{B(0,0)}$ of constants. For convenience in notation, we henceforth denote by C_p the class of polynomial functions of the form (3.3). (It is certainly more precise to denote the class by $C_p^{K_t,K_x}$ instead, while we suppress K_t and K_x in what follows. There is

no possibility of confusion since they are fixed throughout implementation of each optimization.) We also need to set V to be a *piecewise* polynomial in both t and x, and both a_0 and a_1 are in $C_p([0,T]\times\mathscr{X};\mathbb{R})$. We have now arrived at the following optimization problem

$$\left|\begin{array}{ll}\min & f(0,X_0)\\ \text{s.t.} & f(T,x)\geq V(T,x)\quad \text{on } \mathscr{X},\\ & \mathscr{A}f(t,x)\leq 0\quad \text{on } [0,T]\times\mathscr{X},\\ & f\in C_p([0,T]\times\mathscr{X};\mathbb{R}).\end{array}\right. \tag{3.4}$$

Suppose, for a moment, that there is no jump in the formulation (3.1), that is, $b\equiv 0$ as in the setting of [16]. Then, $\mathscr{A}f$ is trivially polynomial, and consequently the entire setting (3.4) reduces to a polynomial optimization problem. Let us return to our setting with jumps. It turns out that $\mathscr{A}f$ is not necessarily polynomial due to the jump component. We can circumvent this difficulty by decomposing the coefficient b as

$$b(t,x,z)=b_1(t,x)b_2(z), \tag{3.5}$$

where $b_1\in C_p([0,T]\times\mathscr{X};\mathbb{R})$ and where $b_2:\mathbb{R}_0\mapsto\mathbb{R}$ such that

$$\int_{\mathbb{R}_0}|b_2(z)|^k\,\nu(dz)<+\infty,\quad k=2,\ldots,k_x. \tag{3.6}$$

and then, a simple algebra yields

$$\begin{aligned}\mathscr{A}f(t,x)=&\sum_{B(1,0)}c_{k_t,k_x}k_t t^{k_t-1}x^{k_x}+\sum_{B(0,1)}c_{k_t,k_x}t^{k_t}k_x x^{k_x-1}a_0(t,x)\\ &+\frac{1}{2}\sum_{B(0,2)}c_{k_t,k_x}t^{k_t}k_x(k_x-1)x^{k_x-2}a_1(t,x)^2\\ &+\sum_{B(0,2)}c_{k_t,k_x}t^{k_t}\sum_{k=0}^{k_x-2}{}_{k_x}C_k x^k b_1(t,x)^{k_x-k}\int_{\mathbb{R}_0}b_2(z)^{k_x-k}\nu(dz).\end{aligned}$$

Since the integral $\int_{\mathbb{R}_0}b_2(z)^{k_x-k}\nu(dz)$ is independent of t and x, the optimization problem (3.4) falls in the framework of polynomial programming.

In general, polynomial optimization problems are still NP hard. However, if the degrees K_t and K_x are fixed, *sum of squares relaxation* techniques enable us to solve the problem efficiently. Let us here briefly describe this relaxation procedure in a setting with $\mathscr{X}=\mathbb{R}_+$, the value function V being polynomial and $c_{k_t,k_x}=0$ for $k_t+k_x>2n$, where n is a positive integer. Let $q(t,x)$ be a vector consisting of $(n+1)(n+2)/2$ monomials of the form $t^k x^l$ such that $k\geq 0$, $l\geq 0$ and $k+l\leq n$. Instead of the original problem (3.4), we will solve

$$\left|\begin{array}{ll}\min & f(0,X_0)\\ \text{s.t.} & f(T,x)=V(T,x)+q(t,x)^{\mathsf{T}}Q_1q(t,x)+xq(t,x)^{\mathsf{T}}Q_2q(t,x),\\ & \mathscr{A}f(t,x)=-q(t,x)^{\mathsf{T}}Q_3q(t,x)-xt(T-t)q(t,x)^{\mathsf{T}}Q_4q(t,x),\\ & \{Q_k\}_{k=1,\ldots,4}\text{ are positive semidefinite matrices,}\end{array}\right. \tag{3.7}$$

where decision variables are not only f but also Q_k's. This formulation serves as a relaxation of the original problem (3.4) since the equalities in (3.7) imply the inequalities in (3.4), due to the positive semidefiniteness of Q_k's. The optimization problem (3.7) can be rewritten as an optimization problem under a set of linear equalities and semidefiniteness constraints. This problem conversion is automatically executed by SOSTOOLS. The resulting problem can be solved with semidefinite programming, for which several well-established solvers exist.

Remark 3.2. The decomposition (3.5) and the integrability condition (3.6) are sufficient to guarantee that the Ito formula (3.1) is well defined. If $b_2(z) = \theta z$ for some constant θ, then (2.1) is called a *Lévy-driven* stochastic differential equation, since then the jump component reduces to

$$\int_{\mathbb{R}_0} b(t, X_{t-}, z)(\mu - \nu)(dz, dt) = \theta b_1(t, X_{t-}) \int_{\mathbb{R}_0} z(\mu - \nu)(dz, dt),$$

where the integral of the right hand side corresponds to a Lévy process. □

To obtain a lower bound of $\mathbb{E}[V(T, X_T)]$, we wish to find a $g \in C_p([0,T] \times \mathscr{X}; \mathbb{R})$ via the polynomial programming

$$\left| \begin{array}{ll} \max & g(0, X_0) \\ \text{s.t.} & g(T,x) \le V(T,x) \quad \text{on } \mathscr{X}, \\ & \mathscr{A} g(t,x) \ge 0 \quad \text{on } [0,T] \times \mathscr{X}, \\ & g \in C_p([0,T] \times \mathscr{X}; \mathbb{R}), \end{array} \right. \tag{3.8}$$

in a similar manner to (3.4). Let us remind again that our optimization approach yields upper and lower bounds of $\mathbb{E}[V(T, X_T)]$ without sample paths simulation.

Remark 3.3. The conservativeness of sum of square relaxation has been intensively studied, for example, in Parrilo [14]. It is known that in most settings, the relaxation error can be made arbitrarily small by employing more sophisticated relaxation techniques than those applied in (3.7), in return for rapidly increasing computation burden. Let us also discuss the conservativeness caused by restricting bounding functions f and g to polynomials $C_p([0,T] \times \mathscr{X}; \mathbb{R})$. A similar issue has been discussed in the framework of generalized moment problems, which is the dual formulation of ours, for example, in Lasserre et al. [11]. They show that the gap between optimal upper and lower bounds converges to 0 as the degrees of moments tend to infinity, under suitable conditions on underlying random elements. In fact, thanks to the duality of formulations, this issue can be addressed in a similar manner. Error analysis of this type is, however, not necessarily informative in the framework of semidefinite programming, since it is impossible in practice to solve semidefinite programming problems of arbitrarily high dimension. For this reason, it seems more worthwhile to examine effectiveness of our approach through numerical experiments, which we will demonstrate in Section 4. □

4 Numerical illustrations

In this section, we test our method on a Doléans-Dade stochastic exponential, a truncated stable subordinator and a process of Ornstein-Uhlenbeck type, all with no diffusion component, that is, $a_1(t,x) \equiv 0$. We here estimate moments of those processes, which are available in closed form for the sake of comparison. In the numerical examples presented hereafter, we utilized MATLAB SOSTOOLS combined with SeDuMi [15, 22], using a computer with a Pentium 4 3.2GHz processor and 2 GB memory.

4.1 Doléans-Dade stochastic exponential driven by gamma process

Set $X_0 > 0$, $a_0(t,x) = a_1(t,x) = 0$, $b_1(t,x) = x$, $b_2(z) = z$, and

$$\nu(dz) = a \frac{e^{-bz}}{z} dz, \quad z > 0,$$

for $a > 0$ and $b > 0$, that is a gamma Lévy measure. In this setting, (2.1) reduces to a Doléans-Dade stochastic exponential (2.2). It is clear that $\mathbb{E}[X_T] = X_0$. Moreover, we have $\mathbb{E}[X_t^2] = X_0^2 e^{\frac{a}{b^2} t}$, since by the Ito-Wiener isometry,

$$\mathbb{E}\left[X_T^2\right] = X_0^2 + \int_{\mathbb{R}_+} z^2 \nu(dz) \mathbb{E}\left[\int_0^T X_t^2 dt\right] = X_0^2 + \frac{a}{b^2} \int_0^T \mathbb{E}\left[X_t^2\right] dt,$$

where the interchange of the integrals holds by the Fubini theorem with the almost sure non-negativity of X_t^2. (For more details, we refer to Applebaum [1].)

On one hand, Euler-Maruyama schemes do not guarantee the non-negativity of sample paths. To illustrate this, let $N \in \mathbb{N}$ and $\Delta := T/N$, and consider the equidistant time discretization approximation of $\{X_t : t \in [0,T]\}$, that is,

$$\frac{X_{k\Delta}}{X_{(k-1)\Delta}} = 1 + \gamma_k(a\Delta, b) - a\Delta/b, \tag{4.1}$$

where $\{\gamma_k(a,b)\}_{k\in\mathbb{N}}$ is a sequence of iid gamma random variables with the common distribution $b^a/\Gamma(a)y^{a-1}e^{-by}dy$ on $\mathbb{R}_+$. With a choice of (a,b,Δ) satisfying $1 - a\Delta/b < 0$, discretized sample paths may drop below zero. Such numerical experiments produce a misleading result $\mathbb{E}[X_T] \ll X_0$.

On the other hand, based upon the canonical form (2.3) with a shot noise series representation due to Bondesson [3], sample paths can be simulated as

$$X_t = X_0 \exp\left[-t\frac{a}{b} + \sum_{k=1}^{+\infty} \ln\left(1 + e^{-\frac{\Gamma_k}{aT}}\frac{V_k}{b}\right)\mathbb{1}(T_k \le t)\right], \quad t \in [0,T], \tag{4.2}$$

where $\{\Gamma_k\}_{k\in\mathbb{N}}$ are arrival times of a standard Poisson process, where $\{V_k\}_{k\in\mathbb{N}}$ is a sequence of iid standard exponential random variables, and where $\{T_k\}_{k\in\mathbb{N}}$ is a sequence of iid uniform random variables on $[0,T]$. It is, however, not sensible to generate this infinite series for each sample path.

By noting that $\mathscr{X} = \mathbb{R}_+$ and $\int_{\mathbb{R}_+} z^k \nu(dz) = a(k-1)!/b^k$ for $k = 2,3,\ldots$, it holds that for $f \in C_p([0,T]\times\mathbb{R}_+;\mathbb{R})$,

$$\mathscr{A}f(t,x) = \sum_{B(1,0)} c_{k_t,k_x} k_t t^{k_t-1} x^{k_x} + \sum_{B(0,2)} c_{k_t,k_x} t^{k_t} x^{k_x} \sum_{k=0}^{k_x-2} {}_{k_x}C_k \frac{a(k_x-k-1)!}{b^{k_x-k}}.$$

We can check the condition (3.2) by

$$\mathbb{E}\left[\int_0^T \int_{\mathbb{R}_+} B_z f(t,X_t)\nu(dz)dt\right] = \mathbb{E}\left[\int_0^T \int_{\mathbb{R}_+} \left(\sum_{B(0,1)} c_{k_t,k_x} t^{k_t} X_t^{k_x} \sum_{k=1}^{k_x} {}_{k_x}C_k z^k\right)^2 a\frac{e^{-bz}}{z}dzdt\right].$$

By the Fubini theorem, it thus suffices to have $\mathbb{E}[X_t^{2K_x}] < +\infty$ for each $t \in [0,T]$. In view of (2.3), it holds that for each $k_x \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}\left[X_t^{2K_x}\right] &\le X_0^{2K_x}\mathbb{E}\left[\exp\left(2K_x\int_0^T\int_{\mathbb{R}_+}\ln(1+z)\mu(dz,ds)\right)\right] \\ &= X_0^{2K_x}\exp\left[T\int_{\mathbb{R}_+}\left(e^{2K_x z}-1\right)a\frac{e^{-b(e^z-1)}}{e^z-1}dz\right] < +\infty.\end{aligned}$$

This implies that K_x can be taken arbitrarily large.

Here, we test our optimization approach on estimation of $\mathbb{E}[X_t] = X_0$ and $\mathbb{E}[X_t^2] = X_0^2 e^{\frac{a}{b^2}t}$. Numerical results are presented in Table 1. The two numbers given in the row "SDP" indicate upper and lower bounds obtained through our optimization approach. Note that when $\mathscr{X}$ is unbounded, we must choose $K_x \ge p$ for the estimation of the p-th moment because of the constraint $f(T,x) \ge x^p$ for $x \in \mathscr{X}$. In view of this, we choose the minimal degree $K_x = p$. We also set $K_t = p$. It took at most 1 second to solve each optimization problem to obtain a bound. We can see from Table 1 that even such a low degree polynomial function achieves tight upper and lower bounds. For comparison purpose, we also provide 99%-confidence intervals of Monte Carlo simulations based upon 50000 iid samples. The "MC (ES)" rows present intervals obtained through the Euler-Maruyama Scheme (4.1) with equidistant stepsize of 1e-2, while the "MC (SR)" rows give results through the Series Representation (4.2) with truncation of the infinite series to 100 per unit

time. Recall that Euler-Maruyama schemes produce some amount of discretization error, while an infinite shot noise series representation provides an approximative simulation method as well due to truncation of the infinite series. Any large sample size in Monte Carlo simulations can never be in competition with our results since the upper and lower bounds obtained through our method form nothing but the 100%-confidence interval. By taking into account the computing time required for the Monte Carlo simulations, the superiority of our optimization approach is evident.

		$t=1$	$t=2$	$t=3$
$\mathbb{E}[X_t]$	True Value	1.0000	1.000	1.0000
	SDP	1.0000 – 1.0000	1.0000 – 1.0000	1.0000 – 1.0000
	MC (ES)	[0.9973, 1.0021]	[0.9951, 1.0021]	[0.9934, 1.0022]
	MC (SR)	[0.9986, 1.0036]	[0.9959, 1.0028]	[0.9942, 1.0029]
$\mathbb{E}[X_t^2]$	True Value	1.04545	1.09296	1.14263
	SDP	1.04541 – 1.04550	1.09294 – 1.09300	1.14249 – 1.14288
	MC (ES)	[1.0347, 1.0531]	[1.0727, 1.1099]	[1.1073, 1.1703]
	MC (SR)	[1.0388, 1.0565]	[1.0744, 1.1029]	[1.1178, 1.1593]

Table 1: Numerical results for Doléans-Dade stochastic exponential with $X_0 = 1$ and $(a,b) = (0.1, 1.5)$.

4.2 Short- and long-range behaviors of truncated stable subordinator

Set $X_0 = 0$, $a_1(t,x) = 0$, $b_1(t,x) = 1$, $b_2(z) = z$, and

$$\nu(dz) = \frac{1}{z^{1+\alpha}} \mathbb{1}(z \le \eta)\, dz, \quad z \in \mathbb{R}_+,$$

for some $\alpha \in (0,1)$ and $\eta > 0$, and $a_0(t,x) = \int_{\mathbb{R}_+} z\nu(dz)$. Then, the stochastic differential equation (2.1) reduces to

$$X_t = \int_0^t \int_{\mathbb{R}_0} z\mu(dz, ds),$$

that is, a truncated stable subordinator.

Here, in view of Houdré and Kawai [7] and Rosiński [19], we test our method on short- and long-range behaviors of truncated stable processes, by looking into convergence in terms of moments. Due to the truncation of Lévy measure, its marginal has a finite moment of every polynomial order. (See Theorem 25.3 of Sato [20], for example.) For every $t > 0$, its moments of up to fourth order are given by

$$\begin{aligned}
\mathbb{E}[X_t] &= \int_{\mathbb{R}_+} z\nu(dz) = t\frac{\eta^{1-\alpha}}{1-\alpha}, \\
\mathbb{E}\left[(X_t - \mathbb{E}[X_t])^2\right] &= t\int_{\mathbb{R}_+} z^2\nu(dz) = t\frac{\eta^{2-\alpha}}{2-\alpha}, \\
\mathbb{E}\left[(X_t - \mathbb{E}[X_t])^3\right] &= t\int_{\mathbb{R}_+} z^3\nu(dz) = t\frac{\eta^{3-\alpha}}{3-\alpha}, \\
\mathbb{E}\left[(X_t - \mathbb{E}[X_t])^4\right] &= t\int_{\mathbb{R}_+} z^4\nu(dz) + 3\left(t\int_{\mathbb{R}_+} z^2\nu(dz)\right)^2 = t\frac{\eta^{4-\alpha}}{4-\alpha} + 3\left(t\frac{\eta^{2-\alpha}}{2-\alpha}\right)^2.
\end{aligned}$$

First, for the short-range behavior, we can prove that as $h \downarrow 0$,

$$\left\{h^{-1/\alpha} X_{ht} : t \ge 0\right\} \xrightarrow{\mathscr{L}} \left\{X_t^{(\alpha)} : t \ge 0\right\},$$

where $\{X_t^{(\alpha)} : t \geq 0\}$ is a stable subordinator with Lévy measure $z^{-1-\alpha}dz$ on $\mathbb{R}_+$. To apply our optimization method to the first and the second moments, it is sufficient to first compute moments of X_{ht}, and to multiply next a scaling related to h. Clearly, $\mathscr{X} = \mathbb{R}_+$. We have for $f \in C_p([0,T] \times \mathbb{R}; \mathbb{R})$,

$$\mathscr{A} f(ht,x) = \sum_{B(1,0)} c_{k_t,k_x} k_t (ht)^{k_t-1} x^{k_x} + \sum_{B(0,1)} c_{k_t,k_x} (ht)^{k_t} \sum_{k=0}^{k_x-1} {}_{k_x}C_k x^k \frac{\eta^{k_x-k-\alpha}}{k_x-k-\alpha}.$$

On the other hand, in the long run, as $h \uparrow +\infty$,

$$\left\{ h^{-1/2} \left(X_{ht} - \mathbb{E}\left[X_{ht}\right] \right) : t \geq 0 \right\} \overset{\mathscr{L}}{\to} \left\{ W_t : t \geq 0 \right\},$$

where $\{W_t : t \geq 0\}$ is a (centered) Brownian motion with $\mathbb{E}[|W_1|^2] = \eta^{2-\alpha}/(2-\alpha)$. In this case, $\mathscr{X} = \mathbb{R}$, $a_0(t,x) = 0$, and we take the form $f(ht, X_{ht} - \mathbb{E}[X_{ht}])$. Hence, we have

$$\mathscr{A} f(ht,x) = \sum_{B(1,0)} c_{k_t,k_x} k_t (ht)^{k_t-1} x^{k_x} + \sum_{B(0,2)} c_{k_t,k_x} (ht)^{k_t} \sum_{k=0}^{k_x-2} {}_{k_x}C_k x^k \frac{\eta^{k_x-k-\alpha}}{k_x-k-\alpha}.$$

We present numerical results in Table 2 and in Table 3. Similarly to the previous example, for the estimation of the p-th moment, we set $K_t = K_x = p$.

Unlike in the last example, we do not provide numerical results of Monte Carlo simulations due to extremely large computing effort required. To the best of our knowledge, the only decent method for the sample paths generation is based on its infinite shot noise series representation

$$\{X_t : t \in [0,T]\} \overset{\mathscr{L}}{=} \left\{ \sum_{k=1}^{+\infty} \left(\eta^{-\alpha} + \frac{\alpha \Gamma_k}{T} \right)^{-1/\alpha} \mathbb{1}(T_k \leq t) : t \in [0,T] \right\}, \tag{4.3}$$

where $\{\Gamma_k\}_{k\in\mathbb{N}}$ are arrival times of a standard Poisson process and where $\{T_k\}_{k\in\mathbb{N}}$ is a sequence of iid uniform random variables on $[0,T]$. We have observed through our experiments that the infinite series converges at an extremely slow rate. For instance, even in the shortest time case $h = 0.01$, at least 5000 jumps are required to obtain decent estimation results. This Monte Carlo simulation is too expensive to be competitive against our optimization approach.

		$h=1$	$h=0.1$	$h=0.01$	$h \downarrow 0$
$\mathbb{E}\left[h^{-1/\alpha} X_{ht}\right]$	True Value	5.000	8.8914	15.811	$+\infty$
	SDP	5.000 – 5.000	8.8914 – 8.8914	15.811 – 15.811	n/a
$\mathbb{E}\left[(h^{-1/\alpha} X_{ht})^2\right]$	True Value	25.833	105.41	1083.3	$+\infty$
	SDP	25.833 – 25.833	105.41 – 105.41	1083.3 – 1083.3	n/a

Table 2: Numerical results for short-range behavior with $X_0 = 0$, $\alpha = 0.8$, $\eta = 1$, and $t = 1$.

		$h=1$	$h=5$	$h=10$	$h \uparrow +\infty$
$\mathbb{E}\left[\left(\frac{1}{\sqrt{h}}\left(X_{ht} - \mathbb{E}\left[X_{ht}\right]\right)\right)^2\right]$	True Value	0.8333	0.8333	0.8333	0.8333
	SDP	0.8333 – 0.8333	0.8333 – 0.8333	0.8333 – 0.8333	n/a
$\mathbb{E}\left[\left(\frac{1}{\sqrt{h}}\left(X_{ht} - \mathbb{E}\left[X_{ht}\right]\right)\right)^3\right]$	True Value	0.4545	0.2033	0.1437	0
	SDP	0.4545 – 0.4545	0.2032 – 0.2033	0.1437 – 0.1437	n/a
$\mathbb{E}\left[\left(\frac{1}{\sqrt{h}}\left(X_{ht} - \mathbb{E}\left[X_{ht}\right]\right)\right)^4\right]$	True Value	2.3958	2.1458	2.1146	2.0833
	SDP	2.3958 - 2.3958	2.1458 - 2.1458	2.1127 - 2.1150	n/a

Table 3: Numerical results for long-range behavior with $X_0 = 0$, $\alpha = 0.8$, $\eta = 1$, and $t = 1$.

4.3 Process of Ornstein-Uhlenbeck type with gamma stationary distribution

Let ν be a Lévy measure on $\mathbb{R}_+$ such that $\int_{\mathbb{R}_+} z\nu(dz) < +\infty$. Set $a_0(t,x) = -\lambda x + \int_{\mathbb{R}_+} z\nu(dz)$ for some $\lambda > 0$, $a_1(t,x) = 0$, $b_1(t,x) = 1$, $b_2(z) = z$, and X_0 is independent of μ. Then, the stochastic differential equation (2.1) reduces to

$$dX_t = -\lambda X_t dt + \int_{\mathbb{R}_+} z\mu(dz,dt),$$

which is called a process of Ornstein-Uhlenbeck type. Its strong solution is given in closed form

$$X_t = e^{-\lambda t}X_0 + \int_0^t \int_{\mathbb{R}_+} e^{-\lambda(t-s)} z\mu(dz,ds). \tag{4.4}$$

For simplicity, we fix $X_0 = 0$, $\lambda = 1$ and $\nu(dz) = bae^{-bz}dz$, where $a > 0$ and $b > 0$. We can show that the stationary distribution of $\{X_t : t \geq 0\}$ is gamma with probability density $b^a/\Gamma(a)x^{a-1}e^{-bx}$ defined on $\mathbb{R}_+$. (See, for example, Sato [20] for more details.)

By noting that $\mathscr{X} = \mathbb{R}_+$ and $\int_{\mathbb{R}_+} z^k \nu(dz) = ak!/b^k$ for $k \in \mathbb{N}$, it holds that for $f \in C_p([0,T]\times\mathbb{R}_+;\mathbb{R})$,

$$\begin{aligned}\mathscr{A}f(t,x) = &\sum_{B(1,0)} c_{k_t,k_x} k_t t^{k_t-1} x^{k_x} + \left(-x + \frac{a}{b}\right) \sum_{B(0,1)} c_{k_t,k_x} t^{k_t} k_x x^{k_x-1} \\ &+ \sum_{B(0,2)} c_{k_t,k_x} t^{k_t} \sum_{k=0}^{k_x-2} {}_{k_x}C_k x^k \frac{a(k_x-k)!}{b^{k_x-k}}.\end{aligned}$$

The condition (3.2) holds for each K_t and K_x, since for each $t \in [0,T]$,

$$\int_0^t \int_{\mathbb{R}_+} e^{-\lambda(t-s)} z\mu(dz,ds) \leq \int_0^T \int_{\mathbb{R}_+} z\mu(dz,ds), \quad a.s.,$$

where the right hand side is an infinitely divisible random variable, whose Lévy measure has an exponential decay at infinity.

We test our method on the distribution transition via moment estimations of $\mathbb{E}[X_t] = (1-e^{-t})a/b$, $\mathbb{E}[X_t^2] = (1-e^{-2t})a/b^2 + (1-e^{-t})^2a^2/b^2$, and $\lim_{t\uparrow+\infty}\mathbb{E}[X_t^3] = \Gamma(a+3)/(b^3\Gamma(a))$. Numerical results are presented in Table 4. We set K_t in the same manner as in the previous example. It is well known that computing effort required for solving a polynomial optimization problem through sum of squares decomposition significantly increases as the polynomial has larger degrees. In our experiments, however, even with a relatively large setting $K_x = 10$, computing time was at most 2 seconds.

Since the Lévy measure ν here is finite, sample paths can be simulated in the exact sense through a standard compound Poisson generation applied in the expression (4.4). To compare our methods with such a typical Monte Carlo technique, we also provide 99%-confidence interval based on 1000000 iid samples. As can be observed, even with the extraordinarily large number of samples, the 99%-confidence intervals are far from comparable with our results.

5 Concluding remarks

In this paper, we have developed a new approach to the weak approximation of stochastic differential equations with jumps, through an optimization problem yielding upper and lower bounds of the expectation of interest. It is a major advantage that our approach provides bounds without sample path simulation of jump processes, for which few practical simulation technique exist. In most numerical experiments which we have conducted, we have obtained fairly tight bounds without taking very large degrees of bounding polynomial functions. We conjecture that the tightness of upper and lower bounds may be ruined when the value

		$t=1$	$t=2$	$t=3$	$t \uparrow +\infty$
$\mathbb{E}[X_t]$	True Value	0.042141	0.057644	0.063348	0.06667
	SDP	0.042141 – 0.042141	0.057644 – 0.057644	0.063347 – 0.063348	n/a
	MC	[0.041733, 0.042747]	[0.056985, 0.058058]	[0.062879, 0.063967]	n/a
$\mathbb{E}\left[X_t^2\right]$	True Value	0.040205	0.046953	0.048347	0.04889
	SDP	0.040205 – 0.040205	0.046952 – 0.046955	0.048331 – 0.048347	n/a
	MC	[0.039618, 0.041484]	[0.045721, 0.047600]	[0.047666, 0.049604]	n/a
$\mathbb{E}\left[X_t^3\right]$	True Value	n/a	n/a	n/a	0.06844
	SDP	0.061217 – 0.061268	0.066812 – 0.066886	0.068009 – 0.068051	n/a
	MC	[0.059161, 0.064966]	[0.063362, 0.068839]	[0.066005, 0.071926]	n/a

Table 4: Numerical results with $X_0 = 0$ and $(a,b) = (0.1, 1.5)$. The intervals are 99%-confidence interval with 1000000 iid samples.

function has a highly complicated form, while this could be overcome by relaxing the bounding polynomial functions to piecewise polynomial functions.

In this paper, we have only dealt with stochastic differential equations whose marginal has finite moments of up to a sufficiently large polynomial order. There are, however, various interesting examples that do not meet such moment conditions, such as stochastic differential equations driven by stable Lévy processes and the Heston stochastic volatility model even in the pure diffusion case. In particular, the latter example entails a multidimensional formulation, which will be a numerically challenging problem due to the high dimensionality of semidefinite programming. Those issues will be addressed in our subsequent paper [8].

Besides extensions and improvements of our present approach, it would certainly be worthwhile to apply our method to practical problems, such as option pricing and calibration in finance, the probability tail estimation, and so on.

Acknowledgements

The authors would like to thank Japan Society for the Promotion of Science for Grant-in-Aid for Scientific Research 20740059 and 21686039. Part of this work was carried out while RK was based at Center for the Study of Finance and Insurance, Osaka University, Japan.

References

[1] D. Applebaum, Lévy Processes and Stochastic Calculus, Cambridge University Press, 2004.

[2] R. Bass, Stochastic differential equations with jumps, Probability Surveys, 1 (2004) 1-19.

[3] L. Bondesson, On simulation from infinitely divisible distributions, Advances in Applied Probability, 14(4) (1982) 855-869.

[4] N. Bruti-Liberati, E. Platen, Strong approximations of stochastic differential equations with jumps, Journal of Computational and Applied Mathematics, 205(2) (2007) 982-1001.

[5] J. Jacod, T.G. Kurtz, S. Méléard, P. Protter, The approximate Euler method for Lévy driven stochastic differential equations, Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques, 41 (2005) 523-558.

[6] D.J. Higham, P.E. Kloeden, Strong convergence rates for backward Euler on a class of nonlinear jump-diffusion problems, Journal of Computational and Applied Mathematics, 205(2) (2007) 949-956.

[7] C. Houdré, R. Kawai, On layered stable processes, Bernoulli, 13(1) (2007) 252-278.

[8] K. Kashima, R. Kawai, A weak approximation of stochastic differential equations with jumps through tempered polynomial optimization, Stochastic Models, 27(1) (2010) 26-49.

[9] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Third Printing, Springer, Berlin, 1999.

[10] K. Kubilius, E. Platen, Rate of weak convergence of the Euler approximation for diffusion processes with jumps, Monte Carlo Methods and Applications, 8(1) (2002) 83-96.

[11] J.B. Lasserre, T. Prieto-Rumeau, M. Zervos, Pricing a class of exotic via moments and SDP relaxations, Mathematical Finance, 16(3) (2006) 469-494.

[12] X.Q. Liu, C.W. Li, Weak approximations and extrapolations of stochastic differential equations with jumps, SIAM Journal on Numerical Analysis, 37(6) (2004) 1747-1767.

[13] R. Mikulevičius, E. Platen, Time discrete Taylor approximations for Itô processes with jump component, Math Nachr., 138 (1998) 93-104.

[14] P.A. Parrilo, Semidefinite programming relaxations for semialgebraic problems, Mathematical Programming Series B, 96(2) (2003) 293-320.

[15] S. Prajna, A. Papachristodoulou, P. Seiler, P.A. Parrilo, SOS-TOOLS: Sum of squares optimization toolbox for MATLAB, 2004.

[16] J.A. Primbs, Optimization based option pricing bounds via piecewise polynomial super- and sub-martingales, In: 2008 American Control Conference.

[17] P. Protter, D. Talay, The Euler scheme for Lévy driven stochastic differential equations, Annals of Probability, 25 (1997) 393-423.

[18] J. Rosiński, Series representations of Lévy processes from the perspective of point processes, In: Lévy Processes - Theory and Applications, Eds. Barndorff-Nielsen, O.-E., Mikosch, T., Resnick, S.I., Birkhäuser, 401-415, 2001.

[19] J. Rosiński, Tempering stable processes, Stochastic Processes and their Applications, 117(6) (2001) 677-707.

[20] K. Sato, Lévy processes and infinitely divisible distributions, Cambridge University Press, 1999.

[21] R. Situ., Reflecting Stochastic Differential Equations with Jumps and Applications, Chapman & Hall, 1999.

[22] J. Sturm, SeDuMi version 1.1. (2006)