Markov chain approximations for transition densities of Lévy processes*

Aleksandar Mijatović†   Matija Vidmar‡   Saul Jacka§

Abstract

We consider the convergence of a continuous-time Markov chain approximation X^h, $h > 0$, to an $\mathbb{R}^d$-valued Lévy process X. The state space of X^h is an equidistant lattice and its Q-matrix is chosen to approximate the generator of X. In dimension one ($d = 1$), and then under a general sufficient condition for the existence of transition densities of X, we establish sharp convergence rates of the normalised probability mass function of X^h to the probability density function of X. In higher dimensions ($d > 1$), rates of convergence are obtained under a technical condition, which is satisfied when the diffusion matrix is non-degenerate.

1 Introduction

Discretization schemes for stochastic processes are relevant both theoretically, as they shed light on the nature of the underlying stochasticity, and practically, since they lend themselves well to numerical methods. Lévy processes, in particular, constitute a rich and fundamental class with applications in diverse areas such as mathematical finance, risk management, insurance, queuing, storage and population genetics etc. (see e.g. [22]).

1.1 Short statement of problem and results

In the present paper, we study the rate of convergence of a weak approximation of an $\mathbb{R}^d$-valued ($d \in \mathbb{N}$) Lévy process X by a continuous-time Markov chain (CTMC). Our main aim is to understand the rates of convergence of transition densities. These cannot

---

†Imperial College London, UK. E-mail: a.mijatovic@imperial.ac.uk
‡University of Warwick, UK. E-mail: m.vidmar@warwick.ac.uk
§University of Warwick, UK. E-mail: s.d.jacka@warwick.ac.uk

be viewed as expectations of (sufficiently well-behaved, e.g. bounded continuous) real-valued functions against the marginals of the processes, and hence are in general hard to study.

Since the results are easier to describe in dimension one ($d = 1$), we focus first on this setting. Specifically, our main result in this case, Theorem 2.4, establishes the precise convergence rate of the normalised probability mass function of the approximating Markov chain to the transition density of the Lévy process for the two proposed discretisation schemes, one in the case where X has a non-trivial diffusion component and one when it does not. More precisely, in both cases we approximate X by a CTMC X^h with state space $\mathbb{Z}_h := h\mathbb{Z}$ and Q-matrix defined as a natural discretised version of the generator of X. This makes the CTMC X^h into a continuous-time random walk, which is skip-free (i.e. simple) if X is without jumps (i.e. Brownian motion with drift). The quantity:

$$\kappa(\delta) := \int_{[-1,1]\setminus[-\delta,\delta]} |x| d\lambda(x), \quad \delta \geq 0,$$

where λ is the Lévy measure of X, is related to the activity of the small jumps of X and plays a crucial role in the rate of convergence. We assume that either the diffusion component of X is present ($\sigma^2 > 0$) or the jump activity of X is sufficient (Orey's condition [27], see also Assumption 2.3 below) to ensure that X admits continuous transition densities $p_{t,T}(x,y)$ (from x at time t to y at time $T > t$), which are our main object of study.

Let $P_{t,T}^h(x,y) := \mathsf{P}(X_T^h = y | X_t^h = x)$ denote the corresponding transition probabilities of X^h and let

$$\Delta_{T-t}(h) := \sup_{x,y \in \mathbb{Z}_h} \left| p_{t,T}(x,y) - \frac{1}{h} P_{t,T}^h(x,y) \right|.$$

The following table summarizes our result (for functions $f \geq 0$ and $g > 0$, we shall write $f = O(g)$ (resp. $f = o(g)$, $f \sim g$) for $\limsup_{h\downarrow 0} f(h)/g(h) < \infty$ (resp. $\lim_{h\downarrow 0} f(h)/g(h) = 0$, $\lim_{h\downarrow 0} f(h)/g(h) \in (0,\infty)$) — if g converges to 0, then we will say f decays no slower than (resp. faster than, at the same rate as) g):

	$\sigma^2 > 0$	$\sigma^2 = 0$
$\lambda(\mathbb{R}) = 0$	$\Delta_{T-t}(h) = O(h^2)$	$\times$
$0 < \lambda(\mathbb{R}) < \infty$	$\Delta_{T-t}(h) = O(h)$	$\times$
$\lambda(\mathbb{R}) = \infty$	$\Delta_{T-t}(h) = O(h\kappa(h/2))$	

We also prove that the rates stated here are sharp in the sense that there exist Lévy processes for which convergence is no better than stated.

Note that the rate of convergence depends on the Lévy measure λ, it being best when $\lambda = 0$ (quadratic when $\sigma^2 > 0$), and linear otherwise, unless the pure jump part of X has infinite variation, in which case it depends on the quantity κ. This is due to the nature of the discretisation of the Brownian motion with drift (which gives a quadratic order of convergence, when $\sigma^2 > 0$), and then of the Lévy measure, which is aggregated over intervals of length h around each of the lattice points; see also (v) of Remark 3.11. In the infinite activity case, $\kappa(h) = o(1/h)$, indeed κ is bounded, if in addition $\kappa(0) < \infty$. However, the convergence of $h\kappa(h/2)$ to zero, as $h \downarrow 0$, can be arbitrarily slow. Finally, if X is a compound Poisson process (i.e. $\lambda(\mathbb{R}) \in (0,\infty)$) without a diffusion component, but possibly with a drift, there is always an atom present in the law of X at a fixed time, which is why the finite Lévy measure case is studied only when $\sigma^2 > 0$.

By way of example, note that if $\lambda([-1,1]\setminus[-h,h]) \sim 1/h^{1+\alpha}$ for some $\alpha \in (0,1)$, then $\kappa(h) \sim h^{-\alpha}$ and the convergence of the normalized probability mass function to the transition density is by Theorem 2.4 of order $h^{1-\alpha}$, since $\kappa(0) = \infty$ and Orey's condition

is satisfied. In particular, in the case of the CGMY [5] (tempered stable) or β-stable [30, p. 80] processes with stability parameter $\beta \in (1,2)$, we have $\alpha = \beta - 1$ and hence convergence of order $h^{2-\beta}$. More generally, if $\beta := \inf\{p > 0 : \int_{[-1,1]} |x|^p d\lambda(x) < \infty\}$ is the *Blumenthal-Getoor index* of X, and $\beta \geq 1$, then for any $p > \beta$ we have $\kappa(h) = O(h^{1-p})$. Conversely, if for some $p \geq 1$, $\kappa(h) = O(h^{1-p})$, then $\beta \leq p$.

The proof of Theorem 2.4 is in two steps: we first establish the convergence rate of the characteristic exponent of X_t^h to that of X_t (Subsection 3.2). In the second step we apply this to the study of the convergence of transition densities (Section 4) via their spectral representations (established in Subsection 3.1). Note that in general the rates of convergence of the characteristic functions do not carry over directly to the distribution functions. We are able to follow through the above programme by exploiting the special structure of the infinitely divisible distributions in what amounts to a detailed comparison of the transition kernels $p_{t,T}(x,y)$ and $P_{t,T}^h(x,y)$.

This gives the overall picture in dimension one. In dimensions higher than one ($d > 1$), and then under a straightforward extension of the discretization described above, essentially the same rates of convergence are obtained as in the univariate case; this time under a technical condition (cf. Assumption 2.6), which is satisfied when the diffusion-matrix is non-degenerate. Our main result in this case is Theorem 2.8.

1.2 Literature overview

In general, there has been a plethora of publications devoted to the subject of discretization schemes for stochastic processes, see e.g. [19], and with regard to the pricing of financial derivatives [15] and the references therein. In particular, there exists a wealth of literature concerning approximations of Lévy processes in one form or another and a brief overview of simulation techniques is given by [29].

In continuous time, for example, [18] approximates by replacing the small jumps part with a diffusion, and discusses also rates of convergence for $\mathsf{E}[g \circ X_T]$, where g is real-valued and satisfies certain integrability conditions, T is a fixed time and X the process under approximation; [9] approximates by a combination of Brownian motion and sums of compound Poisson processes with two-sided exponential densities. In discrete time, Markov chains have been used to approximate the much larger class of Feller processes and [4] proves convergence in law of such an approximation in the Skorokhod space of càdlàg paths, but does not discuss rates of convergence; [33] has a finite state-space path approximation and applies this to option pricing together with a discussion of the rates of convergence for the prices. With respect to Lévy process driven SDEs, [21] (resp. [35]) approximates solutions Y thereto using a combination of a compound Poisson process and a high order scheme for the Brownian component (resp. discrete-time Markov chains and an operator approach) — rates of convergence are then discussed for expectations of sufficiently regular real-valued functions against the marginals of the solutions.

We remark that approximation/simulation of Lévy processes in dimensions higher than one is in general more difficult than in the univariate case, see, e.g. the discussion on this in [6] (which has a Gaussian approximation and establishes convergence in the Skorokhod space [6, p. 197, Theorem 2.2]). Observe also that in terms of pricing theory, the probability density function of a process can be viewed as the Arrow-Debreu state price, i.e. the current value of an option whose payoff equals the Dirac delta function. The singular nature of this payoff makes it hard, particularly in the presence of jumps, to study the convergence of the prices under the discretised process to its continuous counterpart.

Indeed, Theorem 2.8 can be viewed as a generalisation of such convergence results for the well-known discretisation of the multi-dimensional Black-Scholes model

(see e.g. [24] for the case of Brownian motion with drift in dimension one). In addition, existing literature, as specific to approximations of densities of Lévy processes (or generalizations thereof), includes [12] (polynomial expansion for a bounded variation driftless pure-jump process) and [13] (density expansions for multivariate affine jump-diffusion processes). [20, 34] study upper estimates for the densities. On the other hand [2] has a result similar in spirit to ours, but for solutions to SDEs: for the case of the Euler approximation scheme, the authors there also study the rate of convergence of the transition densities.

Further, from the point of view of partial integro-differential equations (PIDEs), the density $p : (0, \infty) \times \mathbb{R}^d \to [0, \infty)$ of the Lévy process X is the classical fundamental solution of the Cauchy problem (in $u \in C_0^{1,2}((0, \infty), \mathbb{R}^d))$ $\frac{\partial u}{\partial t} = Lu$, L being the infinitesimal generator of X [7, Chapter 12] [14, Chapter IV]. Note that Assumption 2.3 in dimension one (resp. Assumption 2.6 in the multivariate case) guarantees $p \in C_0^{1,\infty}$. There are numerous numerical methods in dealing with such PIDEs (and PIDEs in general): fast Fourier transform, trees and discrete-time Markov chains, viscosity solutions, Galerkin methods, see, e.g. [8, Subsection 1.1] [7, Subsections 12.3-12.7] and the references therein. In particular, we mention the finite-difference method, which is in some sense the counterpart of the present article in the numerical analysis literature, discretising both in space and time, whereas we do so only in space. In general, this literature often restricts to finite activity processes, and either avoids a rigorous analysis of (the rates of) convergence, or, when it does, it does so for initial conditions $h = u(0, \cdot)$, which exclude the singular δ-distribution. For example, [8, p. 1616, Assumption 6.1] requires h continuous, piecewise C^∞ with bounded derivatives of all orders; compare also Propositions 5.1 and 5.4 concerning convergence of expectations in our setting. Moreover, unlike in our case where the discretisation is made outright, the approximation in [8] is sequential, as is typical of the literature: beyond the restriction to a bounded domain (with boundary conditions), there is a truncation of the integral term in L, and then a reduction to the finite activity case, at which point our results are in agreement with what one would expect from the linear order of convergence of [8, p. 1616, Theorem 6.7].

The rest of the paper is organised as follows. Section 2 introduces the setting by specifying the Markov generator of X^h and precisely states the main results. Then Section 3 provides integral expressions for the transition kernels by applying spectral theory to the generator of the approximating chain and studies the convergence of the characteristic exponents. In section 4 this allows us to establish convergence rates for the transition densities. While Sections 3 and 4 restrict this analysis to the univariate case, explicit comments are made in both, on how to extend the results to the multivariate setting (this extension being, for the most part, direct and trivial). Finally, Section 5 derives some results regarding convergence of expectations $\mathsf{E}[f \circ X_t^h] \to \mathsf{E}[f \circ X_t]$ for suitable test functions f; presents a numerical algorithm, under which computations are eventually done; discusses the corresponding truncation/localization error and gives some numerical experiments.

2 Definitions, notation and statement of results

2.1 Setting

Fix a dimension $d \in \mathbb{N}$ and let $(e_j)_{j=1}^d$ be the standard orthonormal basis of $\mathbb{R}^d$. Further, let X be an $\mathbb{R}^d$-valued Lévy process with characteristic exponent [30, pp. 37-39]:

$$\Psi(p) = -\frac{1}{2} \langle p, \Sigma p \rangle + i \langle \mu, p \rangle + \int_{\mathbb{R}^d} \left(e^{i\langle p, x \rangle} - i\langle p, x \rangle \mathbb{1}_{[-V,V]^d}(x) - 1 \right) d\lambda(x) \qquad (2.1)$$

($p \in \mathbb{R}^d$). Here $(\Sigma, \lambda, \mu)_{\tilde{c}}$ is the characteristic triplet relative to the cut-off function $\tilde{c} = \mathbb{1}_{[-V,V]^d}$; V is 1 or 0, the latter only if $\int_{[-1,1]^d} |x| d\lambda(x) < \infty$. Note that X is then a Markov process with transition function $P_{t,T}(x, B) := \mathsf{P}(X_{T-t} \in B - x)$ ($0 \le t \le T$, $x \in \mathbb{R}^d$ and $B \in \mathcal{B}(\mathbb{R}^d)$) and (for $t \ge 0$, $p \in \mathbb{R}^d$) $\phi_{X_t}(p) := \mathsf{E}[e^{i\langle p, X_t \rangle}] = \exp\{t\Psi(p)\}$. We refer to [3, 30] for the general background on Lévy processes.

Since $\Sigma \in \mathbb{R}^{d \times d}$ is symmetric, nonnegative definite, it is assumed without loss of generality that $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_d^2)$ with $\sigma_1^2 \ge \cdots \ge \sigma_d^2$. We let $l := \max\{k \in \{1, \ldots, d\} : \sigma_k^2 > 0\}$ ($\max \emptyset := 0$). In the univariate case $d = 1$, Σ reduces to the scalar $\sigma^2 := \sigma_1^2$.

Now fix $h > 0$. Consider a CTMC $X^h = (X_t^h)_{t \ge 0}$ approximating our Lévy process X (in law). We describe (see [26] for a general reference on CTMCs) X^h as having state space $\mathbb{Z}_h^d := h\mathbb{Z}^d := \{hk : k \in \mathbb{Z}^d\}$ ($\mathbb{Z}_h := \mathbb{Z}_h^1$), initial state $X_0^h = 0$, a.s. and an infinitesimal generator L^h given by a spatially homogeneous Q-matrix Q^h (i.e. $Q_{ss'}^h$ depends only on $s - s'$, for $\{s, s'\} \subset \mathbb{Z}_h^d$). Thus L^h is a mapping defined on the set $l^\infty(\mathbb{Z}_h^d)$ of bounded functions f on $\mathbb{Z}_h^d$, and $L^h f(s) = \sum_{s' \in \mathbb{Z}_h^d} Q_{ss'}^h f(s')$.

It remains to specify Q^h. To this end we discretise on $\mathbb{Z}_h^d$ the infinitesimal generator L of the Lévy process X, thus obtaining L^h. Recall that [30, p. 208, Theorem 31.5]:

$$Lf(x) = \sum_{j=1}^d \left(\frac{\sigma_j^2}{2} \partial_{jj} f(x) + \mu_j \partial_j f(x) \right) + \int_{\mathbb{R}^d} \left(f(x+y) - f(x) - \sum_{j=1}^d y_j \partial_j f(x) \mathbb{1}_{[-V,V]^d}(y) \right) d\lambda(y)$$

($f \in C_0^2(\mathbb{R}^d)$, $x \in \mathbb{R}^d$). We specify L^h separately in the univariate, $d = 1$, and in the general, multivariate, setting.

2.1.1 Univariate case

In the case when $d = 1$, we introduce two schemes. Referred to as **discretization scheme 1** (resp. **2**), and given by (2.2) (resp. (2.4)) below, they differ in the discretization of the first derivative, as follows.

Under **discretisation scheme 1**, for $s \in \mathbb{Z}_h$ and $f : \mathbb{Z}_h \to \mathbb{R}$ vanishing at infinity:

$$L^h f(s) = \frac{1}{2} \left(\sigma^2 + c_0^h \right) \frac{f(s+h) + f(s-h) - 2f(s)}{h^2} + \left(\mu - \mu^h \right) \frac{f(s+h) - f(s-h)}{2h} +$$
$$\sum_{s' \in \mathbb{Z}_h \setminus \{0\}} [f(s+s') - f(s)] c_{s'}^h \qquad (2.2)$$

where the following notation has been introduced:

- for $s \in \mathbb{Z}_h$:

$$A_s^h := \begin{cases} [s - h/2, s + h/2), & \text{if } s < 0 \\ [-h/2, h/2], & \text{if } s = 0 \ ; \\ (s - h/2, s + h/2], & \text{if } s > 0 \end{cases}$$

- for $s \in \mathbb{Z}_h \setminus \{0\}$: $c_s^h := \lambda(A_s^h)$;

- and finally:

$$c_0^h := \int_{A_0^h} y^2 \mathbb{1}_{[-V,V]}(y) d\lambda(y) \quad \text{and} \quad \mu^h := \sum_{s \in \mathbb{Z}_h \setminus \{0\}} s \int_{A_s^h} \mathbb{1}_{[-V,V]}(y) d\lambda(y).$$

Note that Q^h has nonnegative off-diagonal entries for all h for which:

$$\frac{\sigma^2 + c_0^h}{2h^2} + \frac{\mu - \mu^h}{2h} + c_h^h \ge 0 \quad \text{and} \quad \frac{\sigma^2 + c_0^h}{2h^2} - \frac{\mu - \mu^h}{2h} + c_{-h}^h \ge 0 \qquad (2.3)$$

and in that case Q^h is a genuine Q-matrix. Moreover, due to spatial homogeneity, its entries are then also uniformly bounded in absolute value.

Further, when $\sigma^2 > 0$, it will be shown that (2.3) always holds, at least for all sufficiently small h (see Proposition 3.9). However, in general, (2.3) may fail. It is for this reason that we introduce scheme 2, under which the condition on the nonnegativity of off-diagonal entries of Q^h holds vacuously.

To wit, we use in **discretization scheme 2** the one-sided, rather than the two-sided discretisation of the first derivative, so that (2.2) reads:

$$
\begin{aligned}
L^h f(s) \;=\; & \frac{1}{2}\left(\sigma^2 + c_0^h\right)\frac{f(s+h)+f(s-h)-2f(s)}{h^2} + \sum_{s' \in \mathbb{Z}_h \backslash \{0\}} [f(s+s') - f(s)]c_{s'}^h + \\
& (\mu - \mu^h)\left(\frac{f(s+h)-f(s)}{h}\mathbb{1}_{[0,\infty)}(\mu - \mu^h) + \frac{f(s)-f(s-h)}{h}\mathbb{1}_{(-\infty,0]}(\mu - \mu^h)\right) \quad (2.4)
\end{aligned}
$$

Importantly, while scheme 2 is always well-defined, scheme 1 is not; and yet the two-sided discretization of the first derivative exhibits better convergence properties than the one-sided one (cf. Proposition 3.10). We therefore retain the treatment of both these schemes in the sequel.

For ease of reference we also summarize here the following notation which will be used from Subsection 3.2 onwards:

$$
c := \lambda(\mathbb{R}), \quad b := \kappa(0), \quad d := \lambda(\mathbb{R} \backslash [-1,1])
$$

and for $\delta \in (0,1]$:

$$
\zeta(\delta) := \delta \int_{[-1,1]\backslash[-\delta,\delta]} |x| d\lambda(x) \text{ and } \gamma(\delta) := \delta^2 \int_{[-1,1]\backslash[-\delta,\delta]} d\lambda(x).
$$

2.1.2 Multivariate case

For the sake of simplicity we introduce only one discretisation scheme in this general setting. If necessary, and to avoid confusion, we shall refer to it as the **multivariate scheme**. We choose $V = 0$ or $V = 1$, according as $\lambda(\mathbb{R}^d)$ is finite or infinite. L^h is then given by:

$$
\begin{aligned}
L^h f(s) \;=\; & \frac{1}{2}\sum_{j=1}^d \left(\sigma_j^2 + c_{0j}^h\right)\frac{f(s+he_j)+f(s-he_j)-2f(s)}{h^2} + \sum_{j=1}^l (\mu_j - \mu_j^h)\frac{f(s+he_j)-f(s-he_j)}{2h} + \\
& \sum_{j=l+1}^d (\mu_j - \mu_j^h)\left(\frac{f(s+he_j)-f(s)}{h}\mathbb{1}_{[0,\infty)}(\mu_j - \mu_j^h) + \frac{f(s)-f(s-he_j)}{h}\mathbb{1}_{(-\infty,0]}(\mu_j - \mu_j^h)\right) + \\
& \sum_{s' \in \mathbb{Z}_h^d} \left(f(s+s') - f(s)\right)c_{s'}^h
\end{aligned}
$$

($f \in c_0(\mathbb{Z}_h^d)$, $s \in \mathbb{Z}_h^d$; and we agree $\sum_\emptyset := 0$). Here the following notation has been introduced:

- for $s \in \mathbb{Z}_h^d$: $A_s^h := \prod_{j=1}^d I_{s_j}^h$, where for $s \in \mathbb{Z}_h$:

$$
I_s^h := \begin{cases} [s-h/2, s+h/2), & \text{if } s < 0 \\ [-h/2, h/2], & \text{if } s = 0 \\ (s-h/2, s+h/2], & \text{if } s > 0 \end{cases}
$$

so that $\{A_s^h : s \in \mathbb{Z}_h^d\}$ constitutes a partition of $\mathbb{R}^d$;

- for $s \in \mathbb{Z}_h^d \backslash \{0\}$: $c_s^h := \lambda(A_s^h)$;

- and finally for $j \in \{1, \ldots, d\}$:

$$c_{0j}^h := \int_{A_0^h} x_j^2 \mathbb{1}_{[-V,V]^d}(x) d\lambda(x) \qquad \text{and} \qquad \mu_j^h := \sum_{s \in \mathbb{Z}_h^d \setminus \{0\}} s_j \int_{A_s^h} \mathbb{1}_{[-V,V]^d}(y) d\lambda(y).$$

Notice that when $d = 1$, this scheme reduces to scheme 1 or scheme 2, according as $\sigma^2 > 0$ or $\sigma^2 = 0$. Indeed, statements pertaining to the multivariate scheme will always be understood to include also the univariate case $d = 1$.

Remark 2.1. *The complete analogue of c_0^h from the univariate case would be the matrix c_0^h, entries $(c_0^h)_{ij} := \int_{A_0^h} x_i x_j \mathbb{1}_{[-V,V]^d}(x) d\lambda(x)$, $\{i, j\} \subset \{1, \ldots, d\}$. However, as h varies, so could c_0^h, and thus no diagonalization of $c_0^h + \Sigma$ possible (in general), simultaneously in all (small enough) positive h. Thus, retaining c_0^h in its totality, we should have to discretize mixed second partial derivatives, which would introduce (further) nonpositive entries in the corresponding Q-matrix Q^h of X^h. It is not clear whether these would necessarily be counter-balanced in a way that would ensure nonnegative off-diagonal entries. Retaining the diagonal terms of c_0^h, however, is of no material consequence in this respect.*

It is verified just as in the univariate case, component by component, that there is some $h_\star \in (0, +\infty]$ such that for all $h \in (0, h_\star]$, L^h is indeed the infinitesimal generator of some CTMC (i.e. the off-diagonal entries of Q^h are nonnegative). Q^h is then a regular (as spatially homogeneous) Q-matrix, and X^h is a compound Poisson process, whose Lévy measure we denote λ^h.

2.2 Summary of results

We have (see Remark 3.11(iii) pursuant to Proposition 3.10 for proof):

Remark 2.2 (Convergence in distribution). *X^h converges to X, as $h \downarrow 0$, weakly in finite-dimensional distributions (hence w.r.t. the Skorokhod topology on the space of càdlàg paths [17, p. 415, Corollary 3.9]).*

Next, in order to formulate the rates of convergence, recall that $P_{t,T}^h(x,y)$ (resp. $p_{t,T}(x,y)$) denote the transition probabilities (resp. continuous transition densities, when they exist) of X^h (resp. X) from x at time t to y at time T, $\{x,y\} \subset \mathbb{Z}_h^d$, $0 \leq t < T$. Further, for $0 \leq t < T$ define:

$$\Delta_{T-t}(h) := \sup_{\{x,y\} \subset \mathbb{Z}_h^d} D_{t,T}^h(x,y) \qquad \text{where} \qquad D_{t,T}^h(x,y) := \left| p_{t,T}(x,y) - \frac{1}{h^d} P_{t,T}^h(x,y) \right|. \qquad (2.5)$$

We now summarize the results first in the univariate, and then in the multivariate setting (Remark 2.2 holding true of both).

2.2.1 Univariate case

The assumption alluded to in the introduction is the following (we state it explicitly when it is being used):

Assumption 2.3. *Either $\sigma^2 > 0$ or Orey's condition [27] holds:*

$$\exists \epsilon \in (0,2) \qquad \text{such that} \qquad \liminf_{r \downarrow 0} \frac{1}{r^{2-\epsilon}} \int_{[-r,r]} u^2 d\lambda(u) > 0.$$

The usage of the two schemes and the specification of V is as summarized in Table 1. In short we use scheme 1 or scheme 2, according as $\sigma^2 > 0$ or $\sigma^2 = 0$, and we use $V = 0$

Table 1: Usage of the two schemes and of V depending on the nature of σ^2 and λ.

Lévy measure/diffusion part	$\sigma^2 > 0$	$\sigma^2 = 0$
$\lambda(\mathbb{R}) < \infty$	scheme 1, $V = 0$	scheme 2, $V = 0$
$\lambda(\mathbb{R}) = \infty$	scheme 1, $V = 1$	scheme 2, $V = 1$

or $V = 1$, according as $\lambda(\mathbb{R}) < \infty$ or $\lambda(\mathbb{R}) = \infty$. By contrast to Assumption 2.3 we maintain Table 1 as being in effect throughout this paragraph (Paragraph 2.2.1).

Under Assumption 2.3 for every $t > 0$, $\phi_{X_t} \in L^1(m)$ where m is Lebesgue measure and (for $0 \le t < T$, $\{x, y\} \subset \mathbb{R}$):

$$p_{t,T}(x, y) = \frac{1}{2\pi} \int_{\mathbb{R}} \exp\{ip(x - y)\} \exp\{\Psi(p)(T - t)\}\, dp \tag{2.6}$$

(see Remark 3.1). Similarly, with Ψ^h denoting the characteristic exponent of the compound Poisson process X^h (for $0 \le t < T$, $y \in \mathbb{Z}_h$, $\mathsf{P}_{X_t^h}$-a.s. in $x \in \mathbb{Z}_h$):

$$\frac{1}{h} P_{t,T}^h(x, y) = \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \exp\{ip(x - y)\} \exp\{\Psi^h(p)(T - t)\} dp \tag{2.7}$$

(see Proposition 3.5). Note that the right-hand side is defined even if $\mathsf{P}(X_t^h = x) = 0$ and we let the left-hand side take this value when this is so.

The main result can now be stated.

Theorem 2.4 (Convergence of transition kernels). *Under Assumption 2.3, whenever $s > 0$, the convergence of $\Delta_s(h)$ is summarized in the following table. In general convergence is no better than stipulated.*

	$\lambda(\mathbb{R}) = 0$	$0 < \lambda(\mathbb{R}) < \infty$	$\kappa(0) < \infty = \lambda(\mathbb{R})$	$\kappa(0) = \infty$
$\sigma^2 > 0$	$\Delta_s(h) = O(h^2)$	$\Delta_s(h) = O(h)$	$\Delta_s(h) = O(h)$	$\Delta_s(h) = O(h\kappa(h/2))$
$\sigma^2 = 0$	$\times$	$\times$		

More exhaustive statements, of which this theorem is a summary, are to be found in Propositions 4.5 and 4.6, and will be proved in Section 4. The proof of Theorem 2.4 itself can be found at the end of Section 4.

Remark 2.5. *Assumption 2.3 implies that X_t, for any $t > 0$, has a smooth density [30, p. 190, Proposition 28.3]. It hence appears to be unlikely that this assumption constitutes a necessary condition for the convergence rates of Theorem 2.4 to hold. In particular, Assumption 2.6 with $d = 1$, stipulating a certain exponential decay of the characteristic exponents, is implied by Assumption 2.3 (see Remark 3.1 and Proposition 4.1) but sufficient for the validity of the convergence rates in Theorem 2.4 (see Theorem 2.8).*

2.2.2 Multivariate case

The relevant technical condition here is:

Assumption 2.6. *There are $\{P, C, \epsilon\} \subset (0, \infty)$ and an $h_0 \in (0, h_\star]$, such that for all $h \in (0, h_0)$, $s > 0$ and $p \in [-\pi/h, \pi/h]^d \setminus (-P, P)^d$:*

$$|\phi_{X_s^h}(p)| \le \exp\{-Cs|p|^\epsilon\} \tag{2.8}$$

whereas for $p \in \mathbb{R}^d \setminus (-P, P)^d$:

$$|\phi_{X_s}(p)| \le \exp\{-Cs|p|^\epsilon\}. \tag{2.9}$$

Again we shall state it explicitly when it is being used.

Remark 2.7. *It is shown, just as in the univariate case, that Assumption 2.6 holds if $l = d$, i.e. if Σ is non-degenerate. Moreover, then we may take $P = 0$, $C = \frac{1}{2} \left(\frac{2}{\pi} \right)^2 \left(\wedge_{j=1}^d \sigma_j^2 \right)$, $\epsilon = 2$ and $h_0 = h_\star$.*

It would be natural to expect that the same could be verified for the multivariate analogue of Orey's condition, which we suggest as being:

$$\liminf_{r\downarrow 0} \inf_{e \in S^{d-1}} \int_{\overline{B}(0,r)} |\langle e, x \rangle|^2 d\lambda(x)/r^{2-\epsilon} > 0$$

for some $\epsilon \in (0, 2)$ (with $S^{d-1} \subset \mathbb{R}^d$ (resp. $\overline{B}(0, r) \subset \mathbb{R}^d$) the unit sphere (resp. closed ball of radius r centered at the origin)). Specifically, under this condition, it is easy to see that (2.9) of Assumption 2.6 still holds. However, we are unable to show the validity of (2.8).

Under Assumption 2.6, Fourier inversion yields the integral representation of the continuous transition densities for X (for $0 \le t < T$, $\{x, y\} \subset \mathbb{R}^d$):

$$p_{t,T}(x,y) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\langle p, x-y \rangle} e^{(T-t)\Psi(p)} dp.$$

On the other hand, $L^2([-\pi/h, \pi/h]^d)$ Hilbert space techniques yield for the normalized transition probabilities of X^h (for $0 \le t < T$, $y \in \mathbb{Z}_h^d$ and $\mathsf{P}_{X_t^h}$-a.s. in $x \in \mathbb{Z}_h^d$):

$$\frac{1}{h^d} P_{t,T}(x,y) = \frac{1}{(2\pi)^d} \int_{[-\pi/h, \pi/h]^d} e^{i\langle p, x-y \rangle} e^{(T-t)\Psi^h(p)} dp,$$

where Ψ^h is the characteristic exponent of X^h.

Finally, we state the result with the help of the following notation:

- $\kappa(\delta) := \int_{[-1,1]^d \setminus [-\delta, \delta]^d} |x| d\lambda(x)$, $\zeta(\delta) := \delta \kappa(\delta)$ and $\chi(\delta) := \sum_{1 \le i < j \le d} \int_{[-\delta, \delta]^d} |x_i x_j| d\lambda(x)$ (where $\delta \in [0, \infty)$).

- $\hat{\sigma}^2 := \wedge_{j=1}^d \sigma_j^2$ and $\sigma^2 := \sum_{j=1}^d \sigma_j^2$.

Note that by the dominated convergence theorem, $(\zeta + \chi)(\delta) \to 0$ as $\delta \downarrow 0$ (this is seen as in the univariate case, cf. Lemma 3.8).

Theorem 2.8 (Convergence — multivariate case). *Let $d \in \mathbb{N}$ and suppose Assumption 2.6 holds. Then for any $s > 0$, $\Delta_s(h) = O(h \vee (\zeta + \chi)(h/2))$. Moreover, if $\hat{\sigma}^2 > 0$, then there exists a universal constant $D_d \in (0, \infty)$, such that:*

1. If $\lambda(\mathbb{R}^d) = 0$,

$$\limsup_{h\downarrow 0} \frac{\Delta_s(h)}{h^2} \le D_d \left[\frac{\sigma^2}{\hat{\sigma}^2} \frac{1}{\sqrt{s\hat{\sigma}^2}} + \frac{|\mu|}{\hat{\sigma}^2} \right] \frac{1}{(s\hat{\sigma}^2)^{\frac{d+1}{2}}}.$$

2. If $0 < \lambda(\mathbb{R}^d) < \infty$,

$$\limsup_{h\downarrow 0} \frac{\Delta_s(h)}{h} \le D_d \frac{\lambda(\mathbb{R}^d) s}{(s\hat{\sigma}^2)^{\frac{d+1}{2}}}.$$

3. If $\kappa(0) < \infty = \lambda(\mathbb{R}^d)$,

$$\limsup_{h\downarrow 0} \frac{\Delta_s(h)}{h} \le D_d \left[\lambda(\mathbb{R}^d \setminus [-1,1]^d) s + \frac{\kappa(0)s}{\sqrt{s\hat{\sigma}^2}} \right] \frac{1}{(s\hat{\sigma}^2)^{\frac{d+1}{2}}}.$$

4. If $\kappa(0) = \infty$,

$$\limsup_{h \downarrow 0} \frac{\Delta_s(h)}{(\zeta + \chi)(h/2)} \le D_d \frac{s}{(s\hat{\sigma}^2)^{\frac{d+2}{2}}}.$$

Remark 2.9. *Notice that in the univariate case $\zeta + \chi$ reduces to ζ. The presence of χ is a consequence of the omission of non-diagonal entries of c_0^h in the multivariate approximation scheme (cf. Remark 2.1).*

The proof of Theorem 2.8 is an easy extension of the arguments behind Theorem 2.4 and can be found immediately following the proof of Proposition 4.2.

3 Transition kernels and convergence of characteristic exponents

In the interest of space, simplicity of notation and ease of exposition, the analysis in this and in Section 4 is *restricted* to dimension $d = 1$. Proofs in the multivariate setting are, for the most part, a direct and trivial extension of those in the univariate case. However, when this is not so, necessary and explicit comments will be provided in the sequel, as appropriate.

3.1 Integral representations

First we note the following result (its proof is essentially by the standard inversion theorem, see also [30, p. 190, Proposition 28.3]).

Remark 3.1. *Under Assumption 2.3, for some $\{P, C, \epsilon\} \subset (0, \infty)$ depending only on $\{\lambda, \sigma^2\}$ and then all $p \in \mathbb{R} \backslash (-P, P)$ and $t \ge 0$: $|\phi_{X_t}(p)| \le \exp\{-Ct|p|^\epsilon\}$. Moreover, when $\sigma^2 > 0$, one may take $P = 0$, $C = \frac{1}{2}\sigma^2$ and $\epsilon = 2$, whereas otherwise ϵ may take the value from Orey's condition in Assumption 2.3. Consequently, X_t ($t > 0$) admits the continuous density $f_{X_t}(y) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ipy} \phi_{X_t}(p) dp$ ($y \in \mathbb{R}$). In particular, the law $P_{t,T}(x, \cdot)$ is given by (2.6).*

Second, to obtain (2.7) we apply some classical theory of Hilbert spaces, see e.g. [10].

Definition 3.2. *For $s \in \mathbb{Z}_h$ let $g_s : [-\frac{\pi}{h}, \frac{\pi}{h}] \to \mathbb{C}$ be given by $g_s(p) := \sqrt{\frac{h}{2\pi}} e^{-isp}$. The $(g_s)_{s \in \mathbb{Z}_h}$ constitute an orthonormal basis of the Hilbert space $L^2([-\frac{\pi}{h}, \frac{\pi}{h}])$.*
Let $A \in l^2(\mathbb{Z}_h)$. We define: $\mathcal{F}_h A := \sum_{s \in \mathbb{Z}_h} A(s) g_s$. The inverse of this transform $\mathcal{F}_h^{-1} : L^2([-\frac{\pi}{h}, \frac{\pi}{h}]) \to l^2(\mathbb{Z}_h)$ is given by:

$$(\mathcal{F}_h^{-1} \phi)(s) = \langle \phi, g_s \rangle := \int_{[-\frac{\pi}{h}, \frac{\pi}{h}]} \phi(u) \overline{g_s}(u) du$$

for $\phi \in L^2([-\frac{\pi}{h}, \frac{\pi}{h}])$ and $s \in \mathbb{Z}_h$.

Definition 3.3. *For a bounded linear operator $A : l^2(\mathbb{Z}_h) \to l^2(\mathbb{Z}_h)$, we say $F_A : [-\pi/h, \pi/h] \to \mathbb{R}$ is its diagonalization, if $\mathcal{F}_h A \mathcal{F}_h^{-1} \phi = F_A \phi$ for all $\phi \in L^2([-\frac{\pi}{h}, \frac{\pi}{h}])$.*

We now diagonalize L^h, which allows us to establish (2.7). The straightforward proof is left to the reader.

Proposition 3.4. *Fix $C \in l^1(\mathbb{Z}_h)$. The following introduces a number of bounded linear operators $A : l^2(\mathbb{Z}_h) \to l^2(\mathbb{Z}_h)$ and gives their diagonalization. With $f \in l^2(\mathbb{Z}_h)$, $s \in \mathbb{Z}_h$, $p \in [-\frac{\pi}{h}, \frac{\pi}{h}]$:*

(i) $\Delta_h f(s) := \frac{f(s+h) + f(s-h) - 2f(s)}{h^2}$. $F_{\Delta_h}(p) = 2\frac{\cos(hp) - 1}{h^2}$.

(ii) $\nabla_h f(s) := \frac{f(s+h)-f(s-h)}{2h}$. $F_{\nabla_h}(p) = i\frac{\sin(hp)}{h}$. *Under scheme 2 we let* $\nabla_h^+ f(s) := \frac{f(s+h)-f(s)}{h}$ *(resp.* $\nabla_h^- f(s) := \frac{f(s)-f(s-h)}{h}$*) and then* $F_{\nabla_h^+}(p) = \frac{e^{ihp}-1}{h}$ *(resp.* $F_{\nabla_h^-}(p) = \frac{1-e^{-ihp}}{h}$*).*

(iii) $L_C f(s) := \sum_{s'\in\mathbb{Z}_h}(f(s+s')-f(s))C(s')$. $F_{L_C}(p) = \sum_{s\in\mathbb{Z}_h} C(s)(e^{isp}-1)$.

As λ is finite outside any neighborhood of 0, $L^h|_{l^2(\mathbb{Z}_h)}$ (as in (2.2), resp. (2.4)) is a bounded linear mapping. We denote this restriction by L^h also. Its diagonalization is then given by $\Psi^h := F_{L^h}$, where, under scheme 1,

$$\Psi^h(p) = i(\mu - \mu^h)\frac{\sin(hp)}{h} + (\sigma^2 + c_0^h)\frac{(\cos(hp)-1)}{h^2} + \sum_{s\in\mathbb{Z}_h\setminus\{0\}} c_s^h\left(e^{isp}-1\right) \qquad (3.1)$$

and under scheme 2,

$$
\begin{aligned}
\Psi^h(p) &= (\mu - \mu^h)\left(\frac{e^{ihp}-1}{h}\mathbb{1}_{[0,\infty)}(\mu-\mu^h) + \frac{1-e^{-ihp}}{h}\mathbb{1}_{(-\infty,0]}(\mu-\mu^h)\right) + \\
&\quad + (\sigma^2 + c_0^h)\frac{(\cos(hp)-1)}{h^2} + \sum_{s\in\mathbb{Z}_h\setminus\{0\}} c_s^h\left(e^{isp}-1\right) \qquad (3.2)
\end{aligned}
$$

(with $p \in [-\frac{\pi}{h}, \frac{\pi}{h}]$, but we can and will view Ψ^h as defined for all real p by the formulae above). Under either scheme, Ψ^h is bounded and continuous as the final sum converges absolutely and uniformly.

Proposition 3.5. *For scheme 1 under Equation* (2.3) *and always for scheme 2, for every* $0 \le t < T$, $y \in \mathbb{Z}_h$ *and* $\mathsf{P}_{X_t^h}$*-a.s. in* $x \in \mathbb{Z}_h$ (2.7) *holds, i.e.:*

$$\mathsf{P}(X_T^h = y | X_t^h = x) = \frac{h}{2\pi}\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \exp\{ip(x-y)\}\exp\{\Psi^h(p)(T-t)\}dp.$$

Proof. (Condition (2.3) ensures scheme 1 is well-defined (Q^h needs to have nonnegative off-diagonal entries).) Note that: $\mathsf{P}(X_T^h = y | X_t^h = x) = (e^{(T-t)L^h}\mathbb{1}_{\{y\}})(x)$. Thus (2.7) follows directly from the relation $\mathcal{F}_h L^h \mathcal{F}_h^{-1} = \Psi^h\cdot$ (where $\Psi^h\cdot$ is the operator that multiplies functions pointwise by Ψ^h). $\qquad\square$

In what follows we study the convergence of (2.7) to (2.6) as $h \downarrow 0$. These expressions are particularly suited to such an analysis, not least of all because the spatial and temporal components are factorized.

One also checks that for every $t \ge 0$ and $p \in \mathbb{R}$:

$$\phi_{X_t^h}(p) = \mathsf{E}[e^{ipX_t^h}] = \exp\{t\Psi^h(p)\}.$$

Hence X^h are compound Poisson processes [30, p. 18, Definition 4.2] with characteristic exponents Ψ^h.

In the multivariate scheme, by considering the Hilbert space $L^2([-\pi/h, \pi/h]^d)$ instead, X^h is again seen to be compound Poisson with characteristic exponent given by (for $p \in \mathbb{R}^d$):

$$
\begin{aligned}
\Psi^h(p) &= \sum_{j=1}^d (\sigma_j^2 + c_{0j}^h)\frac{\cos(hp_j)-1}{h^2} + i\sum_{j=1}^l (\mu_j - \mu_j^h)\frac{\sin(hp_j)}{h} + \\
&\quad + \sum_{j=l+1}^d (\mu_j - \mu_j^h)\left(\frac{e^{ihp_j}-1}{h}\mathbb{1}_{[0,\infty)}(\mu_j-\mu_j^h) + \frac{1-e^{-ihp_j}}{h}\mathbb{1}_{(-\infty,0]}(\mu_j-\mu_j^h)\right) +
\end{aligned}
$$

$$+ \sum_{s \in \mathbb{Z}_h^d \setminus \{0\}} \left(e^{i\langle p, s \rangle} - 1 \right) c_s^h. \tag{3.3}$$

In the sequel, we shall let λ^h denote the Lévy measure of X^h.

3.2 Convergence of characteristic exponents

We introduce for $p \in \mathbb{R}$:

$$f_h(p) := \frac{\cos(hp) - 1}{h^2} + \frac{p^2}{2}$$

and, under scheme 1:

$$
\begin{aligned}
g_h(p) &:= i\left(\frac{\sin(hp)}{h} - p \right) \\
l_h(p) &:= c_0^h \frac{\cos(hp) - 1}{h^2} - \mu^h i \frac{\sin(hp)}{h} + \sum_{s \in \mathbb{Z}_h \setminus \{0\}} c_s^h \left(e^{isp} - 1 \right) - \int_{\mathbb{R}} \left(e^{ipu} - 1 - ipu \mathbb{1}_{[-V,V]}(u) \right) d\lambda(u),
\end{aligned}
$$

respectively, under scheme 2:

$$
\begin{aligned}
g_h(p) &:= \frac{e^{ihp} - 1}{h} \mathbb{1}_{(0,\infty)}(\mu - \mu^h) + \frac{1 - e^{-ihp}}{h} \mathbb{1}_{(-\infty,0]}(\mu - \mu^h) - ip; \\
l_h(p) &:= c_0^h \frac{\cos(hp) - 1}{h^2} - \mu^h \left[\frac{e^{ihp} - 1}{h} \mathbb{1}_{(0,\infty)}(\mu - \mu^h) + \frac{1 - e^{-ihp}}{h} \mathbb{1}_{(-\infty,0]}(\mu - \mu^h) \right] + \\
&\quad + \sum_{s \in \mathbb{Z}_h \setminus \{0\}} c_s^h \left(e^{isp} - 1 \right) - \int_{\mathbb{R}} \left(e^{ipu} - 1 - ipu \mathbb{1}_{[-V,V]}(u) \right) d\lambda(u).
\end{aligned}
$$

Thus:

$$\Psi^h - \Psi = \sigma^2 f_h + \mu g_h + l_h.$$

Next, three elementary but key lemmas. The first concerns some elementary trigonometric inequalities as well as the Lipschitz difference for the remainder of the exponential series $f_l(x) := \sum_{k=l+1}^{\infty} \frac{(ix)^k}{k!}$ ($x \in \mathbb{R}$, $l \in \{0,1,2\}$): these estimates will be used again and again in what follows. The second is only used in the estimates pertaining to the multivariate scheme. Finally, the third lemma establishes key convergence properties relating to λ.

Lemma 3.6. *For all real* x: $0 \leq \cos(x) - 1 + \frac{x^2}{2} \leq \frac{x^4}{4!}$, $0 \leq \operatorname{sgn}(x)(x - \sin(x)) \leq \operatorname{sgn}(x)\frac{x^3}{3!}$ *and* $0 \leq x^2 + 2(1 - \cos(x)) - 2x\sin(x) \leq x^4/4$. *Whenever* $\{x, y\} \subset \mathbb{R}$ *we have (with* $\delta := y - x$):

1. $|e^{ix} - 1 - (e^{iy} - 1)|^2 \leq \delta^2$.

2. $|e^{ix} - 1 - ix - (e^{iy} - 1 - iy)|^2 \leq \delta^4/4 + \delta^2 x^2 + |\delta|^3 |x|$.

3. $|e^{ix} - 1 - ix + x^2/2 - (e^{iy} - 1 - iy + y^2/2)|^2 \leq \delta^6/36 + |\delta|^5 |x|/6 + (5/12)\delta^4 x^2 + |\delta|^3 |x|^3/2 + \delta^2 x^4/4$.

Proof. The first set of inequalities may be proved by comparison of derivatives. Then, (1) follows from $|e^{i(x-y)} - 1|^2 = 2(1 - \cos(x - y))$ and $|e^{iy}| = 1$; (2) from

$$|e^{ix} - ix - e^{iy} + iy|^2 = \left(\delta^2 + 2(1 - \cos(\delta)) - 2\delta \sin(\delta) \right) - 2\delta(\cos(x) - 1)\sin(\delta) + 2\delta \sin(x)(1 - \cos(\delta))$$

and finally (3) from the decomposition of $|e^{ix} - ix + x^2/2 - e^{iy} + iy - y^2/2|^2$ into the following terms:

1. $2(1 - \cos(\delta)) + \delta^2 + \delta^4/4 - 2\delta \sin(\delta) - (1 - \cos(\delta))\delta^2 \leq \delta^6/36$ for any real δ.

2. $\delta^3 x - \sin(x)\sin(\delta)\delta^2 = \delta^2(\delta(x - \sin(x)) + \sin(x)(\delta - \sin(\delta))) \leq |\delta|^3 |x|^3/6 + |\delta|^5 |x|/6$.

3. $-2(1-\cos(\delta))\delta x+2\delta x(1-\cos(x))(1-\cos(\delta))+2\delta(1-\cos(\delta))\sin(x) = 2\delta(1-\cos(\delta))(x(1-\cos(x)) + \sin(x) - x) \le |\delta|^3|x|^3/3$, since for all real x one has $|\sin(x) - x\cos(x)| \le |x|^3/3$.

4. $-(\cos(x) - 1)(1 - \cos(\delta))\delta^2 \le x^2\delta^4/4$.

5. $\delta^2 x^2 - 2\delta x\sin(x)\sin(\delta) - 2\delta\sin(\delta)(\cos(x) - 1) = x^2\delta(\delta - \sin(\delta)) + 2\delta\sin(\delta)(1 - \cos(x) - x\sin(x) + x^2/2) \le \delta^4 x^2/6 + \delta^2 x^4/4$ since for all real x, one has $0 \le 1 - \cos(x) - x\sin(x) + x^2/2 \le x^4/8$.

The latter inequalities are again seen to be true by comparing derivatives. □

Lemma 3.7. *Let* $\{p, x, y\} \subset \mathbb{R}^d$*. Then:*

1. $|(e^{i\langle p,x\rangle} - 1) - (e^{i\langle p,y\rangle} - 1)| \le |p||x - y|$.

2. $|(e^{i\langle p,x\rangle} - i\langle p, x\rangle - 1) - (e^{i\langle p,y\rangle} - i\langle p, y\rangle - 1)| \le 2|p|^2(|x| + |y|)|x - y|$.

Proof. This is an elementary consequence of the complex Mean Value Theorem [11, p. 859, Theorem 2.2] and the Cauchy-Schwartz inequality. □

Lemma 3.8. *For any Lévy measure* λ *on* $\mathbb{R}$*, one has for the two functions (given for* $1 \ge \delta > 0$*):* $M_0(\delta) := \delta^2 \int_{[-1,1]\setminus(-\delta,\delta)} d\lambda(x)$ *and* $M_1(\delta) := \delta \int_{[-1,1]\setminus(-\delta,\delta)} |x| d\lambda(x)$ *that* $M_0(\delta) \to 0$ *and* $M_1(\delta) \to 0$ *as* $\delta \downarrow 0$*. If, moreover,* $\int_{[-1,1]} |x| d\lambda(x) < \infty$*, then* $\delta \int_{[-1,1]\setminus(-\delta,\delta)} d\lambda(x) \to 0$ *as* $\delta \downarrow 0$*.*

Proof. Indeed let μ be the finite measure on $([-1,1], \mathcal{B}_{[-1,1]})$ given by $\mu(A) := \int_A x^2 d\lambda(x)$ (where A is a Borel subset of $[-1,1]$) and let $f_\delta^0(x) := \left(\frac{\delta}{x}\right)^2 \mathbb{1}_{[-1,1]\setminus(-\delta,\delta)}(x)$ and $f_\delta^1(x) := \frac{\delta}{|x|}\mathbb{1}_{[-1,1]\setminus(-\delta,\delta)}(x)$ be functions on $[-1,1]$. Clearly $0 \le f_\delta^0, f_\delta^1 \le 1$ and $f_\delta^0, f_\delta^1 \to 0$ pointwise as $\delta \downarrow 0$. Hence by the Lebesgue Dominated Convergence Theorem (DCT), we have $M_0(\delta) = \int f_\delta^0 d\mu$ and $M_1(\delta) = \int f^1(\delta) d\mu$ converging to $\int 0 d\mu = 0$ as $\delta \downarrow 0$. The "finite first absolute moment" case is similar. □

Proposition 3.9. *Under scheme 1, with* $\sigma^2 > 0$*, (2.3) holds for all sufficiently small* h*. Notation-wise, under either of the two schemes, we let* $h_\star \in (0, +\infty]$ *be such that* Q^h *has non-negative off-diagonal entries for all* $h \in (0, h_\star)$*.*

Proof. If $V = 0$ this is immediate. If $V = 1$, then (via a triangle inequality):

$$
\begin{aligned}
h|\mu^h| &\le h\left|\sum_{s\in\mathbb{Z}_h\setminus\{0\}} s \int_{A_s^h} \mathbb{1}_{[-1,1]}(y) d\lambda(y)\right| \le h\sum_{s\in\mathbb{Z}_h\setminus\{0\}} \int_{A_s^h} |s - u + u|\mathbb{1}_{[-1,1]}(y) d\lambda(y) \\
&\le h\left(\frac{h}{2}\lambda([-1,1]\setminus[-h/2, h/2]) + \int_{[-1,1]\setminus[h/2,h/2]} |u| d\lambda(u)\right) \to 0
\end{aligned}
$$

as $h \downarrow 0$ by Lemma 3.8. Eventually the expression is smaller than $\sigma^2 > 0$ and the claim follows. □

Furthermore, we have the following inequalities, which together imply an estimate for $|\Psi^h - \Psi|$. In the following, recall the notation ($\delta \ge 0$): $\zeta(\delta) := \delta \int_{[-1,1]\setminus[-\delta,\delta]} |x| d\lambda(x)$, $\gamma(\delta) := \delta^2 \int_{[-1,1]\setminus[-\delta,\delta]} d\lambda(x)$, $c := \lambda(\mathbb{R})$, $b := \kappa(0)$, $d := \lambda(\mathbb{R}\setminus[-1,1])$. Recall also the definition of the sets A_s^h following (2.2).

Proposition 3.10 (Convergence of characteristic exponents)**.** *For all* $p \in \mathbb{R}$*:* $0 \le f_h(p) \le p^4 h^2/4!$ *and* $0 \le i\text{sgn}(p)g_h(p) \le h^2|p|^3/3!$ *(resp., under scheme 2,* $|g_h(p)| \le hp^2/2!$*). Moreover:*

(i) *when $c < \infty$; with $V = 0$: $|l_h(p)| \leq c|p|h/2$.*

(ii) *when $b < \infty = c$; with $V = 1$; for all $h \leq 2$: $|l_h(p)| \leq \frac{h}{2}\left(|p|d + p^2b\right) + (p^2 + |p|^3 + p^4)o(h)$ (resp. under scheme 2, $|l_h(p)| \leq \frac{h}{2}\left(|p|d + 2p^2b\right) + (p^2 + |p|^3 + p^4)o(h))$ where $o(h)$ depends only on λ.*

(iii) *when $b = \infty$; with $V = 1$; for all $h \leq 2$: $|l_h(p)| \leq p^2\left(\zeta(h/2) + \frac{1}{2}\gamma(h/2)\right) + (|p| + |p|^3 + p^4)O(h)$ (resp. under scheme 2, $|l_h(p)| \leq p^2\left[2\zeta(h/2) + \frac{1}{2}\gamma(h/2)\right] + (|p| + p^2 + |p|^3 + p^4)O(h))$ where again $O(h)$ depends only on λ. Note here that we always have $\gamma \leq \zeta$ and that ζ decays strictly slower than h, as $h \downarrow 0$.*

Remark 3.11.

(i) *We may briefly summarize the essential findings of Proposition 3.10 in Table 2, by noting that the following will have been proved for $p \in \mathbb{R}$ and $h \in (0, h_\star \wedge 2)$:*

$$|\Psi^h(p) - \Psi(p)| \leq f(h)R(|p|) + o(f(h))Q(|p|) \qquad (3.4)$$

where R and Q are polynomials of respective degrees α and β and $f : (0, h_\star \wedge 2) \to (0, \infty)$.

Table 2: Summary of Proposition 3.10 via the triplet $(f(h), \alpha, \beta)$ introduced in (i) of Remark 3.11. We agree $\deg 0 = -\infty$, where 0 is the zero polynomial.

$(f(h), \alpha, \beta)$	$\sigma^2 > 0$ (scheme 1)	$\sigma^2 = 0$ (scheme 2)
$\lambda(\mathbb{R}) = 0$ $(V = 0)$	$(h^2, 4, -\infty)$	$(h, 2, -\infty)$
$\lambda(\mathbb{R}) < \infty$ $(V = 0)$	$(h, 1, 4)$	$(h, 2, -\infty)$
$\kappa(0) < \infty = \lambda(\mathbb{R})$ $(V = 1)$	$(h, 2, 4)$	
$\kappa(0) = \infty$ $(V = 1)$	$(\zeta(h/2), 2, 4)$	

(ii) *An analogue of (3.4) is got in the multivariate case, simply by examining directly the difference of (3.3) and (2.1). One does so either component by component (when it comes to the drift and diffusion terms), the estimates being then the same as in the univariate case; or else one employs, in addition, Lemma 3.7 (for the part corresponding to the integral against the Lévy measure). In particular, (3.4) (with $p \in \mathbb{R}^d$) follows for suitable choices of R, Q and f, and Table 2 remains unaffected, apart from its last entry, wherein ζ should be replaced by $\zeta + \chi$ (one must also replace "$\sigma^2 = 0$" (resp. "$\sigma^2 > 0$") by "Σ (resp. non-) degenerate" (amalgamating scheme 1 & 2 into the multivariate one) and $\lambda(\mathbb{R})$ by $\lambda(\mathbb{R}^d)$).*

(iii) *The above entails, in particular, convergence of $\Psi^h(p)$ to $\Psi(p)$ as $h \downarrow 0$ pointwise in $p \in \mathbb{R}$. Lévy's continuity theorem [10, p. 326] and stationarity and independence of increments yield at once Remark 2.2.*

(iv) *Note that we use $V = 1$ rather than $V = 0$ when $b < \infty = c$, because this choice yields linear convergence (locally uniformly) of $\Psi^h \to \Psi$. By contrast, retaining $V = 0$, would have meant that the decay of $\Psi^h - \Psi$ would be governed, modulo terms which are $O(h)$, by the quantity $Q(h) := \sum_{s \in \mathbb{Z}_h \setminus \{0\}} \int_{A_s^h \cap [-1,1]}(s - u)d\lambda(u)$ (as will become clear from the estimates in the proof of Proposition 3.10 below). But the latter can decay slower than h. In particular, consider the family of Lévy measures, indexed by $\epsilon \in [0, 1)$: $\lambda_\epsilon = \sum_{n=1}^\infty w_n\delta_{-x_n}$, with $h_n = 1/3^n$, $x_n = 3h_n/2$, $w_n = 1/x_n^\epsilon$, $n \geq 1$. For all these measures $b < \infty = c$. Furthermore, it is straightforward to verify that $\liminf_{n\to\infty} Q(h_n)/K(h_n) > 0$, where $K(h)$ is $h^{1-\epsilon}$ or $h\log(1/h)$, according as $\epsilon \in (0, 1)$ or $\epsilon = 0$.*

(v) *It is seen from Table 2 that the order of convergence goes from quadratic (at least when $\sigma^2 > 0$) to linear, to sublinear, according as the Lévy measure is zero, $\lambda(\mathbb{R}) > 0$ & $\kappa(0) < \infty$, or κ becomes more and more singular at the origin. Let us attempt to offer some intuition in this respect. First, the quadratic order of convergence is due to the convergence properties of the discrete second and symmetric first derivative. Further, as soon as the Lévy measure is non-zero, the latter is aggregated over the intervals $(A_s^h)_{s \in \mathbb{Z}_h \setminus \{0\}}$, length h, which (at least in the worst case scenario) commit respective errors of order $\lambda(A_s^h)h$ or $\int_{A_s^h}(|x| \wedge 1)d\lambda(x)h$ ($s \in \mathbb{Z}_h \setminus \{0\}$) each, according as $V = 0$ or $V = 1$. Hence, the more singular the κ, the bigger the overall error. Figure 1 depicts this progressive worsening of the convergence rate for the case of α-stable Lévy processes.*
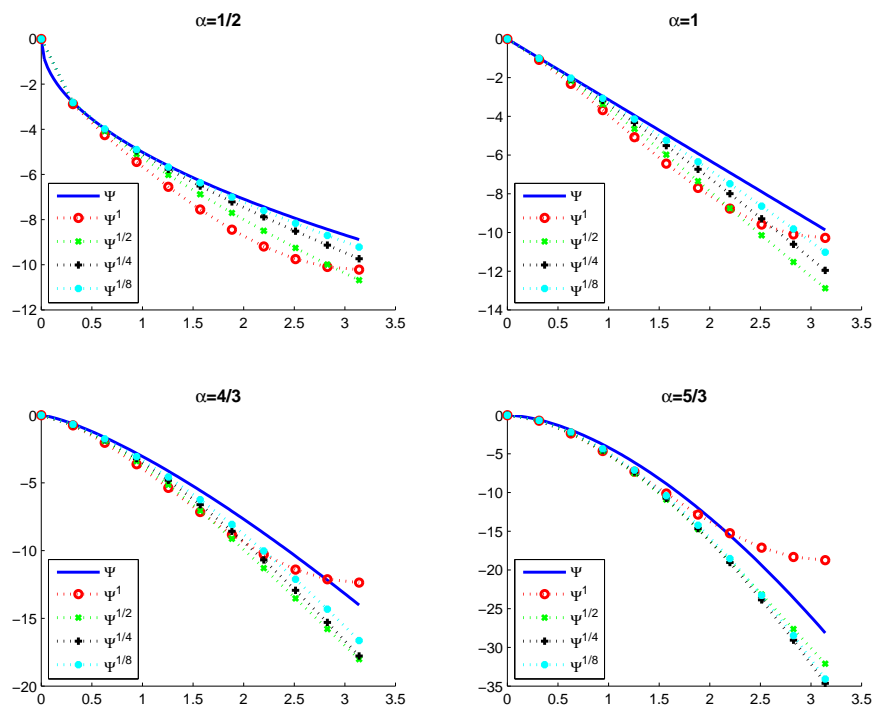

Figure 1: Comparison of the convergence of characteristic exponents for α-stable Lévy processes, $\alpha \in \{1/2, 1, 4/3, 5/3\}$; $\sigma^2 = 0$, $\mu = 0$ and $\lambda(dx) = dx/|x|^{1+\alpha}$ (scheme 2, $V = 1$). Each plot is of Ψ and of Ψ^h ($h \in \{1, 1/2, 1/4, 1/8\}$) on the interval $[0, \pi]$. Note that (i) $\kappa(0) = \infty$, precisely when $\alpha \geq 1$ and (ii) the characteristic exponents are real-valued for the examples shown. The plots are indeed suggestive of a progressive worsening of the rate of convergence as $\alpha \uparrow$.

Proof of Proposition 3.10. The first two assertions are transparent by Lemma 3.6 — with the exception of the estimate under scheme 2, where (with $\delta := hp$):

$$|g_h(p)| = \frac{1}{h}\sqrt{\delta^2 - 2\delta\sin(\delta) + 2(1 - \cos(\delta))} \leq \frac{1}{h}\frac{\delta^2}{2} = hp^2/2!.$$

Further, if $c < \infty$ (under $V = 0$):

$$\left| \sum_{s \in \mathbb{Z}_h \setminus \{0\}} c_s^h (e^{isp} - 1) - \int_{\mathbb{R} \setminus [-\frac{h}{2}, \frac{h}{2}]} (e^{ipu} - 1) d\lambda(u) \right| = \left| \sum_{s \in \mathbb{Z}_h \setminus \{0\}} c_s^h e^{isp} - \int_{\mathbb{R} \setminus [-\frac{h}{2}, -\frac{h}{2}]} e^{ipu} d\lambda(u) \right| =$$

$$= \left| \sum_{s \in \mathbb{Z}_h \setminus \{0\}} \int_{A_s^h} (e^{isp} - e^{ipu}) \, d\lambda(u) \right| \leq \sum_{s \in \mathbb{Z}_h \setminus \{0\}} \int_{A_s^h} \left| 1 - e^{ip(u-s)} \right| d\lambda(u) \leq |p| h \lambda \left(\mathbb{R} \setminus \left[-\frac{h}{2}, \frac{h}{2} \right] \right) / 2,$$

where in the second inequality we apply (1) of Lemma 3.6 and the first follows from the triangle inequalities. Finally, $| \int_{[-h/2, h/2]} (e^{ipu} - 1) d\lambda(u) | \leq \lambda([-h/2, h/2]) |p| h / 2$, again by (1) of Lemma 3.6, and the claim follows.

For the remaining two claims, in addition to recalling the results of Lemma 3.6, we prepare the following specific estimates. Whenever $\{x, y\} \subset \mathbb{R}$, with $\delta := y - x$, $0 \neq |x| \geq |\delta|$, we have:

- using the inequality $\sqrt{1 + z} \leq 1 + z/2$ ($z \geq 0$) and (2) of Lemma 3.6:

$$|e^{ix} - ix - e^{iy} + iy| \leq |\delta x| \left(1 + \frac{1}{2} \left| \frac{\delta}{x} \right| + \frac{1}{8} \frac{\delta^2}{x^2} \right) = |\delta x| + \frac{1}{2} \delta^2 + \frac{1}{8} \left| \frac{\delta^3}{x} \right| \leq |\delta x| + \frac{5}{8} \delta^2. \quad (3.5)$$

- using (3) of Lemma 3.6:

$$|e^{ix} - ix - e^{iy} + iy| \leq |e^{ix} - e^{iy} - ix + iy + x^2/2 - y^2/2| + \frac{1}{2} |x^2 - y^2| \leq \frac{7}{6} |\delta| x^2 + |\delta| |x| + \frac{1}{2} \delta^2. \quad (3.6)$$

Now, when $c = \infty$ (under $V = 1$; for all $h \leq 2$), denoting $\xi(\delta) := \int_{[-\delta, \delta]} x^2 d\lambda(x)$, we have, under scheme 1, as follows:

$$\left| c_0^h \left(\frac{\cos(hp) - 1}{h^2} + \frac{p^2}{2} \right) \right| \leq p^4 h^2 \xi(h/2) / 4!. \quad (3.7)$$

$$\left| \int_{[-\frac{h}{2}, \frac{h}{2}]} u^2 \left(-\frac{p^2}{2} \right) d\lambda(u) - \int_{[-\frac{h}{2}, \frac{h}{2}]} \left(e^{ipu} - 1 - ipu \right) d\lambda(u) \right|$$

$$\leq \left| \int_{[-\frac{h}{2}, \frac{h}{2}]} \left(\cos(pu) - 1 + \frac{p^2 u^2}{2!} \right) d\lambda(u) \right| + \left| \int_{[-\frac{h}{2}, \frac{h}{2}]} (\sin(pu) - pu) d\lambda(u) \right|$$

$$\leq p^4 (h/2)^2 \xi(h/2) / 4! + |p|^3 (h/2) \xi(h/2) / 3!. \quad (3.8)$$

$$| - \mu^h g_h(p) | = \left| -i \mu^h \left(\frac{\sin(hp)}{h} - p \right) \right| \leq \frac{1}{3!} h^2 |p|^3 \left(\zeta(h/2) + \kappa(h/2) \right). \quad (3.9)$$

$$\left| \sum_{s \in \mathbb{Z}_h \setminus \{0\}} c_s^h (e^{isp} - 1) - ip \mu^h - \int_{\mathbb{R} \setminus [-\frac{h}{2}, \frac{h}{2}]} (e^{ipu} - 1 - ipu \mathbb{1}_{[-1,1]}(u)) d\lambda(u) \right|$$

$$\leq \sum_{s \in \mathbb{Z}_h \setminus \{0\}} \int_{A_s^h} \left| e^{ipu} - e^{ips} - ipu \mathbb{1}_{[-1,1]}(u) + ips \mathbb{1}_{[-1,1]}(u) \right| d\lambda(u)$$

$$\leq \sum_{s \in \mathbb{Z}_h \setminus \{0\}} \left[\int_{A_s^h \cap (\mathbb{R} \setminus [-1,1])} + \int_{A_s^h \cap [-1,1]} \right] \left| e^{ipu} - e^{ips} - ipu \mathbb{1}_{[-1,1]}(u) + ips \mathbb{1}_{[-1,1]}(u) \right| d\lambda(u)$$

$$\leq \frac{h}{2} |p| \int_{\mathbb{R} \setminus [-1,1]} d\lambda(u) + p^2 \frac{h}{2} \int_{[-1,1] \setminus [-\frac{h}{2}, \frac{h}{2}]} |u| d\lambda(u) + p^2 \frac{5}{8} \left(\frac{h}{2} \right)^2 \lambda([-1,1] \setminus [-h/2, h/2]), \quad (3.10)$$

where, in particular, we have applied (3.5) to $x = ps$, $y = pu$. If in addition $b = \infty$, we opt rather to use (3.6), again with $x = ps$ and $y = pu$, and obtain instead:

$$\left| \sum_{s \in \mathbb{Z}_h \setminus \{0\}} c_s^h (e^{isp} - 1) - ip \mu^h - \int_{\mathbb{R} \setminus [-\frac{h}{2}, \frac{h}{2}]} (e^{ipu} - 1 - ipu \mathbb{1}_{[-1,1]}(u)) d\lambda(u) \right|$$

$$\leq \frac{h}{2}|p| \int_{\mathbb{R}\setminus[-1,1]} d\lambda(u) + p^2 \frac{h}{2} \int_{[-1,1]\setminus[-\frac{h}{2},\frac{h}{2}]} |u| d\lambda(u) + p^2 \frac{1}{2} \left(\frac{h}{2}\right)^2 \lambda([-1,1]\setminus[-h/2,h/2]) +$$

$$+ \frac{7}{6}|p|^3 \frac{h}{2} \int_{[-1,1]} x^2 d\lambda(x). \tag{3.11}$$

Under scheme 2, (3.7), (3.8) and (3.10)/(3.11) remain unchanged, whereas (3.9) reads:

$$\left|\mu^h g_h(p)\right| \leq \frac{h}{2} p^2 \left(\zeta(h/2) + \kappa(h/2)\right). \tag{3.12}$$

Now, combining (3.7), (3.8), (3.9) and (3.10) under scheme 1 (resp. (3.7), (3.8), (3.12) and (3.10) under scheme 2), yields the desired inequalities when $b < \infty$. If $b = \infty$ use (3.11) in place of (3.10). □

4 Rates of convergence for transition kernels

Finally let us incorporate the estimates of Proposition 3.10 into an estimate of $D_{t,T}^h(x,y)$ (recall the notation in (2.5)). Assumption 2.3 and Table 1 are understood as being in effect throughout this section from this point onwards. Recall that $|\Psi^h - \Psi| \leq \sigma^2|f_h| + \mu|g_h| + |l_h|$ and that the approximation is considered for $h \in (0, h_\star)$ (cf. Proposition 3.9).

First, the following observation, which is a consequence of the h-uniform growth of $-\Re\Psi^h(p)$ as $|p| \to \infty$, will be crucial to our endeavour (compare Remark 3.1).

Proposition 4.1. *For some $\{P, C, \epsilon\} \subset (0,\infty)$ and $h_0 \in (0, h_\star]$, depending only on $\{\lambda, \sigma^2\}$, and then all $h \in (0, h_0)$, $p \in [-\pi/h, \pi/h]\setminus(-P, P)$ and $t \geq 0$: $|\phi_{X_t}^h(p)| \leq \exp\{-C|p|^\epsilon\}$. Moreover, when $\sigma^2 > 0$, we may take $\epsilon = 2$, $P = 0$, $C = \frac{1}{2}\left(\frac{2}{\pi}\right)^2$ and $h_0 = h_\star$, whereas otherwise ϵ may take the same value as in Orey's condition (cf. Assumption 2.3).*

Proof. Assume first $\sigma^2 > 0$, so that we are working under scheme 1. It is then clear from (3.1) that:

$$-\Re\Psi^h(p) \geq \sigma^2 \frac{1 - \cos(hp)}{h^2} \geq \frac{1}{2}\left(\frac{2}{\pi}\right)^2 \sigma^2 p^2,$$

since $1 - \cos(x) = 2\sin^2(x/2) \geq 2\left(\frac{x}{\pi}\right)^2$ for all $x \in [-\pi,\pi]$. On the other hand, if $\sigma^2 = 0$, we work under scheme 2 and necessarily $V = 1$. In that case it follows from (3.2) for $h \leq 2$ and $p \in [-\pi/h, \pi/h]\setminus\{0\}$, that:

$$\begin{aligned}
-\Re\Psi^h(p) &\geq \left[c_0^h \frac{1 - \cos(hp)}{h^2} + \sum_{s \in \mathbb{Z}_h\setminus\{0\}} c_s^h (1 - \cos(sp))\right] \\
&\geq \frac{2}{\pi^2} p^2 \left(\int_{A_0^h} u^2 d\lambda(u) + \sum_{s \in \mathbb{Z}_h\setminus\{0\}, |s|\leq\frac{\pi}{|p|}} s^2 c_s^h\right) \\
&\geq \frac{2}{\pi^2} p^2 \left(\int_{A_0^h} u^2 d\lambda(u) + \frac{4}{9} \sum_{s \in \mathbb{Z}_h\setminus\{0\}, |s|\leq\frac{\pi}{|p|}} \int_{A_s^h} u^2 d\lambda(u)\right) \\
&\geq \frac{2}{\pi^2} p^2 \left(\int_{A_0^h} u^2 d\lambda(u) + \frac{4}{9} \int_{[-(\frac{\pi}{|p|}-\frac{h}{2}), \frac{\pi}{|p|}-\frac{h}{2}]\setminus A_0^h} u^2 d\lambda(u)\right) \\
&\geq \frac{8}{9} \frac{1}{\pi^2} p^2 \int_{[-((\frac{\pi}{|p|}-\frac{h}{2})\vee\frac{h}{2}), ((\frac{\pi}{|p|}-\frac{h}{2})\vee\frac{h}{2})]} u^2 d\lambda(u)
\end{aligned}$$

$$\geq \quad \frac{8}{9}\frac{1}{\pi^2}p^2\int_{[-\frac{1}{2}\frac{\pi}{|p|},\frac{1}{2}\frac{\pi}{|p|}]}u^2 d\lambda(u).$$

Now invoke Assumption 2.3. There are some $\{r_0, A_0\} \in (0, +\infty)$ such that for all $r \in (0, r_0]$: $\int_{[-r,r]}u^2d\lambda(u) \geq A_0 r^{2-\epsilon}$. Thus for $P = \pi/(2r_0)$ and then all $p \in \mathbb{R}\backslash(-P, P)$, we obtain:

$$\int_{[-\frac{1}{2}\frac{\pi}{|p|},\frac{1}{2}\frac{\pi}{|p|}]}u^2d\lambda(u) \geq A_0\left(\frac{1}{2}\frac{\pi}{|p|}\right)^{2-\epsilon},$$

from which the desired conclusion follows at once. Remark that, possibly, r_0 may be taken as $+\infty$, in which case P may be taken as zero. $\qquad\square$

Second, we have the following general observation which concerns the transfer of the rate of convergence from the characteristic exponents to the transition kernels. Its validity is in fact independent of Assumption 2.3.

Proposition 4.2. *Suppose there are $\{P, C, \epsilon\} \subset (0, \infty)$, a real-valued polynomial R, an $h_0 \in (0, h_\star]$, and a function $f : (0, h_0) \to (0, \infty)$, decaying to 0 no faster than some power law, such that for all $h \in (0, h_0)$:*

1. *for all $p \in [-\pi/h, \pi/h]$: $|\Psi^h(p) - \Psi(p)| \leq f(h)R(|p|)$.*

2. *for all $s > 0$ and $p \in [-\pi/h, \pi/h]\backslash(-P, P)$: $|\phi_{X_s^h}(p)| \leq \exp\{-Cs|p|^\epsilon\}$; whereas for $p \in \mathbb{R}\backslash(-P, P)$: $|\phi_{X_s}(p)| \leq \exp\{-Cs|p|^\epsilon\}$.*

Then for any $s > 0$, $\Delta_s(h) = O(f(h))$.

Before proceeding to the proof of this proposition, we note explicitly the following elementary, but key lemma:

Lemma 4.3. *For $\{z, v\} \subset \mathbb{C}$: $|e^z - e^v| \leq (|e^z| \vee |e^v|)|z - v|$.*

Proof. This follows from the inequality $|e^z - 1| \leq |z|$ for $\Re z \leq 0$, whose validity may be seen by direct estimation. $\qquad\square$

Proof of Proposition 4.2. From (2.6) and (2.7) we have for the quantity $\Delta_s(h)$ from (2.5):

$$\Delta_s(h) \leq \int_{\mathbb{R}\backslash(-\pi/h,\pi/h)}|\exp\{\Psi(p)s\}|dp + \int_{[-\pi/h,\pi/h]}|\exp\{\Psi^h(p)s\} - \exp\{\Psi(p)s\}|dp.$$

Then the first term decays faster than any power law in h by (2) and L'Hôpital's rule, say, while in the second term we use the estimate of Lemma 4.3. Since $\exp\{-Ct|p|^\epsilon\}dp$ integrates every polynomial in $|p|$ absolutely, by (1) and (2) integration in the second term can then be extended to the whole of $\mathbb{R}$ and the claim follows. $\qquad\square$

Proposition 4.2 allows us to transfer the rates of convergence directly from those of the characteristic exponents to the transition kernels. In particular, we have, immediately, the following proof of the multivariate result (which we state before the univariate case is dealt with in full detail):

Proof of Theorem 2.8. The conclusions of Theorem 2.8 follow from a straightforward extension (of the proof) of Proposition 4.2 to the multivariate setting, (ii) of Remark 3.11, Assumption 2.6 and Remark 2.7. $\qquad\square$

Returning to the univariate case, analogous conclusions could be got from Remark 3.1, Proposition 4.1 (themselves both consequences of Assumption 2.3) and Proposition 3.10. In the sequel, however, in the case when $\sigma^2 > 0$, we shall be interested in a more precise estimate of the constant in front of the leading order term (D_1 in the statement of Theorem 2.6). Moreover, we shall want to show our estimates are tight in an appropriate precise sense.

To this end we assume given a function K with the properties that:

(F) $0 \le K(h) \to \infty$ as $h \downarrow 0$ and $K(h) \le \frac{\pi}{h}$ for all sufficiently small h;

(E) the quantity

$$\mathcal{A}(h) := \left[\int_{-\infty}^{-K(h)} + \int_{K(h)}^{\infty} \right] |\exp\{\Psi(p)s\}| \, dp + \left[\int_{-\frac{\pi}{h}}^{-K(h)} + \int_{K(h)}^{\frac{\pi}{h}} \right] |\exp\{\Psi^h(p)s\}| \, dp$$

decays faster than the leading order term in the estimate of $D_{t,T}^h(x,y)$ (for which see, e.g., Table 2);

(C) $\sup_{[-K(h),K(h)]} |\Psi^h - \Psi| \le 1$ for all small enough h

(suitable choices of K will be identified later, cf. Table 3 on p. 20). We now comment on the reasons behind these choices.

First, the constants $\{C, P, \epsilon\}$ are taken so as to satisfy simultaneously Remark 3.1 and Proposition 4.1. In particular, if $\sigma^2 > 0$, we take $\epsilon = 2$, $P = 0$, $C = \frac{1}{2}\sigma^2$, and if $\sigma^2 = 0$, we may take ϵ from Orey's condition (cf. Assumption 2.3).

Next, we divide the integration regions in (2.6) and (2.7) into five parts (cf. property (F)): $(-\infty, -\frac{\pi}{h}]$, $(-\frac{\pi}{h}, -K(h))$, $[-K(h), K(h)]$, $(K(h), \frac{\pi}{h})$, $[\frac{\pi}{h}, \infty)$. Then we separate (via a triangle inequality) the integrals in the difference $D_{t,T}^h(x,y)$ accordingly and use the triangle inequality in the second and fourth region, thus (with $s := T - t > 0$):

$$
\begin{aligned}
2\pi D_{t,T}^h(x,y) \quad \le \quad & \left[\int_{-\infty}^{-\pi/h} + \int_{\pi/h}^{\infty} \right] |\exp\{\Psi(p)s\}| \, dp + \\
& \left[\int_{-\frac{\pi}{h}}^{-K(h)} + \int_{K(h)}^{\frac{\pi}{h}} \right] \left(\left|\exp\{\Psi^h(p)s\}\right| + |\exp\{\Psi(p)s\}| \right) dp + \\
& \int_{-K(h)}^{K(h)} \left|\exp\{\Psi(p)s\} - \exp\{\Psi^h(p)s\}\right| dp.
\end{aligned}
$$

Finally, we gather the terms with $|\exp\{\Psi(p)s\}|$ in the integrand and use $|e^z - 1| \le e^{|z|} - 1$ ($z \in \mathbb{C}$) to estimate the integral over $[-K(h), K(h)]$, so as to arrive at:

$$2\pi D_{t,T}^h(x,y) \le \mathcal{A}(h) + \int_{-K(h)}^{K(h)} |\exp\{\Psi(p)s\}| \left(\exp\left\{ s \left|\Psi^h(p) - \Psi(p)\right| \right\} - 1 \right) dp. \qquad (4.1)$$

Now, the rate of decay of $\mathcal{A}(h)$ can be controlled by choosing $K(h)$ converging to $+\infty$ fast enough, viz. property (E). On the other hand, in order to control the second term on the right-hand side of the inequality in (4.1), we choose $K(h)$ converging to $+\infty$ slowly enough so as to guarantee (C). Table 3 lists suitable choices of $K(h)$. It is easily checked from Table 2 (resp. using L'Hôpital's rule coupled with Remark 3.1 and Proposition 4.1), that these choices of $K(h)$ do indeed satisfy (C) (resp. (E)) above. Property (F) is straightforward to verify.

Further, due to (C), for all sufficiently small h, everywhere on $[-K(h), K(h)]$:

$$e^{s|\Psi^h - \Psi|} - 1 = s|\Psi^h - \Psi| + \sum_{k=2}^{\infty} \frac{(s|\Psi^h - \Psi|)^k}{k!} \le s|\Psi^h - \Psi| + (s|\Psi^h - \Psi|)^2 e^{s|\Psi^h - \Psi|} \le s|\Psi^h - \Psi| + e(s|\Psi^h - \Psi|)^2.$$

Manifestly the second term will always decay strictly faster than the first (so long as they are not 0). Moreover, since $\exp\{-Cs|p|^\epsilon\} dp$ integrates every polynomial in $|p|$ (cf. the findings of Proposition 3.10) absolutely, it will therefore be sufficient in the sequel to estimate (cf. (4.1)):

$$\frac{s}{2\pi} \int_{\mathbb{R}} \exp\{-Cs|p|^\epsilon\} \left|\Psi^h(p) - \Psi(p)\right| dp. \qquad (4.2)$$

On the other hand, for the purposes of establishing sharpness of the rates for the quantity $D_{t,T}^h(x,y)$, we make the following:

Table 3: Suitable choices of $K(h)$. For example, the $\sigma^2 > 0$ and $\lambda(\mathbb{R}) = 0$ entry indicates that we choose $K(h) = \sqrt{\frac{2}{Cs} \log \frac{1}{h}}$ and then $\mathcal{A}(h)$ is of order $o(h^2)$.

	$\sigma^2 > 0$ (scheme 1)	$\sigma^2 = 0$ (scheme 2)
$\lambda(\mathbb{R}) = 0$ $(V = 0)$	$K(h) = \sqrt{\frac{2}{Cs} \log \frac{1}{h}} \to \mathcal{A}(h) = o(h^2)$	$\times$
$\lambda(\mathbb{R}) < \infty$ $(V = 0)$	$K(h) = \sqrt{\frac{1}{Cs} \log \frac{1}{h}} \to \mathcal{A}(h) = o(h)$	$\times$
$\kappa(0) < \infty = \lambda(\mathbb{R})$ $(V = 1)$	$K(h) = \sqrt{\frac{1}{Cs} \log \frac{1}{h}} \to \mathcal{A}(h) = o(h)$	$K(h) = \sqrt[c]{\frac{2}{Cs} \log \frac{1}{h}} \to \mathcal{A}(h) = o(h)$
$\kappa(0) = \infty$ $(V = 1)$	$K(h) = \left(\frac{1}{\zeta(h/2)}\right)^{1/4} \to \mathcal{A}(h) = o(\zeta(h/2))$	

Remark 4.4 (RD). *Suppose we seek to prove that $f \geq 0$ converges to 0 no faster than $g > 0$, i.e. that $\limsup_{h\downarrow 0} f(h)/g(h) \geq C > 0$ for some C. If one can show $f(h) \geq A(h) - B(h)$ and $B = o(g)$, then to show $\limsup_{h\downarrow 0} f(h)/g(h) \geq C$, it is sufficient to establish $\limsup_{h\downarrow 0} A(h)/g(h) \geq C$. We refer to this extremely useful principle as* reduction by domination *(henceforth RD).*

In particular, it follows from our above discussion, that it will be sufficient to consider (we shall always choose $x = y = 0$):

$$\frac{s}{2\pi} \int_{-K(h)}^{K(h)} e^{s\Psi(p)} \left(\Psi^h(p) - \Psi(p)\right) dp, \tag{4.3}$$

i.e. in Remark 4.4 this is A, and the difference to $D_{t,T}(0,0)$ represents B. Moreover, we can further replace $\Psi^h(p) - \Psi(p)$ in the integrand of (4.3) by any expression whose difference to $\Psi^h(p) - \Psi(p)$ decays, upon integration, faster than the leading order term. For the latter reductions we (shall) refer to the proof of Proposition 3.10.

We have now brought the general discussion as far as we could. The rest of the analysis must invariably deal with each of the particular instances separately and we do so in the following two propositions. Notation-wise we let DCT stand for the Lebesgue Dominated Convergence Theorem.

Proposition 4.5 (Convergence of transition kernels — $\sigma^2 > 0$). *Suppose $\sigma^2 > 0$. Then for any $s = T - t > 0$:*

1. *If $\lambda(\mathbb{R}) = 0$:*

$$\Delta_s(h) \leq h^2 \left[\frac{1}{3\pi} \frac{|\mu|}{\sigma^4 s} + \frac{1}{8\sqrt{2\pi}} \frac{1}{(s\sigma^2)^{3/2}}\right] + o(h^2).$$

Moreover, with $\sigma^2 s = 1$ and $\mu = 0$ we have $\limsup_{h\downarrow 0} D_{t,T}^h(0,0)/h^2 \geq 1/(8\sqrt{2\pi})$, proving that in general the convergence rate is no better than quadratic.

2. *If $0 < \lambda(\mathbb{R}) < \infty$:*

$$\Delta_s(h) \leq h \frac{1}{2\pi} \frac{c}{\sigma^2} + o(h).$$

Moreover, with $\sigma^2 = s = 1$, $\mu = 0$ and $\lambda = \frac{1}{2}(\delta_{1/2} + \delta_{-1/2})$, $\limsup_{h\downarrow 0} D_{t,T}^h(0,0)/h > 0$, showing that convergence in general is indeed no better than linear.

3. *If $\kappa(0) < \infty = \lambda(\mathbb{R})$:*

$$\Delta_s(h) \leq h \left[\frac{1}{2\pi} \frac{d}{\sigma^2} + \frac{1}{2\sqrt{2\pi}} \frac{bs}{(\sigma^2 s)^{3/2}}\right] + o(h).$$

Moreover, with $\sigma^2 = s = 1$, $\mu = 0$ and $\lambda = \frac{1}{2}(\delta_{3/2} + \delta_{-3/2}) + \frac{1}{2}\sum_{k=1}^{\infty}(\delta_{1/3^k} + \delta_{-1/3^k})$, one has $\limsup_{h\downarrow 0} D_{t,T}^h(0,0)/h > 0$.

4. If $\kappa(0) = \infty$:

$$\Delta_s(h) \leq \frac{1}{\sqrt{2\pi}} \frac{s}{(\sigma^2 s)^{3/2}} \left(\zeta(h/2) + \frac{1}{2}\gamma(h/2) \right) + o(\zeta(h/2)).$$

Moreover, with $\sigma^2 = s = 1$, $\mu = 0$, and $\lambda = \sum_{k=1}^{\infty} w_k(\delta_{x_k} + \delta_{-x_k})$, where $x_n = \frac{3}{2}\frac{1}{3^n}$ and $w_n = 1/x_n$ ($n \in \mathbb{N}$), one has $\limsup_{h\downarrow 0} D_{t,T}^h(0,0)/\zeta(h/2) > 0$.

Proof. Estimates of $\Delta_s(h)$ follow at once from (4.2) and Proposition 3.10, simply by integration. As regards establishing sharpness of the estimates, however, we have as follows (recall that we always take $x = y = 0$):

1. $\lambda(\mathbb{R}) = 0$. Using (4.3) it is sufficient to consider:

$$A(h) := \frac{1}{2\pi} \left| \int_{-K(h)}^{K(h)} \exp\left\{ -\frac{1}{2}p^2 \right\} f_h(p) dp \right|.$$

By DCT, we have $A(h)/h^2 \to \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}p^2\}p^4/4! dp$ and the claim follows.

2. $0 < \lambda(\mathbb{R}) < \infty$. Using (4.3) and further RD via the estimates in the proof of Proposition 3.10, we conclude that it is sufficient to observe for the sequence $(h_n = \frac{1}{3^n})_{n\geq 1} \downarrow 0$ that:

$$\limsup_{n\to\infty} \frac{1}{2\pi h_n} \left| \int_{-K(h_n)}^{K(h_n)} \exp\left\{ -\frac{1}{2}p^2 - 1 + \cos(p/2) \right\} l_{h_n}(p) dp \right| > 0.$$

It is also clear that we may express:

$$l_{h_n}(p) = 2\frac{1}{2}\Re\left(e^{ip(1/2 - h_n/2)} - e^{ip/2} \right) = \cos(p/2)(\cos(ph_n/2) - 1) + \sin(p/2)\sin(ph_n/2).$$

Therefore, by further RD, it will be sufficient to consider:

$$\limsup_{n\to\infty} \frac{1}{2\pi h_n} \left| \int_{-K(h_n)}^{K(h_n)} \exp\left\{ -\frac{1}{2}p^2 - 1 + \cos(p/2) \right\} \sin(ph_n/2)\sin(p/2) dp \right|.$$

By DCT it is equal to:

$$I := \frac{1}{2\pi} \int_0^{\infty} p\sin(p/2) \exp\{-\frac{1}{2}p^2 - 1 + \cos(p/2)\} dp.$$

The numerical value of this integral is (to one decimal place in units of 2π) $0.4/(2\pi)$, but we can show that $I > 0$ analytically. Indeed the integrand is positive on $[0,6]$. Hence $2\pi e I \geq \sin(1/2)e^{\cos(3/2)}\int_1^3 pe^{-p^2/2}dp - e\int_6^{\infty} pe^{-p^2/2}dp = \sin(1/2)e^{\cos(3/2)}[e^{-1/2} - e^{-9/2}] - e^{-17}$. Now use $\sin(1/2) \geq (1/2)\cdot(2/\pi)$ (which follows from the concavity of $\sin$ on $[0,\pi/2]$), so that, very crudely: $2\pi e I \geq (1/\pi)e^{-1/2}(1 - e^{-4}) - e^{-17} \geq (1/\pi)e^{-1/2}(1/2) - e^{-17} \geq (1/e^2)e^{-1/2}(1/e) - e^{-17} \geq e^{-4} - e^{-17} > 0$.

3. $\kappa(0) < \infty = \lambda(\mathbb{R})$. Let $h_n = 1/3^n$, $n \geq 1$. Because the second term in λ lives on $\cup_{n\in\mathbb{N}}\mathbb{Z}_{h_n}$, it is seen quickly (via RD) that one need only consider (to within non-zero multiplicative constants):

$$\limsup_{n\to\infty} \int_{-K(h_n)}^{K(h_n)} \frac{1}{h_n} \sin(ph_n/2)\sin(3p/2) \exp\left\{ -\frac{1}{2}p^2 + (\cos(3p/2) - 1) + \sum_{k=1}^{\infty}(\cos(p/3^k) - 1) \right\} dp.$$

By DCT it is sufficient to observe that:

$$\int_0^{2\pi/3} \sin(3p/2)p \exp\left\{ -\frac{1}{2}p^2 + (\cos(3p/2) - 1) - \frac{p^2}{2}\sum_{k=1}^{\infty}\frac{1}{9^k} \right\} dp - \int_{2\pi/3}^{\infty} p\exp\left\{ -\frac{1}{2}p^2 \right\} dp > 0.$$

To see the latter, note that the second integral is immediate and equals: $e^{-(2\pi/3)^2/2}$. As for the first one, make the change of variables $u = 3p/2$. Thus we need to establish that:

$$A := (4/(9e)) \int_0^\pi \sin(u)u \exp\{-u^2/4 + \cos(u)\}du - e^{-(2\pi/3)^2/2} > 0.$$

Next note that $-u^2/4 + \cos(u)$ is decreasing on $[0, \pi]$ and the integrand in A is positive. It follows that:

$$A \geq \frac{4}{9e} \int_0^{\pi/3} u \sin(u) \exp\left\{ -\frac{1}{4}\left(\frac{\pi}{3}\right)^2 + \cos\left(\frac{\pi}{3}\right) \right\} du +$$
$$\frac{4}{9e} \int_{\pi/3}^{\pi/2} u \sin(u) \exp\left\{ -\frac{1}{4}\left(\frac{\pi}{2}\right)^2 + \cos\left(\frac{\pi}{2}\right) \right\} du - e^{-2\pi^2/9}.$$

Using integration by parts, it is now clear that this expression is algebraic over the rationals in e, $\sqrt{3}$ and the values of the exponential function at rational multiples of π^2. Since this explicit expression can be estimated from below by a positive quantity, one can check that $A > 0$.

4. $\kappa(0) = \infty$. Let again $h_n = 1/3^n$, $n \geq 1$. Notice that:

$$\sigma_1 := \int_{[-1,1]\setminus[-h_n/2, h_n/2]} u^2 d\lambda(u) = 2\sum_{k=1}^n x_k^2 w_k, \text{ and } \sigma_2 := \sum_{s\in\mathbb{Z}_{h_n}\setminus\{0\}} c_s^{h_n} s^2 = 2\sum_{k=1}^n \left(x_k - \frac{h_n}{2}\right)^2 w_k,$$

so that $\Delta := \sigma_1 - \sigma_2 = 2\zeta(h_n/2) - \gamma(h_n/2) \geq \zeta(h_n/2)$. Using (3) of Lemma 3.6 in the estimates of Proposition 3.10, it is then not too difficult to see that it is sufficient to show $\int_{-K(h_n)}^{K(h_n)} p^2 \exp\{\Psi(p)\}dp$ converges to a strictly positive value as $n \to \infty$, which is transparent (since Ψ is real valued).

$\square$

Proposition 4.6 (Convergence of transition kernels — $\sigma^2 = 0$). *Suppose $\sigma^2 = 0$. For any $s = T - t > 0$:*

1. *If Orey's condition is satisfied and $\kappa(0) < \infty = \lambda(\mathbb{R})$, then $\Delta_s(h) = O(h)$. Moreover, with $\sigma^2 = 0$, $s = 1$, $\mu = 0$ and $\lambda = \frac{1}{2}\sum_{k=1}^\infty w_k(\delta_{x_k} + \delta_{-x_k})$, where $x_n = \frac{3}{2}\frac{1}{3^n}$ and $w_n = 1/\sqrt{x_n}$ ($n \in \mathbb{N}$), Orey's condition holds with $\epsilon = 1/2$ and one has $\limsup_{h\downarrow 0} D_{t,T}^h(0,0)/h > 0$.*

2. *If Orey's condition is satisfied and $\kappa(0) = \infty$, then $\Delta_s(h) = O(\zeta(h/2))$. Moreover, with $\sigma^2 = 0$, $s = 1$, $\mu = 0$, and $\lambda = \sum_{k=1}^\infty w_k(\delta_{x_k} + \delta_{-x_k})$, where $x_n = \frac{3}{2}\frac{1}{3^n}$ and $w_n = 1/x_n$ ($n \in \mathbb{N}$), Orey's condition holds with $\epsilon = 1$ and one has $\limsup_{h\downarrow 0} D_{t,T}^h(0,0)/\zeta(h/2) > 0$.*

Proof. Again the rates of convergence for $\Delta_s(h)$ follow at once from (4.2) and Proposition 3.10 (or, indeed, from Proposition 4.2). As regards sharpness of these rates, we have (recall that we take $x = y = 0$):

1. $\kappa(0) < \infty = \lambda(\mathbb{R})$. Let $h_n = 1/3^n$, $n \geq 1$. By checking Orey's condition on the decreasing sequence $(h_n)_{n\geq 1}$, Assumption 2.3 is satisfied with $\epsilon = 1/2$ and we have $b < \infty = c$. $\mu^h = 0$ by symmetry. Moreover by (4.3), and by further going through the estimates of Proposition 3.10 using RD, it suffices to show:

$$\limsup_{n\to\infty} \frac{1}{h_n} \left| \int_{-K(h_n)}^{K(h_n)} \exp\{\Psi(p)\} \left(\sum_{s\in\mathbb{Z}_{h_n}\setminus\{0\}} \int_{A_s^{h_n}} (\cos(ps) - \cos(pu))\, d\lambda(u) \right) dp \right| > 0.$$

Now, one can write for $s \in \mathbb{Z}_{h_n} \backslash \{0\}$ and $u \in A_s^{h_n}$,

$$\cos(sp) - \cos(pu) = \cos(pu)(\cos((s-u)p)-1) - \sin(pu)(\sin((s-u)p)-(s-u)p) - \sin(pu)(s-u)p$$

and via RD get rid of the first two terms (i.e. they contribute to B rather than A in Remark 4.4). It follows that it is sufficient to observe:

$$\limsup_{n \to \infty} \frac{1}{h_n} \left| \int_{-K(h_n)}^{K(h_n)} \exp\left\{ \sum_{k=1}^{\infty} (\cos(px_k)-1)w_k \right\} \left(\sum_{k=1}^{n} w_k \sin(px_k) \right) h_n p\, dp \right| > 0.$$

Finally, via DCT and evenness of the integrand, we need only have:

$$\int_0^{\infty} \left(\sum_{k=1}^{\infty} w_k \sin(px_k) \right) p \exp\left\{ \sum_{k=1}^{\infty} (\cos(px_k)-1)w_k \right\} dp \neq 0.$$

One can in fact check that the integrand is strictly positive, as Lemma 4.7 shows, and thus the proof is complete.

2. $\kappa(0) = \infty$. The example here works for the same reasons as it did in (4) of the proof of Proposition 4.5 (but here benefiting explicitly also from $\mu^h = 0$). We only remark that Orey's condition is of course fulfilled with $\epsilon = 1$, by checking it on the decreasing sequence $(h_n)_{n \geq 1}$.

$\square$

Lemma 4.7. Let $\psi(p) := \sum_{k=1}^{\infty} 3^{k/2} \sin(\frac{3}{2}p/3^k)$. Then ψ is strictly positive on $(0, \infty)$.

Proof. We observe that, (i) $\psi|_{(0, \frac{\pi}{2}]} > 0$ and (ii) for $p \in (\pi/2, 3\pi/2]$ we have: $\psi(p) > \sqrt{3}/(\sqrt{3}-1) =: A_0$. Indeed, (i) is trivial since for $p \in (0, \pi/2]$, $\psi(p)$ is defined as a sum of strictly positive terms. We verify (ii) by observing that (ii.1) $\psi(\pi/2) > A_0$ and (ii.2) ψ is nondecreasing on $[\pi/2, 3\pi/2]$. Both these claims are tedious but elementary to verify by hand. Indeed, with respect to (ii.1), summing three terms of the series defining $\psi(\pi/2)$ is sufficient. Specifically we have $\psi(\pi/2) > \sqrt{3}\sin(\pi/4) + 3\sin(\pi/12) + 3\sqrt{3}\sin(\pi/36)$ and we estimate $\sin(\pi/36) \geq \frac{\pi}{36}\sin(\pi/3)/(\pi/3)$. With respect to (ii.2) we may differentiate under the summation sign, and then $\psi'(p) \geq \frac{\sqrt{3}}{2}\cos(3\pi/4) + \frac{1}{2}\cos(\pi/4) + \frac{\sqrt{3}}{6}\cos(\pi/12)$. The final details of the calculations are left to the reader.

Finally, we claim that if for some $B > 0$ we have $\psi|_{(0,B]} > 0$ and $\psi|_{(B,3B]} > A_0$, then $\psi|_{(0,3B]} > 0$ and $\psi|_{(3B,9B]} > A_0$, and hence the assertion of the lemma will follow at once (by applying the latter first to $B = \pi/2$, then $B = 3\pi/2$ and so on). So let $3p \in (3B, 9B]$, i.e. $p \in (B, 3B]$. Then $\psi(3p) = \sqrt{3}(\sin(3p/2) + \psi(p)) > \sqrt{3}(-1 + A_0) = A_0$, as required. $\square$

Finally, note that:
Proof of Theorem 2.4. The conclusions of Theorem 2.4 follow from Propositions 4.5 and 4.6. $\square$

5 Convergence of expectations and algorithm

5.1 Convergence of expectations

For the sake of generality we state the results in the multivariate setting, but only do so, when this is not too burdensome on the brevity of exposition. For $d = 1$, either the multivariate or the univariate schemes may be considered.

Let $f : \mathbb{R}^d \to \mathbb{R}$ be bounded Borel measurable and define for $t \geq 0$ and $h \in (0, h_\star)$: $p_t := p_{0,t}$ and $P_t^h := P_{0,t}^h$, whereas for $x \in \mathbb{Z}_h^d$, we let $p_t(x) := p_t(0, x)$ and $P_t^h(x) =$

$P_t^h(0, x)$ (assuming the continuous densities exist). Note that for $t \geq 0$, and then for $x \in \mathbb{R}^d$,

$$\mathsf{E}^x[f \circ X_t] = \int_\mathbb{R} f(y)p_t(x, y)dy, \tag{5.1}$$

whereas for $x \in \mathbb{Z}_h^d$ and $h \in (0, h_\star)$:

$$\mathsf{E}^x[f \circ X_t^h] = \sum_{y \in \mathbb{Z}_h^d} f(y)P_t^h(x, y). \tag{5.2}$$

Moreover, if f is continuous, we know that, as $h \downarrow 0$, $\mathsf{E}^x[f \circ X_t^h] \to \mathsf{E}^x[f \circ X_t]$, since $X_t^h \to X_t$ in distribution. Next, under additional assumptions on the function f, we are able to establish the rate of this convergence and how it relates to the convergence rate of the transition kernels, to wit:

Proposition 5.1. *Assume* (2.9) *of Assumption 2.6. Let* $h_0 \in (0, \infty)$, $g : (0, h_0) \to (0, \infty)$ *and* $t > 0$ *be such that* $\Delta_t = O(g)$. *Suppose furthermore that the following two conditions on* f *are satisfied:*

(i) f *is (piecewise[1], if* $d = 1$*) Lipschitz continuous.*

(ii) $\sup_{h \in (0, h_0)} h^d \sum_{x \in \mathbb{Z}_h^d} |f(x)| < \infty$.

Then:

$$\sup_{x \in \mathbb{Z}_h^d} |\mathsf{E}^x[f \circ X_t] - \mathsf{E}^x[f \circ X_t^h]| = O(h \vee g(h)).$$

Remark 5.2.

(1) *Condition (ii) is fulfilled in the univariate case* $d = 1$, *if, e.g.:* $f \in L^1(\mathbb{R})$, *w.r.t. Lebesgue measure,* f *is locally bounded and for some* $K \in [0, \infty)$, $|f|\|_{(-\infty, -K]}$ *(restriction of* $|f|$ *to* $(-\infty, -K]$*) is nondecreasing, whereas* $|f|\|_{[K, \infty)}$ *is nonincreasing.*

(2) *The rate of convergence of the expectations is thus got by combining the above proposition with the findings of Theorems 2.4 and 2.8.*

(3) *Note also that the convergence rate in Proposition 5.1 is never established as better than linear, albeit the transition kernels may converge faster, e.g. at a quadratic rate in the case of Brownian motion with drift. This is so, because we are not only approximating the density with the normalized probability mass function, but also the integral is substituted by a sum (cf. (5.1) and (5.2)). One thus has to invariably estimate* $f(y) - f(z)$ *for* $z \in A_y^h$, $y \in \mathbb{Z}_h^d$. *Excluding the trivial case of a constant* f, *however, this estimate can be at best linear in* $|y - z|$ *(α-Hölder continuous functions on* $\mathbb{R}^d$ *with Hölder exponent* $\alpha > 1$ *being, in fact, constant). Moreover, it appears that this problem could not be avoided using Fourier inversion techniques (as opposed to the direct estimate given in the proof below). Indeed, one would then need to estimate, in particular, the difference of the Fourier transforms* $\int_\mathbb{R} e^{-ipy} f(y)dy - h \sum_{y \in \mathbb{Z}_h^d} f(y)e^{-ipy}$, *wherein again an integral is substituted by a discrete sum and a similar issue arises.*

Proof. Decomposing the difference $\mathsf{E}^x[f \circ X_t] - \mathsf{E}^x[f \circ X_t^h]$ via (5.1) and (5.2), we have:

$$\mathsf{E}^x[f \circ X_t] - \mathsf{E}^x[f \circ X_t^h] \quad = \quad \sum_{y \in \mathbb{Z}_h^d} \int_{A_y^h} (f(z) - f(y)) \, p_t(x, z)dz + \tag{5.3}$$

---

[1] In the sense that there exists some natural n, and then disjoint open intervals $(I_i)_{i=1}^n$, whose union is cofinite in $\mathbb{R}$, and such that $f|_{I_i}$ is Lipschitz for each $i \in \{1, \ldots, n\}$.

$$+ \sum_{y \in \mathbb{Z}_h^d} \int_{A_y^h} f(y) \left(p_t(x,z) - p_t(x,y) \right) dz + \qquad (5.4)$$

$$+ \sum_{y \in \mathbb{Z}_h^d} f(y) h^d \left[p_t(x,y) - \frac{1}{h^d} P_t^h(x,y) \right]. \qquad (5.5)$$

Now, (5.5) is of order $O(g(h))$, by condition (ii) and since $\Delta_t = O(g)$. Further, (5.3) is of order $O(h)$ on account of condition (i), and since $\int p_t(x,z)dz = 1$ for any $x \in \mathbb{R}^d$ (to see piecewise Lipschitzianity is sufficient in dimension one ($d = 1$), simply observe $\sup_{\{x,y\} \subset \mathbb{R}} p_t(x,y)$ is finite, as follows immediately from the integral representation of p_t). Finally, note that $p_t(x, \cdot)$ is also Lipschitz continuous (uniformly in $x \in \mathbb{R}^d$), as follows again at once from the integral representation of the transition densities. Thus, (5.4) is also of order $O(h)$, where again we benefit from condition (ii) on the function f. $\qquad \square$

In order to be able to relax condition (ii) of Proposition 5.1, we first establish the following Proposition 5.3, which concerns finiteness of moments of X_t.

In preparation thereof, recall the definition of submultiplicativity of a function $g : \mathbb{R}^d \to [0, \infty)$:

$$g \text{ is submultiplicative} \Leftrightarrow \exists a \in (0, \infty) \text{ such that } g(x+y) \leq ag(x)g(y), \text{ whenever } \{x,y\} \subset \mathbb{R}^d \qquad (5.6)$$

and we refer to [30, p. 159, Proposition 25.4] for examples of such functions. Any submultiplicative locally bounded function g is necessarily bounded in exponential growth [30, p. 160, Lemma 25.5], to wit:

$$\exists \{b,c\} \subset (0, \infty) \text{ such that } g(x) \leq be^{c|x|} \text{ for } x \in \mathbb{R}^d. \qquad (5.7)$$

Proposition 5.3. *Let $g : \mathbb{R}^d \to [0, \infty)$ be measurable, submultiplicative and locally bounded, and suppose $\int_{\mathbb{R}^d \setminus [-1,1]^d} g d\lambda < \infty$. Then for any $t > 0$, $\mathsf{E}[g \circ X_t] < \infty$ and, moreover, there is an $h_0 \in (0, h_\star)$ such that*

$$\sup_{h \in (0,h_0)} \mathsf{E}[g \circ X_t^h] < \infty.$$

Conversely, if $\int_{\mathbb{R}^d \setminus [-1,1]^d} g d\lambda = \infty$, then for all $t > 0$, $\mathsf{E}[g \circ X_t] = \infty$.

Proof. The argument follows the exposition given in [30, pp. 159-162], modifying the latter to the extent that uniform boundedness over $h \in (0, h_0)$ may be got. In particular, we refer to [30, p. 159, Theorem 25.3] for the claim that $\mathsf{E}[g \circ X_t] < \infty$, if and only if $\int_{\mathbb{R}^d \setminus [-1,1]^d} g d\lambda < \infty$. We take $\{a,b,c\} \subset (0, \infty)$ satisfying (5.6) and (5.7) above. Recall also that λ^h is the Lévy measure of the process X^h, $h \in (0, h_\star)$.

Now, decompose $X = X^1 + X^2$ and $X^h = X^{h1} + X^{h2}$, $h \in (0, h_\star)$ as independent sums, where X^1 is compound Poisson, Lévy measure $\lambda_1 := \mathbb{1}_{\mathbb{R}^d \setminus [-1,1]^d} \cdot \lambda$, and X^{h1} are also compound Poisson, Lévy measures $\lambda_1^h := \mathbb{1}_{\mathbb{R}^d \setminus [-1,1]^d} \cdot \lambda^h$, $h \in (0, h_\star)$. Consequently X^2 is a Lévy process with characteristic triplet $(\Sigma, \mathbb{1}_{[-1,1]^d} \cdot \lambda, \mu)_{\tilde{c}}$ and X^{h2} are compound Poisson, Lévy measures $\mathbb{1}_{[-1,1]^d} \cdot \lambda^h$, $h \in (0, h_\star)$. Moreover, for $h \in (0, h_\star)$, by submultiplicativity and independence:

$$\mathsf{E}[g \circ X_t^h] = \mathsf{E}[g \circ (X_t^{h1} + X_t^{h2})] \leq a\mathsf{E}[g \circ X_t^{h1}]\mathsf{E}[g \circ X_t^{h2}].$$

We first estimate $\mathsf{E}[g \circ X_t^{h1}]$. Let $(J_n)_{n \geq 1}$ (resp. N_t) be the sequence of jumps (resp. number of jumps by time t) associated to (resp. of) the compound Poisson process X^{h1}. Then $X_t^{h1} = \sum_{j=1}^{N_t} J_j$ and so by submultiplicativity:

$$\mathsf{E}[g \circ X_t^{h1}] \quad \leq \quad \mathsf{E}\left[g(0)\mathbb{1}_{\{N_t=0\}} + a^{N_t-1} \prod_{j=1}^{N_t} g(J_j)\mathbb{1}_{\{N_t>0\}} \right]$$

$$= g(0)e^{-t\lambda_1^h(\mathbb{R}^d)} + \sum_{n=1}^{\infty} \frac{t^n a^{n-1}}{n!} e^{-t\lambda_1^h(\mathbb{R}^d)} \left(\int g d\lambda_1^h \right)^n.$$

We also have for all $h \in (0, 1 \wedge h_\star)$:

$$
\begin{aligned}
\int g d\lambda_1^h &= \sum_{s \in \mathbb{Z}_h^d \setminus [-1,1]^d} \int_{A_s^h} g(s) d\lambda = \sum_{s \in \mathbb{Z}_h^d \setminus [-1,1]^d} \int_{A_s^h} g(u + (s-u)) d\lambda(u) \\
&\leq a \left(\sup_{k \in A_h^0} g(k) \right) \sum_{s \in \mathbb{Z}_h^d \setminus [-1,1]^d} \int_{A_s^h} g d\lambda, \text{ by submultiplicativity} \\
&\leq a \left(\sup_{k \in A_1^0} g(k) \right) \int_{\mathbb{R}^d \setminus [-1/2,1/2]^d} g d\lambda.
\end{aligned}
$$

Now, since g is locally bounded, λ is finite outside neighborhoods of 0, and since by assumption $\int_{\mathbb{R}^d \setminus [-1,1]^d} g d\lambda < \infty$, we obtain: $\sup_{h \in (0, 1 \wedge h_\star)} \mathsf{E}[g \circ X_t^{h1}] < \infty$.

Second, we consider $\mathsf{E}[g \circ X_t^{h2}]$. First, by boundedness in exponential growth and the triangle inequality:

$$\mathsf{E}[g \circ X_t^{h2}] \leq b\mathsf{E}[e^{c|X_t^{h2}|}] \leq b\mathsf{E}[e^{c\sum_{j=1}^d |X_{tj}^{h2}|}] = b\mathsf{E}\left[\prod_{j=1}^d e^{c|X_{tj}^{h2}|} \right].$$

It is further seen by a repeated application of the Cauchy-Schwartz inequality that it will be sufficient to show, for each $j \in \{1, \ldots, d\}$, that for some $h_0 \in (0, h_\star)$:

$$\sup_{h \in (0, h_0)} \mathsf{E}\left[e^{2^{d-1}c|X_{tj}^{h2}|} \right] < \infty.$$

Here $X_t^{h2} = (X_{t1}^{h2}, \ldots, X_{td}^{h2})$ and likewise for X_t^2. Fix $j \in \{1, \ldots, d\}$.

The characteristic exponent of X_j^{h2}, denoted Ψ_2^h, extends to an entire function on $\mathbb{C}$. Likewise for the characteristic exponent of X_j^2, denoted Ψ_2 [30, p. 160, Lemma 25.6]. Moreover, since, by expansion into power series, one has, locally uniformly in $\beta \in \mathbb{C}$, as $h \downarrow 0$:

- $\frac{e^{\beta h} + e^{-\beta h} - 2}{2h^2} \to \frac{1}{2}\beta^2$;

- $\frac{e^{\beta h} - e^{-\beta h}}{2h} \to \beta$;

- $\frac{e^{\beta h} - 1}{h} \to \beta$ and $\frac{1 - e^{-\beta h}}{h} \to \beta$;

since furthermore:

- $\left((\beta, u) \mapsto \frac{e^{\beta u} - \beta u - 1}{u^2} \right) : \mathbb{R} \setminus \{0\} \times \mathbb{C} \to \mathbb{C}$ is bounded on bounded subsets of its domain;

and since finally by the complex Mean Value Theorem [11, p. 859, Theorem 2.2]:

- as applied to the function $(x \mapsto e^{\beta x}) : \mathbb{C} \to \mathbb{C}$; $|e^{\beta x} - e^{\beta y}| \leq |x - y||\beta|(|e^{\beta z_1}| + |e^{\beta z_2}|)$ for some $\{z_1, z_2\} \subset \text{conv}(\{x, y\})$, for all $\{x, y\} \subset \mathbb{R}$;

- as applied to the function $(x \mapsto e^{\beta x} - \beta x) : \mathbb{C} \to \mathbb{C}$; $|e^{\beta x} - \beta x - (e^{\beta y} - \beta y)| \leq |x - y||\beta| \left(|e^{\beta z_1} - 1| + |e^{\beta z_2} - 1| \right)$ for some $\{z_1, z_2\} \in \text{conv}(\{x, y\})$, for all $\{x, y\} \subset \mathbb{R}$;

then the usual decomposition of the difference $\Psi_2^h - \Psi_2$ (see proof of Proposition 3.10) shows that $\Psi_2^h \to \Psi_2$ locally uniformly in $\mathbb{C}$ as $h \downarrow 0$. Next let ϕ_2^h and ϕ_2 be the characteristic functions of X_{tj}^{h2} and X_{tj}^2, respectively, $h \in (0, h_\star)$; themselves entire functions on

$\mathbb{C}$. Using the estimate of Lemma 4.3, we then see, by way of corollary, that also $\phi_2^h \to \phi_2$ locally uniformly in $\mathbb{C}$ as $h \downarrow 0$.

Now, since ϕ_2^h is an entire function, for $n \in \mathbb{N} \cup \{0\}$, $i^n \mathsf{E}[(X_{tj}^{h2})^n] = (\phi_2^h)^{(n)}(0)$ and it is Cauchy's estimate [32, p. 184, Lemma 10.5] that, for a fixed $r > 2^{d-1}c$, $\left|(\phi_2^h)^{(n)}(0)\right| \leq \frac{n!}{r^n} M^h$, where $M^h := \sup_{\{z \in \mathbb{C}: |z|=r\}} |\phi_2^h|$. Observe also that for some $h_0 \in (0, h_\star]$, $\sup_{h \in (0,h_0)} M^h < \infty$, since $\phi_2^h \to \phi_2$ locally uniformly as $h \downarrow 0$ and ϕ_2 is continuous (hence locally bounded).

Further to this $\mathsf{E}[|X_{tj}^{h2}|^{2k+1}] \leq 1 + \mathsf{E}[|X_{tj}^{h2}|^{2k+2}]$ ($k \in \mathbb{N} \cup \{0\}$) and $\mathsf{E}\left[e^{2^{d-1}c|X_{tj}^{h2}|}\right] = \sum_{n=0}^{\infty} \frac{1}{n!} \mathsf{E}[|X_{tj}^{h2}|^n](c2^{d-1})^n$. From this the desired conclusion finally follows. $\square$

The following result can now be established in dimension $d = 1$:

Proposition 5.4. *Let $d = 1$ and $t > 0$. Let furthermore:*

(i) *$g : \mathbb{R} \to [0, \infty)$, measurable, satisfy $\mathsf{E}[g \circ X_t] < \infty$, g locally bounded, submultiplicative, $g \neq 0$.*

(ii) *$f : \mathbb{R} \to \mathbb{C}$, measurable, be locally bounded, $\int_{\mathbb{R}} |f| \in (0, \infty]$, $|f|$ ultimately monotone (i.e. $|f||_{[K,\infty)}$ and $|f||_{(-\infty,-K]}$ monotone for some $K \in [0, \infty)$), $|f|/|g|$ ultimately nonincreasing (i.e. $(|f|/|g|)|_{[K,\infty)}$ and $(-|f|/|g|)|_{(-\infty,-K]}$ nonincreasing for some $K \in [0, \infty)$), and with the following Lipschitz property holding for some $\{a, A\} \in (0, \infty)$: $f|_{[-A,A]}$ is piecewise Lipschitz, whereas*

$$|f(z) - f(y)| \leq a|z - y|(g(z) + g(y)), \text{ whenever } \{z, y\} \subset \mathbb{R} \backslash (-A, A).$$

(iii) *$K : (0, \infty) \to [0, \infty)$, with $\lim_{0+} K = +\infty$.*

Then $|\mathsf{E}[f \circ X_t] - \mathsf{E}[f \circ X_t^h]|$ is of order:

$$O\left(\left(\int_{[-K(h),K(h)]} |f(x)|dx\right)(h \vee \Delta_t(h)) + \left(\frac{|f|}{|g|} \vee \frac{|f|}{|g|} \circ (-\mathrm{id}_{\mathbb{R}})\right)(K(h) - 3h/2)\right), \quad (5.8)$$

where $\Delta_t(h)$ is defined in (2.5).

Remark 5.5.

(1) *In (5.8) there is a balance of two terms, viz. the choice of the function K. Thus, the slower (resp. faster) that K increases to $+\infty$ at $0+$, the better the convergence of the first (resp. second) term, provided $f \notin L^1(\mathbb{R})$ (resp. $|f|/|g|$ is ultimately converging to 0, rather than it just being nonincreasing). In particular, when so, then the second term can be made to decay arbitrarily fast, whereas the first term will always have a convergence which is strictly worse than $h \vee \Delta_t(h)$. But this convergence can be made arbitrarily close to $h \vee \Delta_t(h)$ by choosing K increasing all the slower (this since f is locally bounded). In general the choice of K would be guided by balancing the rate of decay of the two terms.*

(2) *Since, in the interest of relative generality, (further properties of) f and λ are not specified, thus also g cannot be made explicit. Confronted with a specific f and Lévy process X, we should like to choose g approaching infinity (at $\pm \infty$) as fast as possible, while still ensuring $\mathsf{E}[g \circ X_t] < \infty$ (cf. Proposition 5.3). This makes, ceteris paribus, the second term in (5.8) decay as fast as possible.*

(3) *We exemplify this approach by considering two examples. Suppose for simplicity $\Delta_t(h) = O(h)$.*

(a) Let first $|f|$ be bounded by $(x \mapsto A|x|^n)$ for some $A \in (0, \infty)$ and $n \in \mathbb{N}$, and assume that for some $m \in (n, \infty)$, the function $g = (x \mapsto |x|^m \vee 1)$ satisfies $\mathsf{E}[g \circ X_t] < \infty$ (so that (i) holds). Suppose furthermore condition (ii) is satisfied as well (as it is for, e.g., $f = (x \mapsto x^n)$). It is then clear that the first term of (5.8) will behave as $\sim K(h)^{n+1}h$, and the second as $\sim K(h)^{-(m-n)}$, so we choose $K(h) \sim 1/h^{1/(1+m)}$ for a rate of convergence which is of order $O(h^{\frac{m-n}{m+1}})$.

(b) Let now $|f|$ be bounded by $(x \mapsto Ae^{\alpha|x|})$ for some $\{A, \alpha\} \subset (0, \infty)$, and assume that for some $\beta \in (\alpha, \infty)$, the function $g = (x \mapsto e^{\beta|x|})$ indeed satisfies $\mathsf{E}[g \circ X_t] < \infty$ (so that (i) holds). Suppose furthermore condition (ii) is satisfied as well (as it is for, e.g., $f = (x \mapsto (e^{\alpha x} - k)^+)$, where $k \in [0, \infty)$ — use Lemma 4.3). It is then clear that the first term of (5.8) will behave as $\sim e^{\alpha K(h)}h$, and the second as $\sim e^{-(\beta-\alpha)K(h)}$, so we choose, up to a bounded additive function of h, $K(h) = \log(1/h^{1/\beta})$ for a rate of convergence which is of order $O(h^{1-\frac{\alpha}{\beta}})$.

(4) Finally, note that Proposition 5.4 can, in particular, be applied to f, which is the mapping $(x \mapsto e^{ipx})$, $p \in \mathbb{R}$, once suitable functions g and K have been identified. This, however, would give weaker results than what can be inferred regarding the rate of the convergence of the characteristic functions $\phi_{X_t^h}(p) \to \phi_{X_t}(p)$ from Remark 3.11(i) (using Lemma 4.3, say). This is so, because the characteristic exponents admit the Lévy-Khintchine representation (allowing for a very detailed analysis of the convergence), a property that is lost for a general function f (cf. Remark 5.2(3)).

Proof of Proposition 5.4. This is a simple matter of estimation; for all sufficiently small $h > 0$:

$$|\mathsf{E}[f \circ X_t] - \mathsf{E}[f \circ X_t^h]| = \left| \int_{\mathbb{R}} f(z) p_t(z) dz - \sum_{y \in \mathbb{Z}_h} f(y) P_t^h(y) \right|$$

$$\leq \left| \sum_{y \in [-K(h), K(h)] \cap \mathbb{Z}_h} \left(\int_{A_y^h} f(z) p_t(z) dz - f(y) P_t^h(y) \right) \right| + \sum_{y \in \mathbb{Z}_h \setminus [-K(h), K(h)]} |f(y)| P_t^h(y) + \int_{\mathbb{R} \setminus [-(K(h)-h/2), K(h)-h/2]} |f(z)| p_t(z) dz$$

$$\leq \underbrace{\left| \sum_{y \in \mathbb{Z}_h \cap [-K(h), K(h)]} \int_{A_y^h} (f(z) - f(y)) p_t(z) dz \right|}_{(A)} + \underbrace{\left| \sum_{y \in \mathbb{Z}_h \cap [-K(h), K(h)]} \int_{A_y^h} f(y) (p_t(z) - p_t(y)) dz \right|}_{(B)} +$$

$$\underbrace{\left| \sum_{y \in \mathbb{Z}_h \cap [-K(h), K(h)]} f(y) h \left[p_t(y) - \frac{1}{h} P_t^h(y) \right] \right|}_{(C)} +$$

$$\underbrace{\left(\frac{|f|}{|g|} \vee \frac{|f|}{|g|} \circ (-\mathrm{id}_{\mathbb{R}}) \right) (K(h)) \, \mathsf{E}[g \circ X_t^h]}_{(D)} + \underbrace{\left(\frac{|f|}{|g|} \vee \frac{|f|}{|g|} \circ (-\mathrm{id}_{\mathbb{R}}) \right) (K(h) - h/2) \, \mathsf{E}[g \circ X_t]}_{(E)}.$$

Thanks to Proposition 5.3, and the fact that $|f|/|g|$ is ultimately nonincreasing, (D) & (E) are bounded (modulo a multiplicative constant) by $\frac{|f|}{|g|}(K(h)-h/2) \vee \frac{|f|}{|g|}(-(K(h)-h/2))$. From the Lipschitz property of f, submultiplicativity and local boundedness of g, and the fact that $\mathsf{E}[g \circ X_t] < \infty$, we obtain (A) is of order $O(h)$. By the local boundedness and eventual monotonicity of $|f|$, the Lipschitz nature of p_t and the fact that $\int |f| > 0$, (B) is bounded (modulo a multiplicative constant) by $h \int_{[-(K(h)+h), K(h)+h]} |f|$. Finally, a similar remark pertains to (C), but with $\Delta_t(h)$ in place of h. Combining these, using

once again $\int |f| > 0$, yields the desired result, since we may finally replace $K(h)$ by $(K(h) - h) \vee 0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

5.2 Algorithm

From a numerical perspective we must ultimately consider the processes X^h on a finite state space, which we take to be $S_M^h := \{x \in \mathbb{Z}_h^d : |x| \le M\}$ ($M > 0$, $h \in (0, h_\star)$). We let $\hat{Q}^h$ denote the sub-Markov generator got from Q^h by restriction to S_M^h, and we let $\hat{X}^h$ be the corresponding Markov chain got by killing X^h at the time $T_M^h := \inf\{t \ge 0 : |X_t^h| > M\}$, sending it to the coffin state ∂ thereafter.

Then the basis for the numerical evaluations is the observation that for a (finite state space) Markov chain Y with generator matrix Q, the probability $\mathsf{P}^y(Y_t = z)$ (resp. the expectation $\mathsf{E}^y[f \circ Y]$, when defined) is given by $(e^{tQ})_{yz}$ (resp. $(e^{tQ}f)(y)$). With this in mind we propose the:

Sketch algorithm

(i) Choose $\{h, M\} \subset (0, \infty)$.

(ii) Calculate, for the truncated sub-Markov generator $\hat{Q}^h$, the matrix exponential $\exp\{t\hat{Q}^h\}$ or action $\exp\{t\hat{Q}^h\}f$ thereof (where f is a suitable vector).

(iii) Adjust truncation parameter M, if needed, and discretization parameter h, until sufficient precision has been established.

Two questions now deserve attention: (1) what is the truncation error and (2) what is the expected cost of this algorithm. We address both in turn.

First, with a view to the localization/truncation error, we shall find use of the following:

Proposition 5.6. *Let $g : [0, \infty) \to [0, \infty)$ be nondecreasing, continuous and submultiplicative, with $\lim_{+\infty} g = +\infty$. Let $t > 0$ and denote by:*
$$X_t^\star = \sup_{s \in [0,t]} |X_s|, \quad X_t^{h\star} = \sup_{s \in [0,t]} |X_s^h|,$$
the running suprema of $|X|$ and of $|X^h|$, $h \in (0, h_\star)$, respectively. Suppose furthermore $\mathsf{E}[g \circ |X_t|] < \infty$. Then $\mathsf{E}[g \circ X_t^\star] < \infty$ and, moreover, there is some $h_0 \in (0, h_\star]$ such that
$$\sup_{h \in (0, h_0)} \mathsf{E}[g \circ X_t^{h\star}] < \infty.$$

Remark 5.7. *The function $g \circ |\cdot| : \mathbb{R}^d \to [0, \infty)$ is measurable, submultiplicative and locally bounded, so for a condition on the Lévy measure equivalent to $\mathsf{E}[g \circ X_t] < \infty$ see Proposition 5.3.*

We prove Proposition 5.6 below, but first let us show its relation to the truncation error. For a function $f : \mathbb{Z}_h^d \to \mathbb{R}$, we extend its domain to $\mathbb{Z}_h^d \cup \{\partial\}$, by stipulating that $f(\partial) = 0$. The following (very crude) estimates may then be made:

Corollary 5.8. *Fix $t > 0$. Assume the setting of Proposition 5.6. There is some $h_0 \in (0, h_\star]$ and then $C := \sup_{h \in (0, h_0)} \mathsf{E}[g \circ X_t^{h\star}] < \infty$, such that the following two claims hold:*

(i) For all $h \in (0, h_0)$:
$$\sum_{x \in \mathbb{Z}_h^d} |\mathsf{P}(X_t^h = x) - \mathsf{P}(\hat{X}_t^h = x)| = \mathsf{P}(T_M^h < t) \le C/g(M).$$

(ii) *Let* $f : \mathbb{Z}_h^d \to \mathbb{R}$ *and suppose* $|f| \leq \tilde{f} \circ |\cdot|$*, with* $\tilde{f} : [0, \infty) \to [0, \infty)$ *nondecreasing and such that* $\tilde{f}/g$ *is (resp. ultimately) nonincreasing. Then for all (resp. sufficiently large)* $M > 0$ *and* $h \in (0, h_0)$:

$$|\mathsf{E}[f \circ X_t^h] - \mathsf{E}[f \circ \hat{X}_t^h]| \leq C \left(\frac{\tilde{f}}{g} \right) (M).$$

Remark 5.9.

1. *Ad (i). Note that* M *may be taken fixed (i.e. independent of* h*) and chosen so as to satisfy a prescribed level of precision. In that case such a choice may be verified explicitly at least retrospectively: the sub-Markov generator* $\hat{Q}^h$ *gives rise to the sub-Markov transition matrix* $\hat{P}_t^h := e^{t\hat{Q}^h}$*; its deficit (in the row corresponding to state* 0*) is precisely the probability* $\mathsf{P}(T_M^h < t)$*.*

2. *Ad (ii). But* M *may also be made to depend on* h*, and then let to increase to* $+\infty$ *as* $h \downarrow 0$*, in which case it is natural to balance the rate of decay of* $|\mathsf{E}[f \circ X_t^h] - \mathsf{E}[f \circ \hat{X}_t^h]|$ *against that of* $|\mathsf{E}[f \circ X_t] - \mathsf{E}[f \circ X_t^h]|$ *(cf. Proposition 5.4). In particular, since* $\mathsf{E}[g \circ |X_t|] < \infty \Leftrightarrow \mathsf{E}[g \circ X_t^\star] \Leftrightarrow \int_{\mathbb{R}^d \setminus [-1,1]^d} g \circ |\cdot| d\lambda < \infty$ *[30, p. 159, Theorem 25.3 & p. 166, Theorem 25.18], this problem is essentially analogous to the one in Proposition 5.4. In particular, Remark 5.5 extends in a straightforward way to account for the truncation error, with* M *in place of* $K(h) - 3h/2$*.*

Proof. (i) follows from the estimate $\sum_{x \in \mathbb{Z}_h^d} |\mathsf{P}(X_t^h = x) - \mathsf{P}(\hat{X}_t^h = x)| = \mathsf{P}(T_M^h < t) = \mathsf{P}(X_t^{h\star} > M) \leq \frac{\mathsf{E}[g \circ X_t^{h\star}]}{g(M)}$, which is an application of Markov's inequality. When it comes to (ii), we have for all (resp. sufficiently large) $M > 0$:

$$|\mathsf{E}[f \circ X_t^h] - \mathsf{E}[f \circ \hat{X}_t^h]| \leq \mathsf{E}\left[(|f| \circ X_t^h) \mathbb{1}(T_M^h < t) \right] \leq \mathsf{E}\left[\left(\tilde{f} \circ |X_t^h| \right) \mathbb{1}(T_M^h < t) \right]$$

$$\leq \mathsf{E}\left[\left(\tilde{f} \circ X_t^{h\star} \right) \mathbb{1}(T_M^h < t) \right] = \mathsf{E}\left[\left(\left(\frac{\tilde{f}}{g} \right) \circ X_t^{h\star} \right) \left(g \circ X_t^{h\star} \right) \mathbb{1}(X_t^{h\star} > M) \right]$$

$$\leq \left(\frac{\tilde{f}}{g} \right) (M) \mathsf{E}[g \circ X_t^{h\star}],$$

whence the desired conclusion follows. $\square$

Proof of Proposition 5.6. We refer to [30, p. 166, Theorem 25.18] for the proof that $\mathsf{E}[g \circ X_t^\star] < \infty$. Next, by right continuity of the sample paths of X, we may choose $b > 0$, such that $\mathsf{P}(X_t^* \leq b/2) > 0$ and we may also insist on $b/2$ being a continuity point of the distribution function of $X_t^\star$ (there being only denumerably many points of discontinuity thereof). Now, $X^h \to X$ as $h \downarrow 0$ w.r.t. the Skorokhod topology on the space of càdlàg paths. Moreover, by [17, p. 339, 2.4 Proposition], the mapping $\Phi := (\alpha \mapsto \sup_{s \in [0,t]} |\alpha(s)|) : \mathbb{D}([0, \infty), \mathbb{R}^d) \to \mathbb{R}$ is continuous at every point α in the space of càdlàg paths $\mathbb{D}([0, \infty), \mathbb{R}^d)$, which is continuous at t. In particular, Φ is continuous, a.s. w.r.t. the law of the process X on the Skorokhod space [30, p. 59, Theorem 11.1]. By the Portmanteau Theorem, it follows that there is some $h_0 \in (0, h_\star]$ such that $\inf_{h \in (0,h_0)} \mathsf{P}(X_t^{h\star} \leq b/2) > 0$.

Moreover, from the proof of [30, p. 166, Theorem 25.18], by letting $\tilde{g} : [0, \infty) \to [0, \infty)$ be nondecreasing, continuous, vanishing at zero and agreeing with g on restriction to $[1, \infty)$, we may then show for each $h \in (0, h_\star)$ that:

$$\mathsf{E}[\tilde{g} \circ (X_t^{h\star} - b); X_t^{h\star} > b] \leq \mathsf{E}[\tilde{g} \circ |X_t^h|]/\mathsf{P}(X_t^{h\star} \leq b/2).$$

Now, since $\mathsf{E}[g \circ |X_t|] < \infty$, by Proposition 5.3 (cf. Remark 5.7), there is some $h_0 \in (0, h_\star]$ such that $\sup_{h \in (0, h_0)} \mathsf{E}[g \circ |X_t^h|] < \infty$, and thus $\sup_{h \in (0, h_0)} \mathsf{E}[\tilde{g} \circ |X_t^h|] < \infty$.

Combining the above, it follows that for some $h_0 \in (0, h_\star]$, $\sup_{h \in (0, h_0)} \mathsf{E}[\tilde{g} \circ (X_t^{h\star} - b); X_t^{h\star} > b] < \infty$ and thus $\sup_{h \in (0, h_0)} \mathsf{E}[g \circ (X_t^{h\star} - b); X_t^{h\star} > b] < \infty$. Finally, an application of submultiplicativity of g allows to conclude. □

Having thus dealt with the truncation error, let us briefly discuss the cost of our algorithm.

The latter is clearly governed by the calculation of the matrix exponential, or, resp., of its action on some vector. Indeed, if we consider as fixed the generator matrix $\hat{Q}^h$, and, in particular, its dimension $n \sim (M/h)^d$, then this may typically require $O(n^3)$ [25, 16], resp. $O(n^2)$ [1], floating point operations. Note, however, that this is a notional complexity analysis of the algorithm. A more detailed argument would ultimately have to specify precisely the particular method used to determine the (resp. action of a) matrix exponential, and, moreover, take into account how $\hat{Q}^h$ (and, possibly, the truncation parameter M, cf. Remark 5.9) behave as $h \downarrow 0$. Further analysis in this respect goes beyond the desired scope of this paper.

We finish off by giving some numerical experiments in the univariate case. To compute the action of $\hat{Q}^h$ on a vector we use the MATLAB function `expmv.m` [1], unless $\hat{Q}^h$ is sparse, in which case we use the MATLAB function `expv.m` from [31].

We begin with transition densities. To shorten notation, fix the time $t = 1$ and allow $p := p_1(0, \cdot)$ and $p^h := \frac{1}{h} \hat{P}_1^h(0, \cdot)$ ($\hat{P}^h$ being the analogue of P^h for the process $\hat{X}^h$). Note that to evaluate the latter, it is sufficient to compute $(e^{\hat{Q}^h t})_{0\cdot} = e^{(\hat{Q}^h)' t} \mathbb{1}_{\{0\}}$, where $(\hat{Q}^h)'$ denotes transposition.

Example 5.10. *Consider first Brownian motion with drift, $\sigma^2 = 1$, $\mu = 1$, $\lambda = 0$ (scheme 1, $V = 0$). We compare the density p with the approximation p^h ($h \in \{1/2^n : n \in \{0, 1, 2, 3\}\}$) on the interval $[0, 2]$ (see Figure 2 on p. 32), choosing $M = 5$. The vector of deficit probabilities $(\mathsf{P}(T_M^{1/2^n} < t))_{n=0}^3$ corresponding to using this truncation was $(5.9 \cdot 10^{-4}, 1.5 \cdot 10^{-4}, 5.8 \cdot 10^{-5}, 4.4 \cdot 10^{-5})$. In this case the matrix $\hat{Q}^h$ is sparse.*

Example 5.11. *Consider now α-stable Lévy processes, $\sigma^2 = 0$, $\mu = 0$, $\lambda(dx) = dx/|x|^{1+\alpha}$ (scheme 2, $V = 1$). We compare the density p with p^h on the interval $[0, 1]$ (see Figure 4 on p. 33). Computations are made for the vector of alphas given by $(\alpha_k)_{k=1}^4 := (1/2, 1, 4/3, 5/3)$ with corresponding truncation parameters $(M_k)_{k=1}^4 = (500, 100, 30, 20)$ resulting in the deficit probabilities (uniformly over the h considered) of $(\mathsf{P}(T_{M_k}^h < t))_{k=1}^4 = (1.7 \cdot 10^{-1}, 2.0 \cdot 10^{-2}, (\text{from } 1.7 \text{ to } 1.8) \cdot 10^{-2}, (\text{from } 0.94 \text{ to } 1.01) \cdot 10^{-2})$. The heavy tails of the Lévy density necessitate a relatively high value of M. Nevertheless, excluding the case $\alpha = 5/3$, a reduction of M by a factor of 5 resulted in an absolute change of the approximating densities, which was at most of the order of magnitude of the discretization error itself. Conversely, for $\alpha = 1/2$, when the deficit probability is highest and appreciable, increasing M by a factor of 2, resulted in an absolute change of the calculated densities of the order 10^{-6} (uniformly over $h \in \{1, 1/2, 1/4\}$). Finally, note that $\alpha = 1$ gives rise to the Cauchy distribution, whereas otherwise we use the MATLAB function `stblpdf.m` to get a benchmark density against which a comparison can be made.*

Example 5.12. *A particular VG model [5, 23] has $\sigma^2 = 0$, $\mu = 0$, $\lambda(dx) = \frac{e^{-|x|}}{|x|} \mathbb{1}_{\mathbb{R} \setminus \{0\}}(x) dx$ (scheme 2, $V = 1$). Again we compare p with p^h ($h \in \{1/2^n : n \in \{0, 1, 2, 3\}\}$) on the interval $[0, 1]$ (see Figure 3 on p. 32), choosing $M = 5$. The vector of deficit probabilities $(\mathsf{P}(T_M^{1/2^n} < t))_{n=0}^3$ corresponding to using this truncation was $(5.2 \cdot 10^{-3}, 6.4 \cdot 10^{-3}, 7.2 \cdot 10^{-3}, 7.6 \cdot 10^{-3})$. The density p is given explicitly by $(x \mapsto e^{-|x|}/2)$.*
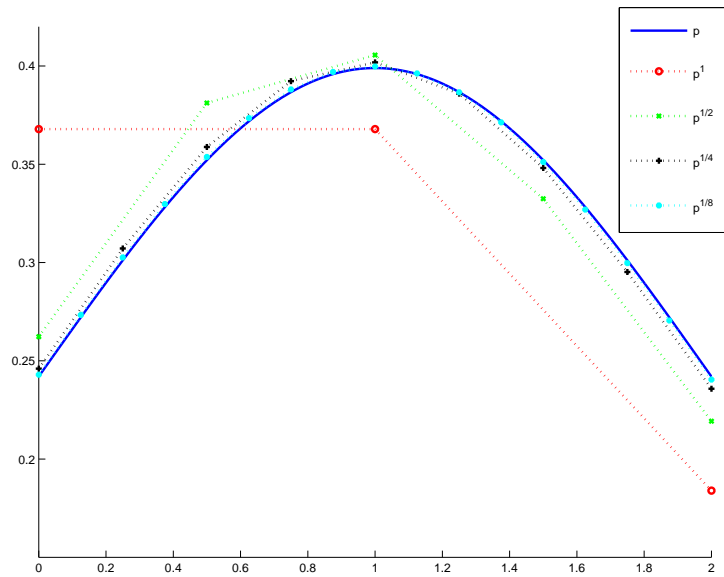
Figure 2: Convergence of p^h to p (as $h \downarrow 0$) on the interval $[0, 2]$ for Brownian motion with drift ($\sigma^2 = \mu = 1$, $\lambda = 0$, scheme 1, $V = 0$). See Example 5.10 for details.
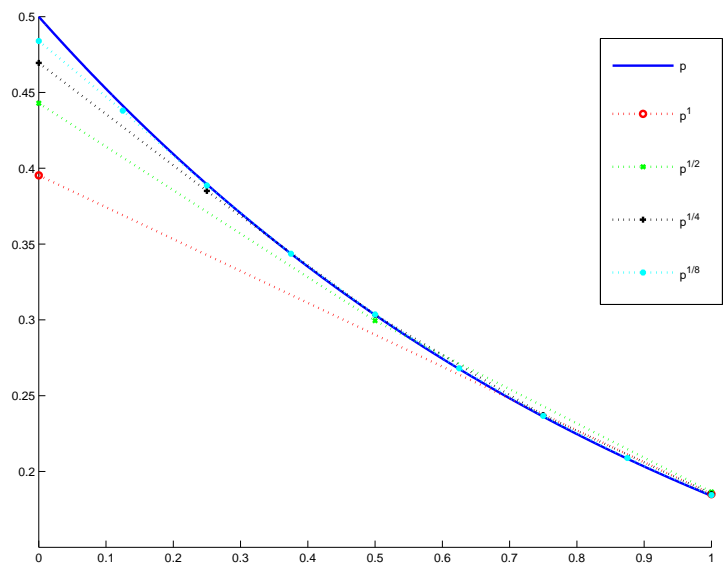

Figure 3: Convergence of p^h to p (as $h \downarrow 0$) on the interval $[0, 1]$ for the VG model ($\sigma^2 = 0$, $\mu = 0$, $\lambda(dx) = \frac{e^{-|x|}}{|x|}\mathbb{1}_{\mathbb{R}\setminus\{0\}}(x)dx$, scheme 2, $V = 1$). Note that in this case Orey's condition fails, but (at least as evidenced numerically) linear convergence does not. See Example 5.12 for details.
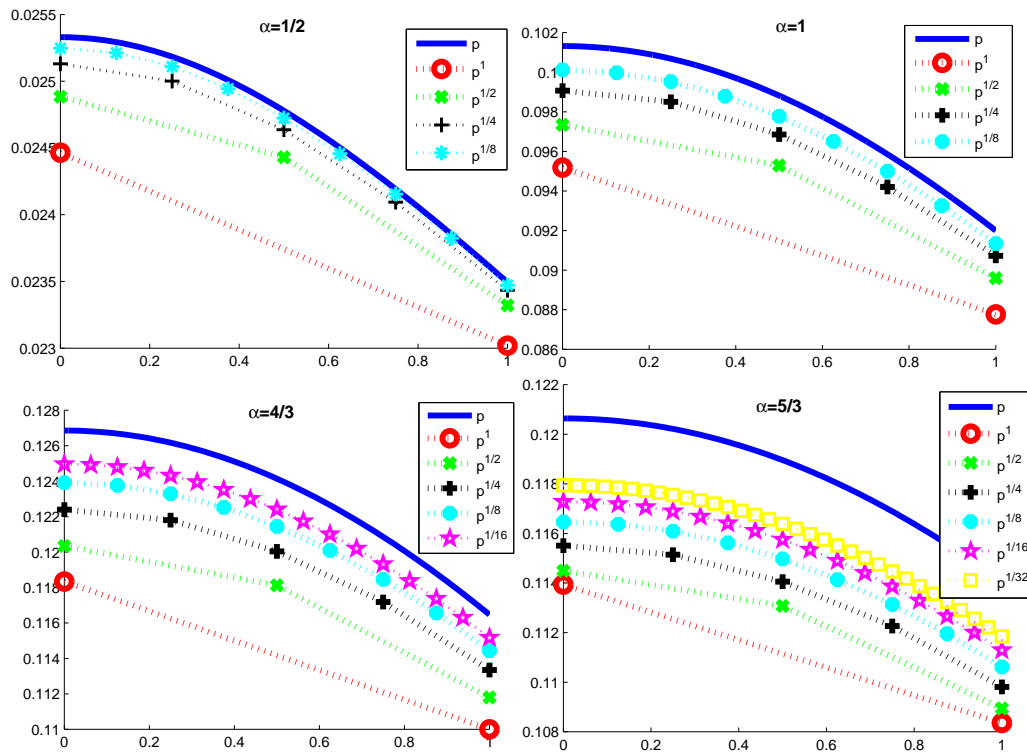
Figure 4: Convergence of p^h to p (as $h \downarrow 0$) on the interval $[0,1]$ for α-stable Lévy processes ($\sigma^2 = 0$, $\mu = 0$, $\lambda(dx) = dx/|x|^{1+\alpha}$, scheme 2, $V = 1$), $\alpha \in \{1/2, 1, 4/3, 5/3\}$. See Example 5.11 for details. Note that convergence becomes progressively worse as $\alpha \uparrow$, which is *precisely consistent* with Figure 1 and the theoretical order of convergence, this being $O(h^{(2-\alpha)\wedge 1})$ (up to a slowly varying factor $\log(1/h)$, when $\alpha = 1$; and noting that Orey's condition is satisfied with $\epsilon = \alpha$). For example, when $\alpha = 5/3$ each successive approximation should be closer to the limit by a factor of $\left(\frac{1}{2}\right)^{1/3} \doteq 0.8$, as it is.

Finally, to illustrate convergence of expectations, we consider a particular option pricing problem.

Example 5.13. *Suppose that, under the pricing measure, the stock price process $S = (S_t)_{t \geq 0}$ is given by $S_t = S_0 e^{rt + X_t}$, $t \geq 0$, where S_0 is the initial price, r is the interest rate, and X is a tempered stable process with Lévy measure given by:*

$$\lambda(dx) = c\left(\frac{e^{-\lambda_+ x}}{x^{1+\alpha}}\mathbb{1}_{(0,\infty)}(x) + \frac{e^{-\lambda_- |x|}}{|x|^{1+\alpha}}\mathbb{1}_{(-\infty,0)}(x)\right)dx.$$

To satisfy the martingale condition, we must have $\mathsf{E}[e^{X_t}] \equiv 1$, which in turn uniquely determines the drift μ (we have, of course, $\sigma^2 = 0$). The price of the European put option with maturity T and strike K at time zero is then given by:

$$P(T,K) = e^{-rT}\mathsf{E}[(K - S_T)^+].$$

We choose the same value for the parameters as [28], namely $S_0 = 100$, $r = 4\%$, $\alpha = 1/2$, $c = 1/2$, $\lambda_+ = 3.5$, $\lambda_- = 2$ and $T = 0.25$, so that we may quote the reference values $P(T,K)$ from there.

Now, in the present case, X is a process of finite variation, i.e. $\kappa(0) < \infty$, hence convergence of densities is of order $O(h)$, since Orey's condition holds with $\epsilon = 1/2$ (scheme 2, $V = 1$). Moreover, $\mathbb{1}_{\mathbb{R}\setminus[-1,1]} \cdot \lambda$ integrates $(x \mapsto e^{2|x|})$, whereas the function $(x \mapsto (K - e^{rt+x})^+)$ is bounded. Pursuant to (2) of Remark 5.9 we thus choose $M = M(h) := \left(\frac{1}{2}\log(1/h)\right) \vee 1$, which by Corollary 5.8 and Proposition 5.4 (with $K(h) = M(h)$) (cf. also ((3)b) of Remark 5.5) ensures that:

$$|\hat{P}^h(T, K) - P(T, K)| = O(h\log(1/h)),$$

where $\hat{P}^h(T, K) := e^{-rT}\mathsf{E}[(K - S_0 e^{rT+\hat{X}^h_T})^+]$. Table 4 on p. 35 summarizes this convergence on the decreasing sequence $h_n := 1/2^n$, $n \geq 1$.

In particular, we wish to emphasize that the computations were all (reasonably) fast. For example, to compute the vector $(\hat{P}^{h_n}(T, K))_{n=1}^9$ with $K = 80$, the times (in seconds; entry-by-entry) $(0.0106, 0.0038, 0.0044, 0.0078, 0.0457, 0.0367, 0.0925, 0.4504, 2.4219)$ were required on an Intel 2.53 GHz processor (times obtained using MATLAB's `tic-toc` facility). This is much better than, e.g., the Monte Carlo method of [28] and comparable with the finite difference method of [8] (VG2 model in [8, p. 1617, Section 7]).

In conclusion, the above numerical experiments serve to indicate that our method behaves robustly when the Blumenthal-Getoor index of the Lévy measure is not too close to 2 (in particular, if the pure-jump part has finite variation). It does less well if this is not the case, since then the discretisation parameter h must be chosen small, which is expensive in terms of numerics (viz. the size of $\hat{Q}^h$).

$K \to$	80	85	90	95	100	105	110	115	120
$P(T,K) \to$	1.7444	2.3926	3.2835	4.5366	6.3711	9.1430	12.7631	16.8429	21.1855
n	$\hat{P}^{h_n}(T,K) - P(T,K)$								
1	0.6411	0.5422	0.2006	-0.5033	-1.7885	-0.8227	0.0970	0.5570	0.7542
2	-0.1089	0.2816	0.4295	0.2151	-0.5806	0.0975	0.5341	0.5109	0.2250
3	-0.2271	-0.1596	-0.1928	0.0920	-0.2046	0.1405	0.0348	-0.4356	-0.3937
4	-0.0904	-0.0753	-0.0517	-0.0442	0.0652	0.1487	0.0057	-0.1511	-0.1838
5	-0.0411	-0.0338	-0.0193	-0.0053	0.0679	0.0569	-0.0073	-0.0616	-0.0833
6	-0.0184	-0.0163	-0.0081	0.0022	0.0347	0.0314	-0.0033	-0.0244	-0.0384
7	-0.0079	-0.0069	-0.0040	0.0019	0.0152	0.0109	-0.0034	-0.0108	-0.0164
8	-0.0034	-0.0029	-0.0016	0.0011	0.0072	0.0053	-0.0012	-0.0048	-0.0070
9	-0.0014	-0.0012	-0.0007	0.0006	0.0033	0.0026	-0.0004	-0.0020	-0.0030

Table 4: Convergence of the put option price for a CGMY model (scheme 2, $V = 1$). See Example 5.13 for details.

References

[1] Awad H. Al-Mohy and Nicholas J. Higham, *Computing the action of the matrix exponential, with an application to exponential integrators*, SIAM J. Sci. Comput. **33** (2011), no. 2, 488–511. MR-2785959

[2] Vlad Bally and Denis Talay, *The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density*, Monte Carlo Methods Appl. **2** (1996), no. 2, 93–128. MR-1401964

[3] Jean Bertoin, *Lévy processes*, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, Cambridge, 1996. MR-1406564

[4] Björn Böttcher and René L. Schilling, *Approximation of Feller processes by Markov chains with Lévy increments*, Stoch. Dyn. **9** (2009), no. 1, 71–80. MR-2502474

[5] Peter P. Carr, Hélyette Geman, Dilip B. Madan, and Marc Yor, *The fine structure of asset returns: an empirical investigation*, Journal of Business **75** (2002), no. 2, 305–332.

[6] Serge Cohen and Jan Rosiński, *Gaussian approximation of multivariate Lévy processes with applications to simulation of tempered stable processes*, Bernoulli **13** (2007), no. 1, 195–210. MR-2307403

[7] Rama Cont and Peter Tankov, *Financial modelling with jump processes*, Chapman & Hall/CRC Financial Mathematics Series, Chapman & Hall/CRC, Boca Raton, FL, 2004. MR-2042661

[8] Rama Cont and Ekaterina Voltchkova, *A finite difference scheme for option pricing in jump diffusion and exponential Lévy models*, SIAM J. Numer. Anal. **43** (2005), no. 4, 1596–1626 (electronic). MR-2182141

[9] John Crosby, Nolwenn Le Saux, and Aleksandar Mijatović, *Approximating Lévy processes with a view to option pricing*, Int. J. Theor. Appl. Finance **13** (2010), no. 1, 63–91. MR-2646974

[10] Richard M. Dudley, *Real analysis and probability*, Cambridge Studies in Advanced Mathematics, vol. 74, Cambridge University Press, Cambridge, 2002, Revised reprint of the 1989 original. MR-1932358

[11] J.-Cl. Evard and F. Jafari, *A complex Rolle's theorem*, Amer. Math. Monthly **99** (1992), no. 9, 858–861. MR-1191706

[12] José E. Figueroa-López, *Approximations for the distributions of bounded variation Lévy processes*, Statist. Probab. Lett. **80** (2010), no. 23-24, 1744–1757. MR-2734238

[13] Damir Filipović, Eberhard Mayerhofer, and Paul Schneider, *Density approximations for multivariate affine jump-diffusion processes*, J. Econometrics **176** (2013), no. 2, 93–111. MR-3084047

[14] M. G. Garroni and J.-L. Menaldi, *Green functions for second order parabolic integro-differential problems*, Pitman Research Notes in Mathematics Series, vol. 275, Longman Scientific & Technical, Harlow, 1992. MR-1202037

[15] Paul Glasserman, *Monte Carlo methods in financial engineering*, Applications of Mathematics (New York), vol. 53, Springer-Verlag, New York, 2004, Stochastic Modelling and Applied Probability. MR-1999614

[16] Nicholas J. Higham, *The scaling and squaring method for the matrix exponential revisited*, SIAM J. Matrix Anal. Appl. **26** (2005), no. 4, 1179–1193 (electronic). MR-2178217

[17] Jean Jacod and Albert N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR-1943877

[18] Jonas Kiessling and Raúl Tempone, *Diffusion approximation of Lévy processes with a view towards finance*, Monte Carlo Methods Appl. **17** (2011), no. 1, 11–45. MR-2784742

[19] Peter E. Kloeden and Eckhard Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics (New York), vol. 23, Springer-Verlag, Berlin, 1992. MR-1214374

[20] Viktorya Knopova and René L. Schilling, *Transition density estimates for a class of Lévy and Lévy-type processes*, J. Theoret. Probab. **25** (2012), no. 1, 144–170. MR-2886383

[21] A. Kohatsu-Higa, S. Ortiz-Latorre, and P. Tankov, *Optimal simulation schemes for Lévy driven stochastic differential equations*, Math. Comp. (to appear).

[22] Andreas E. Kyprianou, *Introductory lectures on fluctuations of Lévy processes with applications*, Universitext, Springer-Verlag, Berlin, 2006. MR-2250061

[23] Dilip B. Madan, Peter Carr, and Eric C. Chang, *The Variance Gamma Process and Option Pricing*, European Finance Review **2** (1998), 79–105.

[24] Aleksandar Mijatović, *Spectral properties of trinomial trees*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **463** (2007), no. 2083, 1681–1696. MR-2331531

[25] Cleve Moler and Charles Van Loan, *Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later*, SIAM Rev. **45** (2003), no. 1, 3–49 (electronic). MR-1981253

[26] J. R. Norris, *Markov chains*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 2, Cambridge University Press, Cambridge, 1998, Reprint of 1997 original. MR-1600720

[27] Steven Orey, *On continuity properties of infinitely divisible distribution functions*, Ann. Math. Statist. **39** (1968), 936–937. MR-0226701

[28] Jérémy Poirot and Peter Tankov, *Monte Carlo Option Pricing for Tempered Stable (CGMY) Processes*, Asia-Pacific Financial Markets **13** (2006), no. 4, 327–344.

[29] Jan Rosiński, *Simulation of Lévy processes*, Encyclopedia of Statistics in Quality and Reliability: Computationally Intensive Methods and Simulation, Wiley, 2008.

[30] Ken-iti Sato, *Lévy processes and infinitely divisible distributions*, Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University Press, Cambridge, 1999, Translated from the 1990 Japanese original, Revised by the author. MR-1739520

[31] Roger B. Sidje, *Expokit: a software package for computing matrix exponentials*, ACM Trans. Math. Softw. **24** (1998), no. 1, 130–156, MATLAB codes at `http://www.maths.uq.edu.au/expokit/`.

[32] Ian Stewart and David Tall, *Complex analysis*, Cambridge University Press, Cambridge, 1983, The hitchhiker's guide to the plane. MR-698076

[33] Alex Szimayer and Ross A. Maller, *Finite approximation schemes for Lévy processes, and their application to optimal stopping problems*, Stochastic Process. Appl. **117** (2007), no. 10, 1422–1447. MR-2353034

[34] Paweł Sztonyk, *Transition density estimates for jump Lévy processes*, Stochastic Process. Appl. **121** (2011), no. 6, 1245–1265. MR-2794975

[35] Hideyuki Tanaka and Arturo Kohatsu-Higa, *An operator approach for Markov chain weak approximations with an application to infinite activity Lévy driven SDEs*, Ann. Appl. Probab. **19** (2009), no. 3, 1026–1062. MR-2537198

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS[1])

- Non profit, sponsored by IMS[2], BS[3], PKP[4]

- Purely electronic and secure (LOCKSS[5])

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]OJS: Open Journal Systems http://pkp.sfu.ca/ojs/
[2]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[3]BS: Bernoulli Society http://www.bernoulli-society.org/
[4]PK: Public Knowledge Project http://pkp.sfu.ca/
[5]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm