The Annals of Applied Probability
2016, Vol. 26, No. 1, 328–359
DOI: 10.1214/14-AAP1093

APPROXIMATING LÉVY PROCESSES WITH COMPLETELY MONOTONE JUMPS

BY DANIEL HACKMANN[1] AND ALEXEY KUZNETSOV[2]

York University

Lévy processes with completely monotone jumps appear frequently in various applications of probability. For example, all popular stock price models based on Lévy processes (such as the Variance Gamma, CGMY/KoBoL and Normal Inverse Gaussian) belong to this class. In this paper we continue the work started in [*Int. J. Theor. Appl. Finance* **13** (2010) 63–91, *Quant. Finance* **10** (2010) 629–644] and develop a simple yet very efficient method for approximating processes with completely monotone jumps by processes with hyperexponential jumps, the latter being the most convenient class for performing numerical computations. Our approach is based on connecting Lévy processes with completely monotone jumps with several areas of classical analysis, including Padé approximations, Gaussian quadrature and orthogonal polynomials.

1. Introduction. Most researchers working in Applied Mathematics are familiar with the problem of choosing the right mathematical objects for their modeling purposes: one needs to strike a balance between the simplicity of the model, its analytical and numerical tractability and its ability to provide a realistic description of the phenomenon. For example, when modeling stock prices in mathematical finance we are faced with the following dilemma: do we choose a process which fits the empirically observed behavior of stock prices (such as having jumps of infinite activity [8]), or do we settle for a simpler model which provides for explicit formulas and efficient numerical algorithms? The first choice would lead to the most popular families of Lévy processes, such as the Variance Gamma (VG), CGMY/KoBoL, Meixner and Normal Inverse Gaussian (NIG) families. These processes, which belong to a wider class of processes with completely monotone jumps, provide a good fit for market data, and they are flexible enough to accommodate for such desirable features as jumps of infinite activity and finite or infinite variation. They also enjoy a certain degree of analytical tractability (e.g., VG and NIG processes have explicit transition probability densities), and European option prices and Greeks can be computed quite easily. However, the computation

Received April 2014; revised December 2014.

[1]Supported by the Ontario Graduate Scholarship Program.

[2]Supported by the Natural Sciences and Engineering Research Council of Canada.

MSC2010 subject classifications. Primary 60G51; secondary 26C15.

Key words and phrases. Lévy processes, complete monotonicity, hyperexponential processes, Padé approximation, rational interpolation, Gaussian quadrature, Stieltjes functions, Jacobi polynomials.

of more exotic option prices (such as barrier, lookback and Asian options) is a much more challenging task. On the other hand, hyperexponential processes (also known as "hyperexponential jump-diffusion processes," see [7]), and more general processes with jumps of rational transform (see [13, 24, 28]) form the most convenient class for performing numerical calculations. This is due to the fact that these processes have an explicit Wiener–Hopf factorization, which leads to simple and efficient numerical algorithms for pricing barrier and lookback options [7, 19] and Asian options [6]. One might think that hyperexponential processes are perfect candidates for modeling stock prices, yet they have a major flaw in that their jumps are necessarily of finite activity, which seems to be incompatible with empirical results [8].

A natural way to reconcile these two competing objectives is to approximate processes with completely monotone jumps with the hyperexponential processes. Two approximations of this sort were developed recently: Jeannin and Pistorius [19] use the least squares optimization in order to find the approximating hyperexponential process, while Crosby, Le Saux and Mijatović [11] use a more direct approach based on the Gaussian quadrature. Our goal in this paper is to present a new method for approximating Lévy processes with completely monotone jumps, and to demonstrate that this method is natural, simple and very efficient.

Let us present the main ideas behind our approach. Approximating a Lévy process X is equivalent to finding an approximation to its *Laplace exponent*, defined as $\psi(z) := \ln \mathbb{E}[\exp(zX_1)]$. The Laplace exponent of a hyperexponential processes is a rational function, therefore, our problem reduces to two steps: (i) finding a good rational approximation $\tilde{\psi}(z) \approx \psi(z)$, and (ii) ensuring that the rational function $\tilde{\psi}(z)$ is itself a Laplace exponent of some Lévy process $\tilde{X}$. For the first step, we rely on the extensive literature on rational approximations and interpolations. One of the simplest and the most natural methods of rational approximation is the *Padé approximation*; see the classical book by Baker [3] for an excellent account of this theory. The Padé approximation $f^{[m/n]}(x)$, of a function $f(x) = \sum_{k\geq 0} c_n x^k$, is defined as a rational function $P_m(x)/Q_n(x)$ [where P and Q are polynomials satisfying $\deg(P) \leq n$ and $\deg(Q) \leq m$] which matches the first $n + m + 1$ Taylor coefficients of $f(x)$. Padé approximations are easy to compute, and there exists a well developed theory related to their various properties (convergence, error estimates, etc.). Thus, the first step of our program is rather simple, but the second step is much more challenging: we need to ensure that the approximating rational function $\psi^{[m/n]}(z)$ is itself the Laplace exponent of some Lévy process Y. First, we can considerably reduce the number of possible cases that we need to study. It is known (see Proposition 2 on page 16 in [4]) that the Laplace exponent of a Lévy process satisfies $\psi(\mathrm{i}z) = O(z^2)$ as $z \to \infty$, therefore, the functions $\psi^{[m/n]}(z)$ cannot be Laplace exponents if $m > n + 2$. If $m < n$ then necessarily $\psi^{[m/n]}(\mathrm{i}z) \to 0$ as $z \to \infty$, and one can prove[3] that a rational function with this property cannot

[3] Assume that a Lévy process Y has a rational function $f(z) = P(z)/Q(z)$ as its Laplace exponent. Do the partial fraction decomposition of $f(z)$ and identify the Lévy measure of Y via the Lévy–

be the Laplace exponent of a Lévy process Y (unless $Y = 0$ almost surely). This shows that in the full table of Padé approximations $\psi^{[m/n]}(z)$ only the following functions:

$$\psi^{[n/n]}(z), \qquad \psi^{[n+1/n]}(z) \quad \text{and} \quad \psi^{[n+2/n]}(z) \tag{1}$$

can possibly be Laplace exponents of a Lévy process.

Checking whether a given function is the Laplace exponent of a Lévy process is a very difficult task: one would need to show that the function can be represented via the Lévy–Khintchine formula [see formula (2) below]. Since it is impossible to verify this property numerically, one would require some additional qualitative information about the function. In our case, this additional information comes from the fact that $\psi(z)$ is the Laplace exponent of a process with *completely monotone jumps*. Using this key fact and utilizing connections with several branches of classical analysis (such as the theory of Padé approximations, orthogonal polynomials, Stieltjes functions and Gaussian quadrature), we are able to completely characterize all cases when the functions in (1) are Laplace exponents of Lévy processes. Our main result states that if the original Lévy process has completely monotone jumps, the function $\psi^{[n+1/n]}(z)$ is a Laplace exponent of a hyperexponential Lévy process $X^{(n)}$, which converges to X in distribution as $n \to +\infty$. Moreover, if the process X has only positive (or only negative) jumps, the same results holds true for $\psi^{[n+2/n]}(z)$ [and for $\psi^{[n/n]}(z)$ under the additional assumption that the process has jumps of finite variation].

The paper is organized as follows. Section 2 contains our main results on approximating Lévy processes with completely monotone jumps (treating the two-sided and one-sided cases separately). Section 3 discusses the important special cases of the Gamma subordinator and of the one-sided tempered stable processes; in both cases the Padé approximation is given explicitly. In this section we also discuss how to use these results to construct explicit approximations to VG, CGMY and NIG processes, and present some extensions of our approximation scheme, including (i) the use of Padé approximation centered at an arbitrary point and (ii) a more general multi-point rational interpolation technique. In Section 4, we present the results of several numerical experiments which demonstrate the efficiency of our approximation method. We compute the Lévy density, the CDF and the prices of various options for the approximating processes and investigate their convergence. In Section 5, we compare our approach with the methods developed in [19] and [11] and we discuss connections with meromorphic processes. For the reader's convenience, in the Appendix we collect some results from the theory of Padé approximations, Stieltjes functions, Gaussian quadrature and orthogonal polynomials, which are used elsewhere in this paper.

Khintchine formula and the inverse Laplace transform. Show that if $\lim_{z\to\infty} f(z) = \lambda < \infty$, then Y must be a compound Poisson process with jump intensity λ. In particular, if $\lim_{z\to\infty} f(z) = 0$ then $Y = 0$ almost surely.

2. Main results. We begin by introducing a number of key definitions and notations. Let X be a Lévy process, and let $\psi(z) := \ln \mathbb{E}[e^{zX_1}]$ denote its *Laplace exponent*, which is initially defined on the vertical line $z \in \mathbb{C}$, $\mathrm{Re}(z) = 0$. The Lévy–Khintchine formula states that

$$\psi(z) = \sigma^2 z^2/2 + az + \int_{\mathbb{R}} \bigl(e^{zx} - 1 - zh(x)\bigr)\Pi(\mathrm{d}x), \tag{2}$$

where $\sigma \geq 0$, $a \in \mathbb{R}$, the Lévy measure $\Pi(\mathrm{d}x)$ satisfies $\int_{\mathbb{R}}(1 \wedge x^2)\Pi(\mathrm{d}x) < \infty$, and $h(x)$ is the *cutoff function*, which is required to ensure the convergence of the integral. Everywhere in this paper we will work under the following assumption.

ASSUMPTION 1. The Lévy measure $\Pi(\mathrm{d}x)$ is absolutely continuous, and its density $\pi(x)$ decreases exponentially fast as $x \to \pm\infty$.

If the cutoff function is fixed [the classical choice is $h(x) \equiv x\mathbf{1}_{\{|x|<1\}}$], then the process X is completely characterized by the triple (a, σ^2, π), which determines the Laplace exponent in (2). It is often convenient, however, to use different cutoff functions depending on the situation. Everywhere in this paper we will follow the convention that if the process X has jumps of finite variation, we will take $h(x) \equiv 0$, otherwise we will set $h(x) \equiv x$ (which is a legitimate choice due to Assumption 1). To distinguish between these two cases we will write the characteristic triple as $(a, \sigma^2, \pi)_{h\equiv 0}$ in the former case and $(a, \sigma^2, \pi)_{h\equiv x}$ in the latter case.

We recall that a function $f : (0, \infty) \mapsto \mathbb{R}$ is called *completely monotone* if $(-1)^k f^{(k)}(x) \geq 0$ for all $k = 0, 1, 2, \ldots$ and $x > 0$.

DEFINITION 1. We say that the process X has *completely monotone jumps* if the functions $\pi(x)$ and $\pi(-x)$ are completely monotone for $x \in (0, \infty)$.

Using Bernstein's theorem (see [31], page 3), we can express the above condition in an equivalent form: X has completely monotone jumps if and only if there exists a positive Radon measure μ, with support in $\mathbb{R} \setminus \{0\}$, such that for all $x \in \mathbb{R}$

$$\pi(x) = \mathbf{1}_{\{x>0\}} \int_{(0,\infty)} e^{-ux}\mu(\mathrm{d}u) + \mathbf{1}_{\{x<0\}} \int_{(-\infty,0)} e^{-ux}\mu(\mathrm{d}u). \tag{3}$$

For our further results, we will need the following two facts [which follow easily from (3) by Fubini's theorem]:

$$\int_{\mathbb{R}} x^2\pi(x)\,\mathrm{d}x < \infty \quad \text{if and only if} \quad \int_{\mathbb{R}} |u|^{-3}\mu(\mathrm{d}u) < \infty, \tag{4}$$

$$\int_{\mathbb{R}} |x|\pi(x)\,\mathrm{d}x < \infty \quad \text{if and only if} \quad \int_{\mathbb{R}} u^{-2}\mu(\mathrm{d}u) < \infty. \tag{5}$$

Condition (4) is required to ensure that the function $\pi(x)$ can be considered as a Lévy density, while the stronger condition (5) ensures that the resulting Lévy process X has jumps of finite variation.

Assuming that the Lévy density $\pi(x)$ is given by (3), we denote

$$\rho := \sup\left\{c \geq 0 : \int_{\mathbb{R}^+} e^{cx}\pi(x)\,\mathrm{d}x < \infty\right\} = \sup\{u \geq 0 : \mu\big((0,u)\big) = 0\},$$

$$\hat{\rho} := \sup\left\{c \geq 0 : \int_{\mathbb{R}^-} e^{-cx}\pi(x)\,\mathrm{d}x < \infty\right\} = \sup\{u \geq 0 : \mu\big((-u,0)\big) = 0\}.$$

Assumption 1 implies that $\rho > 0$ and $\hat{\rho} > 0$. We will denote by $\mathcal{CM}(\hat{\rho}, \rho)$ the class of Lévy processes with completely monotone jumps and parameters ρ and $\hat{\rho}$ defined as above.

Now we consider an important subclass of $\mathcal{CM}(\hat{\rho}, \rho)$.

DEFINITION 2. We say that the process X has *hyperexponential jumps* if the support of the measure $\mu(\mathrm{d}x)$ in (3) consists of finitely many points.

Let us consider a hyperexponential process X. According to Definition 2, the measure μ has finite support, which we will denote $\operatorname{supp}(\mu) = \{\hat{\beta}_i\}_{1\leq i\leq \hat{N}} \cup \{\beta_i\}_{1\leq i\leq N}$, where $\hat{N} \geq 0$ and $N \geq 0$, and $\hat{\beta}_i < 0$ and $\beta_i > 0$. We denote $\mu(\{\hat{\beta}_i\}) = \hat{\alpha}_i$ and $\mu(\{\beta_i\}) = \alpha_i$. Then the Lévy density of X can be represented in the form

$$\pi(x) = \mathbf{1}_{\{x>0\}} \sum_{i=1}^{N} \alpha_i e^{-\beta_i x} + \mathbf{1}_{\{x<0\}} \sum_{i=1}^{\hat{N}} \hat{\alpha}_i e^{-\hat{\beta}_i x}, \tag{6}$$

where one of the sums can be empty (if $\hat{N} = 0$ or $N = 0$). Formula (6) provides another equivalent definition of a hyperexponential process, as having positive/negative jumps equal in law to a finite mixture of exponential distributions.

DEFINITION 3. Let f be a function with a power series representation $f(z) = \sum_{i=0}^{\infty} c_i (z-a)^i$. If there exist polynomials $P_m(z)$ and $Q_n(z)$ satisfying $\deg(P) \leq m$, $\deg(Q) \leq n$, $Q_n(a) \neq 0$ and

$$\frac{P_m(z)}{Q_n(z)} = f(z) + O\big((z-a)^{m+n+1}\big), \qquad z \to a,$$

then we say that $f^{[m/n]}(z) := P_m(z)/Q_n(z)$ is the $[m/n]$ *Padé approximant of* f *at point* a.

Everywhere in this paper we will consider the case when the power series representation for $f(z)$ is convergent in some neighborhood of a (more generally, it can also be considered as a formal power series). When $a = 0$, we will call $f^{[m/n]}(z)$ simply the $[m/n]$ Padé approximation of $f(z)$, without mentioning the reference point.

2.1. *Approximating Lévy processes with two-sided jumps.* For a Lévy process $X \in \mathcal{CM}(\hat{\rho}, \rho)$, we define

$$\mu^*(A) = \mu\big(\{v \in \mathbb{R} : v^{-1} \in A\}\big), \tag{7}$$

for all Borel sets $A \subset \mathbb{R}$, where the measure $\mu(\mathrm{d}v)$ appears in (3). Note that $\operatorname{supp}(\mu^*) \subseteq [-1/\hat{\rho}, 1/\rho]$, and if the measure $\mu(\mathrm{d}v)$ is absolutely continuous with a density $m(v)$, then $\mu^*(\mathrm{d}v)$ also has a density, given by $m^*(v) = m(1/v)/v^2$. The measure $\mu^*(\mathrm{d}v)$ will play a very important role in this paper.

LEMMA 1. *Assume that $X \in \mathcal{CM}(\hat{\rho}, \rho)$. Then*

$$\int_{[-1/\hat{\rho}, 1/\rho]} |v|^3 \mu^*(\mathrm{d}v) < \infty,$$

and

$$\int_{[-1/\hat{\rho}, 1/\rho]} v^2 \mu^*(\mathrm{d}v) < \infty \quad \textit{if and only if} \quad X \textit{ has jumps of finite variation.}$$

PROOF. The result follows from (4) and (5) by change of variables $u = 1/v$. □

Now we are ready to introduce our first approximation. We start with a Lévy process $X \in \mathcal{CM}(\hat{\rho}, \rho)$ defined by the characteristic triple $(a, 0, \pi)_{h \equiv x}$. Note that the process X has zero Gaussian component. However, there is no lack of generality in assuming this: if we know how to approximate a Lévy processes with zero Gaussian component, we know how to approximate a general Lévy process, as we can always add a scaled Brownian motion to our hyperexponential approximation.

According to Lemma 1, $|v|^3 \mu^*(\mathrm{d}v)$ is a finite measure on the interval $[-1/\hat{\rho}, 1/\rho]$. Let $\{x_i\}_{1 \le i \le n}$ and $\{w_i\}_{1 \le i \le n}$ be the nodes and the weights of the Gaussian quadrature of order n with respect to this measure (we have included the definition and several key properties of the Gaussian quadrature in the Appendix. We define

$$\psi_n(z) := az + z^2 \sum_{i=1}^{n} \frac{w_i}{1 - z x_i}. \tag{8}$$

THEOREM 1. (i) *The function $\psi_n(z)$ is the $[n + 1/n]$ Padé approximant of $\psi(z)$.*

(ii) *The function $\psi_n(z)$ is the Laplace exponent of a hyperexponential process $X^{(n)}$ with the characteristic triple $(a, \sigma_n^2, \pi_n)_{h \equiv x}$, where*

$$\sigma_n^2 := \begin{cases} 0, & \text{if } x_i \neq 0 \text{ for all } 1 \le i \le n, \\ 2w_j, & \text{if } x_j = 0 \text{ for some } 1 \le j \le n, \end{cases} \tag{9}$$

and

$$\pi_n(x) := \begin{cases} \displaystyle\sum_{1\le i\le n\,:\,x_i<0} w_i |x_i|^{-3} e^{-x/x_i}, & \text{if } x<0, \\ \displaystyle\sum_{1\le i\le n\,:\,x_i>0} w_i x_i^{-3} e^{-x/x_i}, & \text{if } x>0. \end{cases} \tag{10}$$

If one of the sums in (10) *is empty, it should be interpreted as zero.*

(iii) *The random variables* $X_1^{(n)}$ *and* X_1 *satisfy* $\mathbb{E}[(X_1^{(n)})^j] = \mathbb{E}[(X_1)^j]$ *for* $1 \le j \le 2n+1$.

PROOF. Our first goal is to establish an integral representation of $\psi(z)$ in terms of the measure $\mu^*(\mathrm{d}v)$. Assume that $z \in \mathbb{C}$ with $-\hat{\rho} < \mathrm{Re}(z) < \rho$. We substitute (3) into (2), use Fubini's theorem to interchange the order of integration and obtain

$$\psi(z) = az + z^2 \int_{\mathbb{R}} \frac{\mathrm{sign}(u)}{u-z} \frac{\mu(\mathrm{d}u)}{u^2}. \tag{11}$$

Changing the variable $v = u^{-1}$ in the above integral and using the fact that $\mu((-\hat{\rho}, \rho)) = 0$, we obtain

$$\psi(z) = az + z^2 \int_{[-1/\hat{\rho}, 1/\rho]} \frac{|v|^3 \mu^*(\mathrm{d}v)}{1-vz}, \qquad -\hat{\rho} < \mathrm{Re}(z) < \rho. \tag{12}$$

By analytic continuation, we can see that the above formula is valid in a larger region $\mathbb{C} \setminus \{(-\infty, -\hat{\rho}] \cup [\rho, \infty)\}$.

Let us prove (i). By definition, the Gaussian quadrature of order n is exact for polynomials of degree not greater than $2n-1$, therefore,

$$\int_{[-1/\hat{\rho}, 1/\rho]} v^k |v|^3 \mu^*(\mathrm{d}v) = \sum_{i=1}^n x_i^k w_i, \qquad k = 0, 1, 2, \ldots, 2n-1.$$

The above identity is equivalent to

$$\left(\frac{\mathrm{d}^k}{\mathrm{d}z^k} \int_{[-1/\hat{\rho}, 1/\rho]} \frac{|v|^3 \mu^*(\mathrm{d}v)}{1-vz} \right)\Bigg|_{z=0} = \left(\frac{\mathrm{d}^k}{\mathrm{d}z^k} \sum_{i=1}^n \frac{w_i}{1-zx_i} \right)\Bigg|_{z=0}, \tag{13}$$

for $k = 0, 1, 2, \ldots, 2n-1$. Formulas (8), (12) and (13) imply that

$$\psi^{(k)}(0) = \psi_n^{(k)}(0), \qquad k = 0, 1, 2, \ldots, 2n+1. \tag{14}$$

By definition (8), $\psi_n(z)$ is a rational function, which can be written in the form $P(z)/Q(z)$ with $\deg(P) \le n+1$ and $\deg(Q) = n$. Using this fact and formula (14), we see that $\psi_n(z) \equiv \psi^{[n+1/n]}(z)$, which proves (i).

From (2) and (6), we see that the Laplace exponent of a hyperexponential process Y having triple $(a, 0, \pi)_{h\equiv x}$ is given by

$$\psi_Y(z) = az + z^2 \sum_{i=1}^{\hat{N}} \frac{\hat{\alpha}_i}{|\hat{\beta}_i|^3(1 - z/\hat{\beta}_i)} + z^2 \sum_{i=1}^{N} \frac{\alpha_i}{\beta_i^3(1 - z/\beta_i)}.$$

The result of item (ii) follows at once by comparing the above expression with (8).

Now that we have established that $\psi_n(z)$ is the Laplace exponent of a hyperexponential process $X^{(n)}$, formula (14) shows that the first $2n + 1$ cumulants of $X^{(n)}$ are equal to the corresponding cumulants of X_1, which is equivalent to the equality of corresponding moments and proves item (iii). □

The next important question that we need to address is how fast the approximations $\psi_n(z)$ converge to $\psi(z)$. As we have seen in the proof of Theorem 1 (see also [26, 30]), the Laplace exponent $\psi(z)$ of a process $X \in \mathcal{CM}(\hat{\rho}, \rho)$ is analytic in the cut complex plane $\mathbb{C} \setminus \{(-\infty, -\hat{\rho}] \cup [\rho, \infty)\}$. As we will establish in the next theorem, $\psi_n(z)$ converge to $\psi(z)$ *everywhere* in this region, and the convergence is exponentially fast on compact subsets of $\mathbb{C} \setminus \{(-\infty, -\hat{\rho}] \cup [\rho, \infty)\}$. This behavior should be compared with Taylor approximations, which can converge only in a circle of finite radius [lying entirely in the region of analyticity of $\psi(z)$]. This demonstrates that Padé approximations are very well suited to approximate Laplace exponents of processes in $\mathcal{CM}(\hat{\rho}, \rho)$.

THEOREM 2. *For any compact set* $A \subset \mathbb{C} \setminus \{(-\infty, -\hat{\rho}] \cup [\rho, \infty)\}$, *there exist* $c_1 = c_1(A) > 0$ *and* $c_2 = c_2(A) > 0$ *such that for all* $z \in A$ *and all* $n \geq 1$

$$\left|\psi_n(z) - \psi(z)\right| < c_1 e^{-c_2 n}.$$

Before we can prove Theorem 2, we need to present some auxiliary definitions related to Stieltjes functions. In the Appendix, we collect several relevant results which show the connections between Stieltjes functions and Padé approximations.

DEFINITION 4. A *Stieltjes function* is defined by the Stieltjes-integral representation

$$f(z) := \int_{[0,\infty)} \frac{\nu(\mathrm{d}u)}{1 + zu},$$

where $\nu(\mathrm{d}u)$ is a positive measure on $[0, \infty)$ whose support has infinitely many different points, and which has finite moments

$$m_j := \int_0^\infty u^j \nu(\mathrm{d}u).$$

Formally, we may also express f as a *Stieltjes series*, which may converge only at 0, and has the following form:

$$f(z) = \sum_{j=0}^{\infty} (-z)^j m_j. \tag{15}$$

It is easy to see that the above series converges for $|z| < R$ if and only if $\operatorname{supp}(\nu) \subseteq [0, 1/R]$. In this case we will call $f(z)$ a Stieltjes function (or a Stieltjes series) with the radius of convergence R.

PROOF OF THEOREM 2. Let us denote $\eta(\mathrm{d}v) = |v|^3 \mu^*(\mathrm{d}v)$ and define

$$g(z) := \int_{(0,1/\hat{\rho}+1/\rho]} \frac{\eta(\mathrm{d}(u - 1/\hat{\rho}))}{1 + uz}$$

and $f(z) := zg(z)$. Note that $g(z)$ is a Stieltjes function with the radius of convergence $R = (1/\rho + 1/\hat{\rho})^{-1}$, therefore, according to Theorem 10 in the Appendix, the Padé approximations $g^{[n-1/n]}(z)$ converge to $g(z)$ exponentially fast on compact subsets of $\mathbb{C} \setminus (-\infty, -R]$.

Changing the variable of integration $v = u - 1/\hat{\rho}$ in (12), we obtain

$$\begin{aligned} \psi(z) &= az - zf\left(-\frac{z}{1 + z/\hat{\rho}}\right) \\ &= az + \frac{z^2}{1 + z/\hat{\rho}} g\left(-\frac{z}{1 + z/\hat{\rho}}\right). \end{aligned} \tag{16}$$

According to Theorem 7 in the Appendix, the $[n/n]$ Padé approximation is invariant under rational transformations of the variable. Therefore, if $w = -z/(1 + z/\hat{\rho})$ and $F(z) := f(w)$ then $F^{[n/n]}(z) = f^{[n/n]}(w)$. Theorem 9 in the Appendix shows that $f^{[n/n]}(z) = zg^{[n-1/n]}(z)$. Using these results, formula (16) and the fact that $\psi_n(z) = \psi^{[n+1/n]}(z)$ which was established in Theorem 1, we conclude that

$$\begin{aligned} \psi_n(z) = \psi^{[n+1,n]}(z) &= az - zf^{[n/n]}\left(-\frac{z}{1 + z/\hat{\rho}}\right) \\ &= az + \frac{z^2}{1 + z/\hat{\rho}} g^{[n-1/n]}\left(-\frac{z}{1 + z/\hat{\rho}}\right). \end{aligned} \tag{17}$$

As we have noted above, the functions $g^{[n-1/n]}(z)$ converge to $g(z)$ exponentially fast on compact subsets of $\mathbb{C} \setminus (-\infty, -R]$, and it is easy to see that the function $w(z) = -z/(1 + z/\hat{\rho})$ maps compact subsets of $\mathbb{C} \setminus \{(-\infty, -\hat{\rho}] \cup [\rho, \infty)\}$ onto compact subsets of $\mathbb{C} \setminus (-\infty, -R]$. This fact combined with (16) and (17) completes the proof of Theorem 2. □

The results of Theorem 1 show that the Padé approximant $\psi^{[n+1/n]}(z)$ is always a Laplace exponent of a hyperexponential process. However, as we have discussed

in the Introduction (see the discussion on page 330), there are two other Padé approximants, $\psi^{[n/n]}(z)$ and $\psi^{[n+2/n]}(z)$, which can qualify as Laplace exponents. While we do not have a counterexample, we believe that in general it is not true that for all Lévy processes $X \in \mathcal{CM}(\hat{\rho}, \rho)$ the functions $\psi^{[n/n]}(z)$ and $\psi^{[n+2/n]}(z)$ are Laplace exponents of hyperexponential processes. However, more can be said under the additional assumption that the process has one-sided jumps, and we present these results in the next section.

2.2. *Approximating Lévy processes with one-sided jumps.* In this section, we will consider separately two cases: when the process X has (i) jumps of finite variation or (ii) jumps of infinite variation. In the first case, it is enough to consider subordinators with zero linear drift (if we know how to approximate such subordinators, we can always add a linear drift and a Gaussian component later). Thus, we assume that $X \in \mathcal{CM}(+\infty, \rho)$ is a subordinator with zero linear drift, defined by the characteristic triple $(0, 0, \pi)_{h\equiv 0}$, or, equivalently, by the Laplace exponent

$$\psi(z) = \ln \mathbb{E}\big[e^{zX_1}\big] = \int_0^\infty (e^{zx} - 1)\pi(x)\,\mathrm{d}x. \tag{18}$$

We emphasize that while our definition of the Laplace exponent of a subordinator is consistent with (2), it differs from the classical definition $\phi(z) := -\ln \mathbb{E}[\exp(-zX_1)]$ (see [4, 27]). The justification for this choice comes from the need to have a consistent notation for all Lévy processes under consideration: later we will be approximating subordinators with spectrally positive processes, and the formulas would be very confusing if we have different notations for the Laplace exponents of these two objects.

THEOREM 3. *Assume that $X \in \mathcal{CM}(+\infty, \rho)$ is a subordinator defined by the characteristic triple $(0, 0, \pi)_{h\equiv 0}$. Let $\psi(z)$ denote the Laplace exponent of X, given by (18). Fix $k \in \{0, 1, 2\}$.*

(i) *Let $\{x_i\}_{1\le i\le n}$ and $\{w_i\}_{1\le i\le n}$ be the nodes and the weights of the Gaussian quadrature with respect to the measure $v^{2+k}\mu^*(\mathrm{d}v)$. Then*

$$\psi^{[n+k/n]}(z) = \sum_{j=1}^{k} \psi^{(j)}(0)\frac{z^j}{j!} + z^{k+1}\sum_{i=1}^{n} \frac{w_i}{1 - zx_i}. \tag{19}$$

(ii) *The function $\psi^{[n+k/n]}(z)$ is the Laplace exponent of a hyperexponential process $X^{(n)}$. The process $X^{(n)}$ has a Lévy measure with density,*

$$\pi_n(x) := \mathbf{1}_{\{x>0\}} \sum_{i=1}^{n} w_i x_i^{-2-k} e^{-x/x_i}, \tag{20}$$

and is defined by the characteristic triple

$$(21)\qquad \begin{cases} (0,0,\pi_n)_{h\equiv 0}, & \text{if } k=0,\\ \left(\psi'(0)-\sum_{i=1}^{n} w_i/x_i, 0, \pi_n\right)_{h\equiv 0}, & \text{if } k=1,\\ \left(\psi'(0), \psi''(0)-2\sum_{i=1}^{n} w_i/x_i, \pi_n\right)_{h\equiv x}, & \text{if } k=2. \end{cases}$$

The process $X^{(n)}$ is a subordinator if $k=0$ or $k=1$ (with zero linear drift in the former case and positive linear drift in the latter case), and $X^{(n)}$ is a spectrally positive process with a nonzero Gaussian component if $k=2$.

(iii) *The functions $\psi^{[n+k/n]}(z)$ converge to $\psi(z)$ exponentially fast on compact subsets of $\mathbb{C}\setminus[\rho,\infty)$.*

Before proving Theorem 3, we need to establish the following auxiliary result.

LEMMA 2. *Assume that $\nu(\mathrm{d}x)$ is a finite positive measure on $(0,R]$. Let $\{x_i\}_{1\le i\le n}$ and $\{w_i\}_{1\le i\le n}$ be the nodes and the weights of the Gaussian quadrature with respect to the measure $x\nu(\mathrm{d}x)$ on $(0,R]$. Then*

$$\sum_{i=1}^{n} w_i/x_i < \int_{(0,R]} \nu(\mathrm{d}x).$$

PROOF. Consider two Stieltjes functions

$$f(z) := \int_{(0,R]} \frac{\nu(\mathrm{d}x)}{1+xz}, \qquad g(z) := \int_{(0,R]} \frac{x\nu(\mathrm{d}x)}{1+xz}.$$

It is easy to check that $f(z) = f(0) - zg(z)$. From Theorems 6 and 9 in the Appendix, we find that $f^{[n/n]}(z) = f(0) - zg^{[n-1/n]}(z)$ and $g^{[n-1/n]}(z) = \sum_{1\le i\le n} w_i/(1+x_iz)$. Therefore,

$$(22)\qquad \lim_{z\to+\infty} f^{[n/n]}(z) = f(0) - \sum_{i=1}^{n} w_i/x_i.$$

Consider the function $F(z) := (f(0)/f(z)-1)/z$. Note that $F(z)\to -f'(0)/f(0)$ as $z\to 0$, and that $F(z)$ is analytic in some neighborhood of zero. From Theorems 8 and 9 in the Appendix, we obtain

$$F^{[n-1/n]}(z) = \frac{1}{z}\left(\frac{m_0}{f^{[n/n]}(z)} - 1\right),$$

which can be rewritten as

$$(23)\qquad f^{[n/n]}(z) = m_0/\left(1+zF^{[n-1/n]}(z)\right).$$

Theorem 1.3 in [20] tells us that $F(z)$ is also a Stieltjes function, and since it is analytic in a neighborhood of zero, it has a positive radius of convergence (and therefore, finite moments). Theorem 6 in the Appendix implies that $\lim_{z\to+\infty} zF^{[n-1/n]}(z)$ is finite and positive. This fact combined with (23) shows that $\lim_{z\to+\infty} f^{[n/n]}(z)$ is strictly positive, and applying (22) we obtain the statement of the lemma. □

PROOF OF THEOREM 3. First, we note that since the process X has jumps of finite variation, Lemma 1 ensures that $v^2\mu^*(\mathrm{d}v)$ is a finite measure. Formulas (3) and (18) give us

$$\psi(z) = z\int_{(0,1/\rho]} \frac{v^2\mu^*(\mathrm{d}v)}{1-vz}. \tag{24}$$

We will prove the case $k=2$, as the other two cases can be treated in the same way. We start with the identity (24) and rewrite it in the equivalent form

$$\begin{aligned}\psi(z) &= z\int_{(0,1/\rho]} v^2\mu^*(\mathrm{d}v) + z^2\int_{(0,1/\rho]} v^3\mu^*(\mathrm{d}v) + z^3\int_{(0,1/\rho]} \frac{v^4\mu^*(\mathrm{d}v)}{1-vz}\\ &= \psi'(0)z + \psi''(0)\frac{z^2}{2} + z^3\int_{(0,1/\rho]} \frac{v^4\mu^*(\mathrm{d}v)}{1-vz}.\end{aligned}$$

The result of item (i) follows from the above expression and Theorems 6 and 9 in the Appendix.

Let us prove (ii). We use Lemma 2, from which it follows that

$$\frac{1}{2}\psi''(0) - \sum_{i=1}^n w_i/x_i = \int_{(0,1/\rho]} v^3\mu^*(\mathrm{d}v) - \sum_{i=1}^n w_i/x_i > 0,$$

and thus the coefficient of the Gaussian component is positive. Using (2), we compute the Laplace exponent of the process $X^{(n)}$ corresponding to the characteristic triple $(\psi'(0), \psi''(0) - 2\sum_{i=1}^n w_i/x_i, \pi_n)_{h\equiv x}$:

$$\begin{aligned}\psi_{X^{(n)}}(z) &= \left(\psi''(0) - \sum_{i=1}^n \frac{w_i}{x_i}\right)\frac{z^2}{2} + \psi'(0)z + z^2\sum_{i=1}^n \frac{w_i}{x_i(1-zx_i)}\\ &= \psi'(0)z + \psi''(0)\frac{z^2}{2} + z^3\sum_{i=1}^n \frac{w_i}{1-zx_i} = \psi^{[n+2/n]}(z),\end{aligned}$$

which proves (ii). Item (iii) follows from (24) and Theorem 10 in the Appendix. □

Now we consider the second class of processes with one-sided jumps: spectrally positive Lévy processes with jumps of infinite variation. Again, without loss of generality we assume that there is no Gaussian component. Our results are presented in the following theorem (the proof is omitted, as it is identical to the proof of Theorem 3).

THEOREM 4. *Assume that $X \in \mathcal{CM}(+\infty, \rho)$ is a spectrally positive process having jumps of infinite variation and defined by the characteristic triple $(a, 0, \pi)_{h\equiv x}$. Let $\psi(z)$ be its Laplace exponent defined by (2). Fix $k \in \{1, 2\}$.*

(i) *Let $\{x_i\}_{1\le i\le n}$ and $\{w_i\}_{1\le i\le n}$ be the nodes and the weights of the Gaussian quadrature with respect to the measure $v^{2+k}\mu^*(\mathrm{d}v)$. Then*

$$\psi^{[n+k/n]}(z) = \sum_{j=1}^{k} \psi^{(j)}(0)\frac{z^j}{j!} + z^{k+1}\sum_{i=1}^{n} \frac{w_i}{1 - zx_i}. \tag{25}$$

(ii) *The function $\psi^{[n+k/n]}(z)$ is the Laplace exponent of a hyperexponential process $X^{(n)}$. The process $X^{(n)}$ has a Lévy measure with density,*

$$\pi(x) := \mathbf{1}_{\{x>0\}} \sum_{i=1}^{n} w_i x_i^{-2-k} e^{-x/x_i},$$

and is defined by the characteristic triple

$$\begin{cases} (\psi'(0), 0, \pi)_{h\equiv x}, & \text{if } k = 1, \\ \left(\psi'(0), \psi''(0) - 2\sum_{i=1}^{n} w_i/x_i, \pi\right)_{h\equiv x}, & \text{if } k = 2. \end{cases} \tag{26}$$

(iii) *The functions $\psi^{[n+k/n]}(z)$ converge to $\psi(z)$ exponentially fast on compact subsets of $\mathbb{C} \setminus [\rho, \infty)$.*

REMARK. Let us explain why we have three different approximations in the case of subordinators and only two approximations in the case of spectrally positive processes. For a spectrally positive process with jumps of infinite variation, the measure $v^2\mu^*(\mathrm{d}v)$ is not finite (see Lemma 1), thus we cannot define Gaussian quadrature with respect to this measure and our method of proving that $\psi^{[n/n]}(z)$ is a Laplace exponent (in Theorem 3) will not work. While we do not have a counterexample, we believe that it is not true that for any spectrally positive process X with completely monotone jumps [and Laplace exponent $\psi(z)$] the function $\psi^{[n/n]}(z)$ is a Laplace exponent of a hyperexponential process.

3. Explicit examples and extensions of the algorithm. In this section, we pursue three goals. First, we will show how the results of Theorems 3 and 4 can lead to explicit formulas in the case of Gamma subordinators and one-sided tempered stable processes. Then we use these results to construct explicit hyperexponential approximations to VG, CGMY and NIG processes. Finally, we discuss several extensions of the approximation technique described in the previous section.

The *Jacobi polynomials* $P_n^{(\alpha,\beta)}(x)$ will play an important role in this section. They are defined as follows:

$$P_n^{(\alpha,\beta)}(x) := \sum_{j=0}^{n} \binom{\alpha+n}{n-j} \binom{\alpha+\beta+n+j}{j} \left(\frac{x-1}{2}\right)^j. \tag{27}$$

When $\alpha > -1$ and $\beta > -1$, these polynomials satisfy the orthogonality condition

$$\begin{aligned} &\int_{-1}^{1} P_n^{(\alpha,\beta)}(x) P_m^{(\alpha,\beta)}(x)(1-x)^\alpha (1+x)^\beta \,\mathrm{d}x \\ &\quad = \frac{2^{\alpha+\beta+1}}{2n+\alpha+\beta+1} \frac{\Gamma(n+\alpha+1)}{\Gamma(n+\beta+1)} \Gamma(n+\alpha+\beta+1) n! \delta_{n,m}. \end{aligned} \tag{28}$$

See Section 8.96 in [16] for other results related to Jacobi polynomials.

3.1. *Example* 1: *Gamma subordinator.* Consider a Gamma process X with both mean rate and the variance rate equal to one. In other words, X is a subordinator with zero linear drift, which has Lévy density $\pi(x) = x^{-1}\exp(-x)$ for $x > 0$, and Laplace exponent $\psi(z) = -\ln(1-z)$ [recall that we are using (18) as the definition of the Laplace exponent of a subordinator]. The random variable X_t has Gamma distribution

$$\mathbb{P}(X_t \in \mathrm{d}x) = \frac{1}{\Gamma(t)} x^{t-1} e^{-x} \,\mathrm{d}x, \qquad x > 0.$$

The following proposition gives explicit results for the approximations to X, described in Theorem 3.

PROPOSITION 1. *Let X be a Gamma process defined by the Laplace exponent $\psi(z) = -\ln(1-z)$. Fix $k \in \{0, 1, 2\}$.*

(i) *The denominators of the Padé approximants $\psi^{[n+k/k]}(z) = p_{n,k}(z)/q_{n,k}(z)$ are given by*

$$q_{n,k}(z) = z^n P_n^{(0,k)}(2/z - 1). \tag{29}$$

In the case $k = 0$, the numerators are also given by an explicit formula

$$p_{n,0}(z) = 2 \sum_{j=0}^{n} \binom{n}{j}^2 [H_{n-j} - H_j](1-z)^j, \tag{30}$$

where $H_0 := 0$ and $H_j := 1 + 1/2 + \cdots + 1/j$ for $j \geq 1$.

(ii) *The nodes of the Gaussian quadrature described in Theorem 3 are given by $x_i = (y_i + 1)/2$, where $y_i \in (-1, 1)$ are the roots of the Jacobi polynomials $P_n^{(0,k)}(y)$.*

PROOF. We check that

$$-\ln(1-z) = z\int_0^1 \frac{\mathrm{d}v}{1-zv},$$

and comparing the above result with formula (24) we identity $v^2\mu^*(\mathrm{d}v) = \mathrm{d}v$, which is just the Lebesgue measure on $(0,1)$. The orthogonal polynomials with respect to the measure $\mathbf{1}_{\{0<v<1\}}v^k\,\mathrm{d}v$ are given by the shifted Jacobi polynomials $P_n^{(0,k)}(2z-1)$. Formula (29) follows from this fact and Theorems 3 and 6. Statement (ii) follows from the well-known fact that the nodes of the Gaussian quadrature coincide with the roots of orthogonal polynomials (see the Appendix).

Using an equivalent representation for the Jacobi polynomials (see formula (8.960.1) in [16])

$$P_n^{(\alpha,\beta)}(x) = \sum_{j=0}^{n}\binom{\alpha+n}{n-j}\binom{\beta+n}{j}\left(\frac{x-1}{2}\right)^j\left(\frac{x+1}{2}\right)^{n-j},$$

we find that

$$q_{n,k}(z) = \sum_{j=0}^{n}\binom{k+n}{n-j}\binom{n}{j}(1-z)^j.$$

The above result and formula (5) in [33] give us the explicit expression for $p_{n,0}(z)$ in (30). □

3.2. *Example* 2: *Tempered stable subordinator/spectrally positive process*. Consider a Lévy process X defined by the Laplace exponent

$$\psi(z) = \Gamma(-\alpha)\big((1-z)^\alpha - 1\big), \tag{31}$$

where $\alpha \in (0,1)\cup(1,2)$. It is known (see formula (4.30) in [10]) that the Lévy density of the process X is given by

$$\pi(x) = \mathbf{1}_{\{x>0\}}x^{-1-\alpha}e^{-x}. \tag{32}$$

When $\alpha \in (0,1)$ then X is a subordinator with zero linear drift, and when $\alpha \in (1,2)$ then X is a spectrally positive process with jumps of infinite variation and zero Gaussian component.

PROPOSITION 2. *Let X be a tempered stable process defined by the Laplace exponent* (31). *For $\alpha \in (0,1)$ ($\alpha \in (1,2)$) we fix a value of $k \in \{0,1,2\}$ (resp., $k \in \{1,2\}$).*

(i) *The denominators and the numerators of the Padé approximants $\psi^{[n+k/k]}(z) = p_{n,k}(z)/q_{n,k}(z)$ are given by*

$$q_{n,k}(z) = z^n P_n^{(\alpha,k-\alpha)}(2/z-1), \tag{33}$$

$$p_{n,k}(z) = \Gamma(-\alpha)\left[\frac{1}{n!}\sum_{j=0}^{n+k}\frac{(2n+k-j)!(-n-\alpha)_j}{j!(n+k-j)!}z^j - q_{n,k}(z)\right]. \tag{34}$$

(ii) *The nodes of the Gaussian quadratures described in Theorems 3 and 4 are given by* $x_i = (y_i + 1)/2$, *where* $y_i \in (-1, 1)$ *are the roots of the Jacobi polynomials* $P_n^{(\alpha,k-\alpha)}(y)$.

PROOF. We check that for $x > 0$ and $\alpha > 0$

$$x^{-1-\alpha}e^{-x} = \frac{1}{\Gamma(1+\alpha)}\int_1^\infty e^{-ux}(u-1)^\alpha \,\mathrm{d}u.$$

The above result combined with formulas (3), (7) and (32) gives us

$$\mu(\mathrm{d}u) \sim \mathbf{1}_{\{u>1\}}(u-1)^\alpha \,\mathrm{d}u \quad \text{and} \quad v^2\mu^*(\mathrm{d}v) \sim \mathbf{1}_{\{0<v<1\}}v^{-\alpha}(1-v)^\alpha \mathrm{d}v,$$

where the symbol "$\sim$" means "equal, up to a multiplicative constant." This shows that the orthogonal polynomials with respect to the measure $v^{2+k}\mu^*(\mathrm{d}v)$ are given by the shifted Jacobi polynomials $P_n^{(\alpha,k-\alpha)}(2z-1)$. Formula (33) follows from this fact and Theorems 3 and 6, and statement (ii) follows from the fact that the nodes of the Gaussian quadrature coincide with the roots of orthogonal polynomials (see the Appendix). Formula (34) follows from the last equation in [18], and the following fact: *if* $m \geq n \geq 1$ *and* $p(z)/q(z)$ *is the* $[m/n]$ *Padé approximant to* $f(z)$, *then* $a(p(z) - q(z))/q(z)$ *is the* $[m/n]$ approximant *to* $a(f(z) - 1)$. The above fact is easy to deduce from Definition 3. □

3.3. *Approximating VG, CGMY and NIG processes.* The results of Propositions 1 and 2 can be used to construct explicit approximations to VG, NIG and CGMY processes. There are two methods for doing this: (i) we can construct the process with two-sided jumps as a difference of processes with only positive jumps or (ii) we can express the process as a Brownian motion with drift, time-changed by a subordinator.

Let us describe the first approach using the example of the VG process X (see [29]). We will denote by $\Gamma(t; \mu, \nu)$ the Gamma process with mean rate μ and variance rate ν, defined by the Laplace exponent

$$\psi_\Gamma(z) = -\frac{\mu^2}{\nu}\ln\left(1 - \frac{\nu}{\mu}z\right).$$

The Variance Gamma process is defined as the Brownian motion with drift $\sigma W_t + \theta t$ subordinated by an independent Gamma process Y_t with mean rate one and variance rate ν. The Laplace exponent of X is given by

$$\psi_X(z) = \psi_\Gamma\left(\theta z + \frac{\sigma^2}{2}z^2\right) = -\frac{1}{\nu}\ln\left(1 - \nu\theta z - \nu\frac{\sigma^2}{2}z^2\right). \tag{35}$$

Define $\mu_p = \frac{1}{2}\sqrt{\theta^2 + 2\sigma^2/\nu} + \theta/2$ and $\mu_n = \mu_p - \theta$. The identity

$$-\frac{1}{\nu}\ln\left(1 - \nu\theta z - \nu\frac{\sigma^2}{2}z^2\right) = -\frac{1}{\nu}\ln(1 - \mu_p\nu z) - \frac{1}{\nu}\ln(1 + \mu_n\nu z) \tag{36}$$

allows us to write X as the difference of two independent Gamma subordinators

$$X_t = \Gamma\big(t; \mu_p, \mu_p^2 \nu\big) - \Gamma\big(t; \mu_n, \mu_n^2 \nu\big). \tag{37}$$

In order to approximate the VG process X by a hyperexponential process, we use Proposition 1 and approximate each Gamma process in (37) by a hyperexponential subordinator [equivalently, we approximate each logarithm in (36) by a rational function].

The same procedure works for CGMY processes. They are defined by the Laplace exponent

$$\psi_X(z) = C\Gamma(-Y)\big[(M-z)^Y - M^Y + (G+z)^Y - G^Y\big], \tag{38}$$

where $Y \in (0,1) \cup (1,2)$ and all remaining parameters are positive. We see that X can be obtained as a linear drift plus a difference of two scaled tempered stable processes with only positive jumps. Proposition 2 gives us an explicit approximation to the one-sided processes [equivalently, explicit approximations to each power function in (38)], and as a result we obtain explicit hyperexponential approximations to general two-sided CGMY processes.

The second procedure for obtaining explicit approximations uses the representation of the process as a Brownian motion with drift, time-changed by a subordinator Y. The main idea is that we approximate the subordinator Y by a hyperexponential subordinator $\tilde{Y}$, which we then use as a time-change process, instead of Y. The following proposition ensures that the resulting approximation is also hyperexponential.

PROPOSITION 3. *Assume that $\tilde{Y}$ is a hyperexponential subordinator and W is an independent Brownian motion. Then for all $\sigma > 0$ and $a \in \mathbb{R}$ the process $Z_t := \sigma W_{\tilde{Y}_t} + a\tilde{Y}_t$ is also hyperexponential.*

PROOF. Denote the Laplace exponent of $\tilde{Y}$ as $\psi_{\tilde{Y}}(z)$. Since $\tilde{Y}$ is hyperexponential, $\psi_{\tilde{Y}}(z)$ is a rational function. It is well known that the Laplace exponent of the subordinated process Z is given by $\psi_Z(z) = \psi_{\tilde{Y}}(\sigma^2 z^2/2 + az)$, therefore, it is also a rational function. Proposition 2.1 in [19] tells us that the process Z has completely monotone Lévy density. This fact and rationality of ψ_Z prove that Z is hyperexponential. □

As we have discussed above, the VG process can be obtained as a Brownian motion with drift, time-changed by a Gamma process Y. Proposition 1 gives us explicit hyperexponential approximations to the Gamma process Y, therefore, from Proposition 3 we obtain explicit hyperexponential approximations to the original VG process.

The same ideas can be applied to the NIG process (see Section 4.4.3 in [10]), which is defined as a Brownian motion with drift time-changed by an inverse

Gaussian subordinator Y, defined by Laplace exponent $\psi_Y(z) = (1 - \sqrt{1 - \kappa z})/\kappa$. Proposition 2 gives us explicit hyperexponential approximations to Y and, therefore, we obtain explicit hyperexponential approximations to the NIG process itself.

The approximations described above have a number of desirable features. They are quite explicit, and the nodes of Gaussian quadratures which are needed to compute the characteristic triples of the approximating processes are expressed in terms of the roots of Jacobi polynomials (for which there exist extensive tables, and which can also be computed very easily by numerical means). The first method, based on decomposing the process into a difference of one-sided processes, is also quite flexible: we are free to choose the degree of the Padé approximation for each one-sided process independent of another. This may be helpful in applications, such as when pricing down-and-out barrier options: we may want to approximate negative jumps more accurately than positive jumps. However, we would like to emphasize that these approximations are not optimal, in the sense of property (iii) in Theorem 1: the general method for approximating two-sided Lévy processes gives a hyperexponential process with a Laplace exponent of smaller degree (the degree of a rational function is defined as the maximum of the degree of the numerator and denominator), which matches more moments of the original process. In Section 4, we compare the numerical efficiency of these two methods.

3.4. *Extensions of the approximation algorithm.* There are two ways in which Theorems 1, 3 and 4 can be generalized. First, there is an almost trivial (but potentially useful) generalization, in that instead of considering the Padé approximation at 0, we can consider the Padé approximation centered at another point $a \in (-\hat{\rho}, \rho)$. Then the statements of Theorems 1, 2, 3 and 4 would still be true, provided that we replace the Padé approximation $\psi^{[n+k/n]}(z)$ (centered at 0) by $\psi^{[n+k/n]}(z) - \psi^{[n+k/n]}(a)$ (centered at a). This fact can be easily established using the Esscher transform, which maps a Lévy process $X \in \mathcal{CM}(\hat{\rho}, \rho)$ defined by Laplace exponent $\psi(z)$ into a process $\tilde{X} \in \mathcal{CM}(\hat{\rho} + a, \rho - a)$, defined by Laplace exponent $\tilde{\psi}(z) = \psi(a + z) - \psi(a)$.

The second generalization is that instead of a Padé approximation one can use a general rational interpolation, which can informally be defined as a *multi-point Padé approximation*; see [12]. The following algorithm describes how to approximate the Laplace exponent $\psi(z)$ of a Lévy process $X \in \mathcal{CM}(\hat{\rho}, \rho)$.

A general approximation algorithm:

(i) For $k \geq 1$, choose k distinct points $\{z_i\}_{1 \leq i \leq k}$ inside the interval $(-\hat{\rho}, \rho)$.

(ii) Choose nonnegative integers $\{\beta_i\}_{1 \leq i \leq k}$, such that $k + \sum_{i=1}^{k} \beta_i = 2n + 1$ for some integer n.

(iii) We want to find a rational function $\tilde{\psi}(z) = zP(z)/Q(z)$ with $\deg(P) \leq n$ and $\deg(Q) \leq n$ and which satisfies

$$\frac{\mathrm{d}^j}{\mathrm{d}z^j}\big(\tilde{\psi}(z)/z\big)\big|_{z=z_i} = \frac{\mathrm{d}^j}{\mathrm{d}z^j}\big(\psi(z)/z\big)\big|_{z=z_i}, \qquad 1 \leq i \leq k, 0 \leq j \leq \beta_i. \tag{39}$$

THEOREM 5. *Assume that $\psi(z)$ is the Laplace exponent of a Lévy process $X \in \mathcal{CM}(\hat{\rho}, \rho)$. There exists a unique rational function $\tilde{\psi}(z)$ which satisfies the conditions of item* (iii). *Moreover, $\tilde{\psi}(z)$ is the Laplace exponent of a hyperexponential process $\tilde{X} \in \mathcal{CM}(\hat{\rho}, \rho)$.*

PROOF. We recall that $f(z)$ is called a *Pick function* if $f(z)$ is analytic in the upper-half plane $\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ and satisfies $f(\mathbb{H}) \subseteq \mathbb{H}$. There exists a bijection between Lévy processes with completely monotone jumps and Pick functions: $X \in \mathcal{CM}(\hat{\rho}, \rho)$ if and only if $\psi(z)/z$ is a Pick function analytic in $\mathbb{C} \setminus \{(-\infty, -\hat{\rho}] \cup [\rho, \infty)\}$ (see [26], Theorem 5.1, or [30]). The result of Theorem 5 now follows easily from the fact that $\psi(z)/z$ is a Pick function and [12], Theorem 4, which guarantees the existence of a rational functions $\tilde{\psi}(z)$ satisfying conditions (39) and states that $\tilde{\psi}(z)/z$ is a also Pick function, which is analytic in $\mathbb{C} \setminus \{(-\infty, -\hat{\rho}] \cup [\rho, \infty)\}$. □

4. Numerical results. In this section we present a number of numerical experiments, which demonstrate the efficiency of our approximations. As a first example, we consider the Gamma process X defined by the Laplace exponent $\psi(z) = -\ln(1 - z)$. We compute the Lévy density $\pi_n(x)$ corresponding to the approximation $\psi^{[n/n]}(z)$, which is given explicitly in Proposition 1. The Lévy density of the Gamma process is given by $\pi(x) = \exp(-x)/x$, thus in order to avoid the singularity at $x = 0$ we compare the graphs of $x\pi(x) \equiv \exp(-x)$ and $x\pi_n(x)$. The results are presented on Figure 1. We see that even with a small value of $n = 5$ the

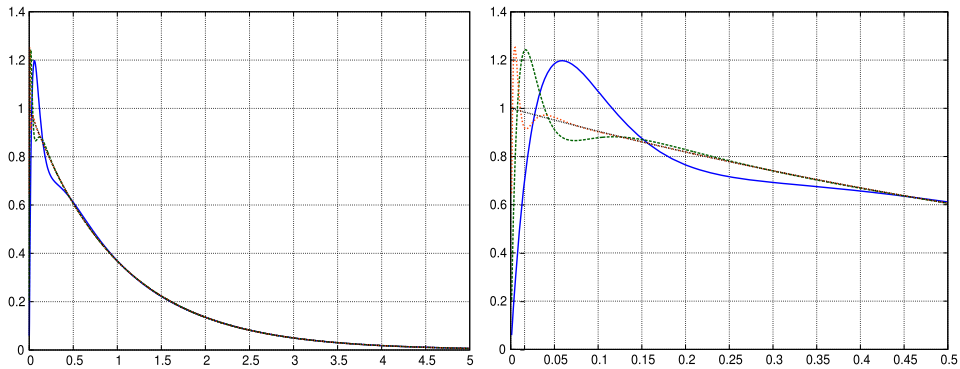

FIG. 1. *The graph of $x\pi(x)$ (black curve) and $x\pi^{[n/n]}(x)$, where $\pi(x) = x^{-1}\exp(-x)\mathbf{1}_{\{x>0\}}$ is the Lévy density of the Gamma subordinator, and $\pi^{[n/n]}(x)$ is the Lévy density corresponding to $\psi^{[n/n]}(z)$ Padé approximation, given by formula* (20). *Blue, green and red curves correspond to $n \in \{5, 10, 20\}$.* (b) *depicts the magnification by a factor of ten of the region near the origin in* (a).

tail of $\pi_n(x)$ matches the tail of $\pi(x)$ very well, and as n increases the approximation converges very rapidly (as long as x is not too close to zero).

Next, we compare the cumulative distribution function (CDF) of X_t for the same Gamma process X and its approximations $X^{(n,k)}$, which are defined by the Laplace exponents $\psi^{[n+k/n]}(z)$, $k \in \{0, 1, 2\}$, see Proposition 1. We compute the CDF for two values of $t \in \{1, 2\}$. The CDF of X_t for the Gamma process is known explicitly:

$$\mathbb{P}(X_1 \le x) = 1 - e^{-x} \quad \text{and} \quad \mathbb{P}(X_2 \le x) = 1 - (x+1)e^{-x}.$$

We define the numbers r_i as the coefficients in the asymptotic expansion

$$\psi^{[n+k/n]}(z) = r_2 z^2 + r_1 z + r_0 + O(1/z), \qquad z \to \infty,$$

and define

$$\phi_{n,k}(z) := \begin{cases} e^{t\psi^{[n/n]}(z)} - e^{\mathrm{tr}_0}, & \text{if } k = 0, \\ e^{t\psi^{[n+1/n]}(z)} - e^{\mathrm{tr}_0 + \mathrm{tr}_1 z}, & \text{if } k = 1, \\ e^{t\psi^{[n+2/n]}(z)}, & \text{if } k = 2. \end{cases}$$

The CDF of the approximating process is computed by the Fourier inversion

$$\begin{aligned} \mathbb{P}\big(X_t^{(n,k)} \le x\big) = 1 - \frac{e^{-cx}}{\pi} \mathrm{Re}\bigg[\int_0^\infty \phi_{n,k}(c + \mathrm{i}u) e^{-\mathrm{i}ux} \frac{\mathrm{d}u}{c + \mathrm{i}u}\bigg] \\ - e^{\mathrm{tr}_0} \mathbf{1}_{\{x \le r_1 t\}} \mathbf{1}_{\{k=1\}}, \end{aligned} \tag{40}$$

where $x > 0$ and $c \in (0, 1)$.

Let us explain the intuition behind the formula corresponding to $k = 0$, the other cases can be treated similarly. According to Theorem 3, the process $X^{(n,0)}$ is a compound Poisson hyperexponential process with intensity $-r_0$, thus its distribution has an atom at zero: $\mathbb{P}(X_t^{(n,0)} = 0) = \exp(\mathrm{tr}_0)$. If we subtract the atom at zero, we obtain an absolutely continuous positive measure

$$\nu_t(\mathrm{d}x) := \mathbb{P}\big(X_t^{(n,0)} \in \mathrm{d}x\big) - e^{\mathrm{tr}_0} \delta_0(\mathrm{d}x) \tag{41}$$

which has Fourier transform

$$\int_{\mathbb{R}} e^{\mathrm{i}tz} \nu_t(\mathrm{d}x) = e^{t\psi^{[n/n]}(z)} - e^{\mathrm{tr}_0} = \phi_{n,0}(z).$$

Since $\nu_t(\mathrm{d}x)$ is absolutely continuous with total mass $1 - \exp(\mathrm{tr}_0)$, we can find the CDF corresponding to this measure by the inverse Fourier transform

$$\nu_t\big((0, x)\big) = 1 - e^{\mathrm{tr}_0} - \frac{e^{-cx}}{\pi} \mathrm{Re}\bigg[\int_0^\infty \phi_{n,0}(c + \mathrm{i}u) e^{-\mathrm{i}ux} \frac{\mathrm{d}u}{c + \mathrm{i}u}\bigg]. \tag{42}$$

Note that the integral in (42) converges absolutely, since $\phi_{n,0}(c + \mathrm{i}u) = O(1/u)$ as $u \to \infty$. Formula (40) follows at once from (41) and (42).

TABLE 1
The values of $\varepsilon_{n,k}(t) := \max_{x\geq 0} |\mathbb{P}(X_t \leq x) - \mathbb{P}(X_t^{(n,k)} \leq x)|$, *where* X *is the Gamma process with* $\psi(z) = -\ln(1-z)$ *and the process* $X^{(n,k)}$ *has Laplace exponent* $\psi^{[n+k/n]}$

$\varepsilon_{n,k}(1)$	$k=0$	$k=1$	$k=2$	$\varepsilon_{n,k}(2)$	$k=0$	$k=1$	$k=2$
$n=5$	1.1e–2	1.1e–2	8.8e–3	$n=5$	3.3e–4	3.2e–4	5.4e–4
$n=10$	2.8e–3	3.4e–3	2.8e–3	$n=10$	2.6e–5	2.8e–5	5.6e–5
$n=15$	1.3e–3	1.6e–3	1.4e–3	$n=15$	5.4e–6	6.4e–6	1.3e–5
$n=20$	7.5e–4	9.3e–4	8.1e–4	$n=20$	1.8e–6	2.1e–6	4.6e–6

The results of our computations are presented in Table 1. We see that the CDF of $X_t^{(n,k)}$ does converge to X_t, and the convergence seems to be faster for $t=2$ than it is for $t=1$.

Our remaining examples are all related to pricing European and various exotic options in Lévy driven models. We will work with the following two processes: the VG process V defined by the Laplace exponent

$$\psi(z) = \mu z - \frac{1}{\nu}\ln\left(1-\frac{z}{a}\right) - \frac{1}{\nu}\ln\left(1+\frac{z}{\hat{a}}\right),$$

and parameters

$$(a, \hat{a}, \nu) = (21.8735, 56.4414, 0.20),$$

and the CGMY process Z defined by the Laplace exponent

$$\psi(z) = \mu z + C\Gamma(-Y)\big[(M-z)^Y - M^Y + (G+z)^Y - G^Y\big],$$

and parameters

$$(C, G, M, Y) = (1, 8.8, 14.5, 1.2).$$

Note that V is a process with jumps of infinite activity and finite variation, whereas Z has jumps of infinite variation. Both of these processes have zero Gaussian component. The process V with the same parameters was considered in [21], and later we will use their numerical results as a benchmark for our computations.

Our approach from here on is to compare a benchmark option price (for a variety of options) with a price calculated using one of four possible approximations. The first approximation is based on the $[n+1/n]$ Padé approximant for the process with two-sided jumps from Theorem 1. The other three approximations are based on the algorithm presented in Section 3.3, which considers the process as a difference of two processes having only positive jumps, and uses the explicit $[N+k/N]$ Padé approximations from Propositions 1 and 2. Note that the first approximation will result in a rational function of degree $n+1$, while the other three approximations result in a rational function of degree $2N+k$. In instances where we calculate multiple approximations, we set $n=2N$ in order to make a

TABLE 2
The error in computing the price of the European call option for the VG V-model. The benchmark price is 2.5002779303

	Two-sided	**One-sided**		
	$[2N+1/2N]$	$[N/N]$	$[N+1/N]$	$[N+2/N]$
$N=1$	−1.58e–2	9.12e–2	7.02e–3	−3.02e–2
$N=2$	1.66e–3	−6.16e–3	4.80e–3	−7.82e–4
$N=3$	6.20e–4	−1.28e–3	−4.32e–5	6.78e–4
$N=4$	1.25e–4	1.88e–4	−1.98e–4	9.81e–5
$N=5$	−7.19e–5	8.82e–5	−2.62e–5	−2.40e–5
$N=7$	4.34e–6	−8.48e–6	5.82e–6	−1.71e–6
$N=9$	−7.72e–8	3.31e–7	−6.99e–7	7.35e–7
$N=12$	4.85e–7	−1.81e–8	4.97e–8	−6.10e–8
$N=15$	−8.56e–8	−1.37e–9	−3.31e–9	6.06e–9

fair comparison between different approximations. In all examples, we define the stock price process as $S_t = S_0 \exp(X_t)$ (where $X \equiv V$ in the VG case or $X \equiv Z$ in the CGMY case). Further, we choose the value of the linear drift μ such that the process $S_t \exp(-rt)$ is a martingale.

First, we compute the price of a European call option with $S_0 = 100$, strike price $K = 100$, maturity $T = 0.25$ and interest rate $r = 0.04$. All option prices are computed using the Fourier transform approach from [9]. When dealing with hyperexponential processes, we have slightly modified this approach by removing the possible atom in the distribution of $X_t^{(n,k)}$, in the same way as we did earlier in equation (40). The benchmark prices for the original VG process V and the CGMY process Z were computed multiple times, with different discretizations of the Fourier integral, and seem to be correct to at least ±1.0e–9. The results of our computations for the approximations to VG (CGMY) model are presented in Table 2 (resp., 3). We see that all four approximations are doing an excellent job, and

TABLE 3
The error in computing the price of the European call option for the CGMY Z-model. The benchmark price is 11.9207826467

	Two-sided	**One-sided**	
	$[2N+1/2N]$	$[N+1/N]$	$[N+2/N]$
$N=1$	−2.75e–2	1.93e–2	−3.72e–3
$N=2$	−4.86e–6	−4.19e–6	9.5e–5
$N=3$	4.80e–7	−1.48e–5	−2.54e–7
$N=4$	2.9e–8	6.41e–7	−1.55e–7
$N=5$	1.14e–9	5.58e–9	6.95e–9

already for $N = 4$ we obtain acceptable accuracy of around 1.0e–4. We would like to point out that the three approximations based on explicit one-sided approximations have remarkably good accuracy. As we have discussed on page 345, these approximations are not optimal in the sense that one can find a rational Laplace exponent of lower degree which matches more moments of the original process. However, this nonoptimality does not seem to play any role here. These three one-sided approximations are superior to the two-sided approximation, in the sense that they have very good accuracy *and* explicit formulas.

We also note that all four approximations seem to be doing a better job in the case of the CGMY process Z. We think that the likely cause is that the process Z has jumps of infinite variation and Z_t has smooth density, which is not the case for the process V.

Next, we compute the price of a continuously sampled arithmetic Asian call option with fixed strike. That is, we calculate the following quantity:

$$C(S_0, K, T) := e^{-rT}\mathbb{E}\left[\left(\int_0^T S_u \,\mathrm{d}u - K\right)^+\right]. \tag{43}$$

We set the parameters as follows:

$$S_0 = 100, \qquad r = 0.03, \qquad T = 1$$

and $K = 90$ for the VG process and $K = 110$ for the CGMY process. In order to compute the price of the Asian option, we use the technique pioneered for hyperexponential processes by Cai and Kou [6] (see also [17], Section 4.2). Since we were unable to find any results in the literature for pricing such options for either the VG or CGMY process (other than by Monte Carlo methods), we use our own benchmark calculated using a significantly larger N. By experimenting with different ways of discretizing the resulting integrals in the inverse Laplace and inverse Mellin transform (see [17], Section 4.2), we arrive at a benchmark price of 11.18859 for the process V and 9.95930 for the process Z. These benchmark prices seem to be correct to within $\pm$1.0e–5. The results for each N are compared to the benchmark price, the errors are gathered in Table 4 for the process V and in Table 5 for the process Z.

We observe again, that convergence to the benchmark price is very rapid and that there is little difference in the rate of convergence between the one-sided and two-sided approximations. We note that we achieve an acceptable error of $\pm$1.0e–4 with a rational approximation of degree 5. We would like to emphasize that the numbers in Table 4 and Table 5 represent the difference between the approximate price and the benchmark price, and the benchmark itself is only an approximation to the exact price with accuracy of the order of $\pm$1.0e–5. The two most important factors influencing the accuracy of the benchmark price are (i) the accuracy of the approximation of the target Lévy process X by a hyperexponential process $X^{(n)}$, and (ii) the error in the discretization of the inverse Laplace and inverse Mellin transform needed to compute the price of Asian option (see [17], Section 4.2) in

TABLE 4
The error in computing the price of the Asian option for the VG V-model. The benchmark price is 11.188589 (*calculated using the* [91/90] *two-sided approximation*)

	Two-sided	**One-sided**		
	[2N + 1/2N]	**[N/N]**	**[N + 1/N]**	**[N + 2/N]**
$N = 1$	−1.87e–3	1.01e–3	−1.82e–3	9.88e–4
$N = 2$	9.49e–5	2.89e–4	−6.33e–5	3.27e–5
$N = 3$	1.30e–6	8.85e–6	−4.24e–6	3.99e–6
$N = 4$	−2.83e–6	1.07e–6	−1.36e–6	3.16e–7
$N = 5$	−1.11e–7	−2.48e–8	−5.91e–7	−3.81e–7

the model driven by the process $X^{(n)}$. The results presented above only measure the effect of the first of these two factors, which explains why some numbers are of the order ±1.0e–7 whereas our benchmark is only correct to within ±1.0e–5.

Our final example is related to pricing down-and-out barrier put option. That is, we wish to calculate

$$D(S_0, K, B, T) := e^{-rT}\mathbb{E}\big[(K - S_T)^+ \mathbf{1}_{\{S_t > B \text{ for } 0 \le t \le T\}}\big],$$

where B is the barrier level. We calculate barrier option prices for the process V, for four values $S_0 \in \{81, 91, 101, 111\}$ and with other parameters given by $K = 100$, $B = 80$, $r = 0.04879$ and $T = 0.5$. We use the prices computed in [21] as the benchmark (these prices seem to be accurate to about ±1.0e–3). In order to compute the prices of down-and-out put options for hyperexponential processes, we use the Laplace transform inversion method by Jeannin and Pistorius [19]. In this case, we present the results only for the one-sided $[N + 1/N]$ approximations. The results are presented in Table 6. We see that in almost all cases the convergence is very rapid, and we are able to match the first four digits of the benchmark price. The convergence is somewhat slower for $S_0 = 81$, which is to be expected: it is

TABLE 5
The error in computing the price of the Asian option for the CGMY Z-model. The benchmark price is 9.959300 (*calculated using the* [91/90] *two-sided approximation*)

	Two-sided	**One-sided**	
	[2N + 1/2N]	**[N + 1/N]**	**[N + 2/N]**
$N = 1$	1.88e–4	7.42e–4	−1.19e–3
$N = 2$	4.03e–6	9.05e–5	5.39e–6
$N = 3$	−3.58e–7	−2.64e–6	7.93e–8
$N = 4$	−3.88e–7	−1.01e–7	−1.21e–7
$N = 5$	−5.26e–7	−2.47e–7	−2.49e–7

TABLE 6
Barrier Option prices calculated for the VG process V-model. Benchmark prices were obtained from [21], *Table* 4, *column* 2

	$S_0 = 81$	$S_0 = 91$	$S_0 = 101$	$S_0 = 111$
Benchmark	**3.39880**	**7.38668**	**1.40351**	**0.04280**
$N = 2$	3.44551	7.39225	1.40527	0.04233
$N = 4$	3.40209	7.38957	1.40329	0.04258
$N = 6$	3.39910	7.38939	1.40332	0.04258
$N = 8$	3.39856	7.38936	1.40332	0.04258
$N = 10$	3.39853	7.38936	1.40332	0.04258

known that when pricing barrier options, the behavior of the price near the barrier is very sensitive to the nature of the small jumps of the underlying process (see [5]). Therefore, we may expect that our results will not be very precise when S_0 is close to B, since we are approximating a process with jumps of infinite activity by a compound Poisson process with drift.

Let us describe the computing environment used for our numerical experiments. The code was written in Fortran-90, and we used a standard 2011 laptop (with an Intel Core i5-2540M CPU). All Padé approximations were computed using the most basic algorithm based on solving the system of linear equations (47), which we describe below in the Appendix. Since this system of linear equations is typically ill-conditioned, all computations related to Padé approximations were performed with a high precision of 200 digits, using the MPFUN multiple precision package [2]. The computation time of the Padé approximation and its partial fraction decomposition was on the order of 0.1 seconds. Our goal in this section was to demonstrate the accuracy of Padé-based hyperexponential approximations, therefore, we did not try to write the most efficient code for computing the option prices. However, our computations were reasonably fast: the computation time for a single European (resp., Asian or barrier) option price was around 2 (resp., 5 or 15) seconds.

5. Concluding remarks. As we have mentioned in the Introduction, there exist other methods for approximating processes with completely monotone jumps by hyperexponential processes. The first of these was proposed by Jeannin and Pistorius in [19], and the second one by Crosby, Le Saux and Mijatović in [11]. Our research in this field was initially inspired by these two papers, and we would like to summarize their methods and highlight the similarities and differences with our method.

The approach of Jeannin and Pistorius is essentially based on minimizing the L_2 distance between the target Lévy density $\pi(x)$ and the approximating hyperexponential Lévy density $\pi_n(x)$. More precisely, we are looking for a hyperexponential

Lévy density $\pi_n(x)$ of the form (6) which minimizes

$$\Delta_{n,\varepsilon} = \int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]} \bigl(\pi(x) - \pi_n(x)\bigr)^2 \,\mathrm{d}x, \tag{44}$$

where $\pi(x)$ is the target Lévy density of a process with completely monotone jumps. Note that we do need to remove an ε-neighborhood of zero in the domain of integration in (44), because otherwise the integral may not converge. According to the definition of $\pi_n(x)$ in (6), the quantity $\Delta_{n,\varepsilon}$ can be considered as a function of $2N + 2\hat{N}$ parameters $\{\alpha_i, \beta_i : 1 \le i \le N\}$ and $\{\hat{\alpha}_i, \hat{\beta}_i : 1 \le i \le \hat{N}\}$, and ideally one would try to find the absolute minimum of this function in order to get the best fit of the hyperexponential Lévy density $\psi_n(x)$ to the target density $\pi(x)$. Since this optimal approach would result in a complicated nonlinear minimization problem, it is much easier to fix the parameters β_i and $\hat{\beta}_i$ [which specify the exponents of the exponential functions in (6)] and to minimize over the remaining parameters α_i and $\hat{\alpha}_i$. This simplification results in a simpler linear problem, which can be easily solved numerically.

Next, let us summarize the main ideas behind the method of Crosby, Le Saux and Mijatović [11]. We start with the Lévy process with completely monotone jumps and zero Gaussian component. We use formula (11), choose a parameter $A > 0$ large enough and derive the following approximation:

$$\begin{aligned}
\psi(z) &= az + z^2 \int_{\mathbb{R}} \frac{\operatorname{sign}(u)}{u - z} \frac{\mu(\mathrm{d}u)}{u^2} \\
&= az + z^2 \int_{\mathbb{R}\setminus[-A,A]} \frac{\operatorname{sign}(u)}{u - z} \frac{\mu(\mathrm{d}u)}{u^2} + z^2 \int_{[-A,A]} \frac{\operatorname{sign}(u)}{u - z} \frac{\mu(\mathrm{d}u)}{u^2} \\
&= az + z^2 \int_{\mathbb{R}\setminus[-A,A]} \frac{1}{1 - z/u} \frac{\mu(\mathrm{d}u)}{|u|^3} + z^2 \int_{[-A,A]} \frac{\operatorname{sign}(u)}{u - z} \frac{\mu(\mathrm{d}u)}{u^2} \\
&\approx az + z^2 \int_{\mathbb{R}\setminus[-A,A]} \frac{\mu(\mathrm{d}u)}{|u|^3} + z^2 \int_{[-A,A]} \frac{\operatorname{sign}(u)}{u - z} \frac{\mu(\mathrm{d}u)}{u^2} =: \tilde{\psi}(z),
\end{aligned}$$

where in the last step we used the fact that $|u| > A \gg 1$ and, therefore, $1 - z/u$ can be approximated by 1. The above approximation is the first step in the method of Crosby et al., and it gives us the Laplace exponent of a Lévy process $\tilde{X}$ with a small (but nonzero) Gaussian coefficient

$$\sigma^2 = 2 \int_{\mathbb{R}\setminus[-A,A]} \frac{\mu(\mathrm{d}u)}{|u|^3}.$$

The process $\tilde{X}$ has Lévy measure $\tilde{\pi}(x)$, given by (3) with $\mu(\mathrm{d}x)$ replaced by $\mu(\mathrm{d}x)\mathbf{1}_{\{|x|\le A\}}$. It is easy to see that $\tilde{\pi}(x)$ is a finite measure, thus $\tilde{X}$ has compound Poisson jumps. Intuitively, the effect of this first step is to replace the jumps of X (which could be of infinite activity or infinite variation) by compound Poisson

jumps and a small Gaussian component. The second step in the method of Crosby et al. consists in discretizing the integral

$$\int_{[-A,A]} \frac{\operatorname{sign}(u)}{u-z}\frac{\mu(\mathrm{d}u)}{u^2} \approx \sum \frac{\operatorname{sign}(x_i)}{x_i-z}\frac{w_i}{x_i^2}$$

via the Gauss–Legendre quadrature (a Gaussian quadrature on the interval $[-A, A]$ with respect to the Lebesgue measure). Combining these two steps results in a Laplace exponent of approximating hyperexponential process.

Our method is quite similar to the approach of Crosby, Le Saux and Mijatović. Instead of their first approximating step, we perform a change of variables $u = 1/v$ in the integral (11). This simple trick and Assumption 1 give us a finite domain of integration in the v-variable in (12), so that we can apply Gaussian quadrature with respect to the measure $|v|^3\mu^*(\mathrm{d}v)$. It turns out that this seemingly small modification has profound consequences. First of all, we do not need to truncate the integrals and we do not require any external parameters (such as ε or A in the above two methods). Second, our approximating Laplace exponents $\psi_n(z)$ have a simple analytic interpretation as Padé approximations of the target Laplace exponent $\psi(z)$, which allows us to borrow tools and ideas from the well developed theory of rational approximations and orthogonal polynomials. Third, our approximation turns out to be optimal in the sense that the hyperexponential process $X^{(n)}$ constructed in Theorem 1 matches $2n + 1$ moments of the target process X [see the statement of Theorem 1(iii)]. Note that this is the best that one can hope for: according to formula (8) the process $X^{(n)}$ has $2n + 1$ free parameters, thus we cannot expect to be able to match more than $2n + 1$ moments of X. Finally, we show in Theorem 2 that our approximations converge exponentially in n, where n is the number of terms in the Lévy density, and this fast convergence is confirmed by our numerical experiments.

In conclusion, we would like to discuss how our current results fit in the context of recent developments on meromorphic processes [22, 23, 25]. The main motivation for introducing meromorphic processes was the perceived lack of explicit examples of Lévy processes which would be useful for modeling purposes and convenient for numerical calculations. Meromorphic processes serve this purpose quite well: they are flexible enough to allow for jumps of infinite activity or infinite variation, they have many parameters and are very similar to the widely used CGMY and VG processes, and at the same time, they are analytically tractable and enjoy an explicit Wiener–Hopf factorization. A meromorphic process can be informally defined as a hyperexponential process with infinitely many terms in the Lévy density [so that the Lévy measure is given by (6) with the finite sum replaced by infinite series]. Hyperexponential processes can be considered as a subclass of meromorphic processes, in the same way that rational functions can be considered a subclass of meromorphic functions. This turns out to be a very useful analogy, and it seems that every formula related to hyperexponential processes has a corresponding analogue for meromorphic processes, with the only change that the

finite sums or products would be replaced by appropriate infinite series or products. While hyperexponential processes are much simpler objects to work with, compared with meromorphic processes, their big disadvantage is that they do not allow for jumps of infinite activity or infinite variation.

As an example of how our current work complements the previous developments on meromorphic processes, consider the following hypothetical situation. Suppose that we have data on European options for a certain stock and we want to price barrier options on the same underlying stock and we want to use the CGMY model to describe the stock price dynamics. We face a problem in that numerical computation of prices of barrier options is not so easy in the CGMY model: algorithms based on Monte Carlo technique are not very accurate and rather time consuming, whereas analytical methods [19] are not possible since we do not know the Wiener–Hopf factors of a CGMY process. One way to solve this problem would be to use meromorphic processes. We would just replace the family of CGMY processes by a very similar family of beta-processes [22], and then calibrate the parameters of a beta-process to the available data on European options and price barrier options in the model driven by a beta-process using the results of [22, 25]. Another way would be to adhere to our original choice of the CGMY model: we would calibrate the parameters of a CGMY process to the available data and then approximate the calibrated CGMY process with hyperexponential processes as described in this paper. The prices of barrier options for hyperexponential processes can be computed easily [19]. It is not clear which of these two approaches would be a better solution in practice. The first one requires that we abandon the CGMY model and instead use meromorphic processes, plus we have to be careful with truncating infinite products and sums when doing numerical computations for meromorphic processes. The second approach allows us to keep our favorite CGMY model and simplifies the numerical computations (dealing with hyperexponential processes is easier compared to meromorphic processes). The downside of the second approach is that we introduce a new source of error when we approximate a CGMY process by a hyperexponential process. Overall, we feel that both approaches have merit and that they deserve further investigation.

APPENDIX: GAUSSIAN QUADRATURE, PADÉ APPROXIMATIONS AND STIELTJES FUNCTIONS

Consider a finite positive measure $\nu(\mathrm{d}x)$ on an interval $[0, a]$. The main idea behind Gaussian quadrature is that we want to find a measure $\tilde{\nu}(\mathrm{d}x)$, supported on n points inside $[0, a]$, which matches the first $2n-1$ moments of $\nu(\mathrm{d}x)$. Thus, the weights $\{w_i\}_{1\le i\le n}$ and the nodes $\{x_i\}_{1\le i\le n}$ of the Gaussian quadrature are uniquely defined by equations

$$\int_{[0,a]} x^k \nu(\mathrm{d}x) = \sum_{i=1}^{n} x_i^k w_i, \qquad k = 0, 1, \ldots, 2n-1.$$

Let $\{p_n(x)\}_{n\geq 0}$ be the sequence of orthogonal polynomials with respect to the measure $\nu(\mathrm{d}x)$: $\deg(p_n) = n$ and $(p_n, p_m)_\nu := \int_{[0,a]} p_n(x)p_m(x)\nu(\mathrm{d}x) = d_n\delta_{n,m}$. It is known [32], Theorems 3.4.1 and 3.4.2, that the nodes $\{x_j\}_{1\leq j\leq n}$ of the Gaussian quadrature of order n are given by the zeros of the polynomial $p_n(x)$, and the weights are given by

$$w_j = \frac{a_n}{a_{n-1}} \frac{(p_{n-1}, p_{n-1})_\nu}{p_{n-1}(x_j)p_n'(x_j)}, \tag{45}$$

where a_k is the coefficient of x^k in $p_k(x)$.

The following result demonstrates close connections between Gaussian quadrature, orthogonal polynomials, Padé approximations and Stieltjes functions.

THEOREM 6 (Theorems 2.2 and 3.1 in [1]). *Consider a Stieltjes function*

$$f(z) := \int_{[0,a]} \frac{\nu(\mathrm{d}x)}{1+xz}.$$

Then

$$f^{[n-1/n]}(z) = \frac{(-z)^{n-1}q_{n-1}(-1/z)}{(-z)^n p_n(-1/z)} = \sum_{i=1}^{n} \frac{w_i}{1+x_i z}, \tag{46}$$

where $\{x_i\}_{1\leq i\leq n}$ and $\{w_i\}_{1\leq i\leq n}$ are the nodes and weights of the Gaussian quadrature with respect to the measure $\nu(\mathrm{d}x)$, $p_n(z)$ is the nth orthogonal polynomial with respect to ν and $q_{n-1}(z)$ is the associated polynomial of degree $n-1$, defined by

$$q_{n-1}(z) := \int_{[0,a]} \frac{p_n(z) - p_n(w)}{z-w} \nu(\mathrm{d}w).$$

Next, we will discuss how one can compute the coefficients of the Padé approximation. Consider a function $f(z)$ given by a formal series expansion $f(z) = \sum_{i\geq 0} c_i z^i$. Then the Padé approximation $f^{[m/n]}(z) = P_m(z)/Q_n(z)$ with $m \geq n$ can be found as follows (provided it exists): first, we solve the system of n linear equations

$$\begin{bmatrix} c_{m-n+1} & c_{m-n+2} & c_{m-n+3} & \cdots & c_m \\ c_{m-n+2} & c_{m-n+3} & c_{m-n+4} & \cdots & c_{m+1} \\ c_{m-n+3} & c_{m-n+4} & c_{m-n+5} & \cdots & c_{m+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_m & c_{m+1} & c_{m+2} & \cdots & c_{m+n-1} \end{bmatrix} \begin{bmatrix} b_n \\ b_{n-1} \\ b_{n-2} \\ \vdots \\ b_1 \end{bmatrix} = - \begin{bmatrix} c_{m+1} \\ c_{m+2} \\ c_{m+3} \\ \vdots \\ c_{m+n} \end{bmatrix} \tag{47}$$

and find b_i, $1 \leq i \leq n$. These coefficients give us the denominator $Q_n(z) := 1 + b_1 z + b_2 z^2 + \cdots + b_n z^n$. Then, the coefficients of the numerator $P_m(z) := a_0 +$

$a_1 z + a_2 z^2 + \cdots + a_m z^m$ can be calculated recursively

$$\begin{aligned} a_0 &= c_0, \\ a_1 &= c_1 + b_1 c_0, \\ a_2 &= c_2 + b_1 c_2 + b_2 c_0, \\ &\vdots \\ a_m &= c_m + \sum_{i=1}^{n} b_i c_{m-i}. \end{aligned} \tag{48}$$

In practice, when n is even moderately large, the system in (47) will have a very large condition number, and solving the system of linear equations (47) would involve a loss of accuracy. This can be avoided by using higher precision arithmetic. Another way to deal with this problem is to use expressions for Padé approximations given in terms of Gaussian quadrature [such as (8), (19) and (25)]. There exist several very fast and accurate methods for computing the weights and nodes of the Gaussian quadrature; see [14, 15].

Below we collect some other results on Padé approximations, which are used elsewhere in this paper.

THEOREM 7 (Theorem 1.5.2 in [3]). *Given a formal series* $f(z) = \sum_{i=0}^{\infty} c_i z^i$ *and* $a \neq 0$ *we define* $w = w(z) = az/(1 + bz)$ *and* $g(w) = f(z)$. *If the Padé approximant* $f^{[n/n]}(z)$ *exists, then* $g^{[n/n]}(w) = f^{[n/n]}(z)$.

THEOREM 8 (Theorem 1.5.3 in [3]). *Given a formal series* $f(z) = \sum_{i=0}^{\infty} c_i z^i$ *we define* $g(z) = (a + bf(z))/(c + df(z))$. *If* $c + df(0) \neq 0$ *and the Padé approximant* $f^{[n/n]}$ *exists, then*

$$g^{[n/n]}(z) = \frac{a + bf^{[n/n]}(z)}{c + df^{[n/n]}(z)}.$$

THEOREM 9 (Theorem 1.5.4 in [3]). *Assume that* $k \geq 1$ *and* n, m *are integers such that* $n - k \geq m - 1$. *Given a formal series* $f(z) = \sum_{i=0}^{\infty} c_i z^i$ *we define*

$$g(z) = \left(f(z) - \sum_{i=0}^{k-1} c_i z^i \right) z^{-k}.$$

Then

$$g^{[n-k/m]}(z) = \left(f^{[n/m]}(z) - \sum_{i=0}^{k-1} c_i z^i \right) z^{-k},$$

provided either Padé approximant exists.

THEOREM 10 (Theorem 5.4.4 in [3]). *Let $f(z)$ be a Stieltjes series with radius of convergence $R > 0$. Let A be a compact subset of $\mathbb{C} \setminus (-\infty, -R]$. Define δ to be the distance from A to the set $(-\infty, -R]$ and $\rho := R - \delta$. Then there exists a constant $C = C(A)$ such that for all $z \in A$ and all $n \geq 1$ we have*

$$\left| f(z) - f^{[n-1/n]}(z) \right| < C \left| \frac{\sqrt{\rho + z} - \sqrt{\rho}}{\sqrt{\rho + z} + \sqrt{\rho}} \right|^{2n}.$$

Acknowledgments. The authors would like to thank two anonymous referees for their careful reading of the paper and for suggesting several improvements.

REFERENCES

[1] ALLEN, G. D., CHUI, C. K., MADYCH, W. R., NARCOWICH, F. J. and SMITH, P. W. (1975). Padé approximation of Stieltjes series. *J. Approx. Theory* **14** 302–316. MR0382927

[2] BAILEY, D. H. (1995). A fortran-90 based multiprecision system. *ACM Trans. Math. Software* **21** 379–387.

[3] BAKER, G. A. JR. and GRAVES-MORRIS, P. (1996). *Padé Approximants*, 2nd ed. Cambridge Univ. Press, Cambridge. MR1383091

[4] BERTOIN, J. (1996). *Lévy Processes*. Cambridge Univ. Press, Cambridge. MR1406564

[5] BOYARCHENKO, M., DE INNOCENTIS, M. and LEVENDORSKIĬ, S. (2011). Prices of barrier and first-touch digital options in Lévy-driven models, near barrier. *Int. J. Theor. Appl. Finance* **14** 1045–1090. MR2864298

[6] CAI, N. and KOU, S. (2012). Pricing Asian options under a hyper-exponential jump diffusion model. *Oper. Res.* **60** 64–77. MR2911657

[7] CAI, N. and KOU, S. G. (2011). Option pricing under a mixed-exponential jump diffusion model. *Management Science* **57** 2067–2081.

[8] CARR, P., GEMAN, H., MADAN, D. B. and YOR, M. (2002). The fine structure of asset returns: An empirical investigation. *The Journal of Business* **75** 305–333.

[9] CARR, P. and MADAN, D. (1999). Option valuation using the fast Fourier transform. *J. Comput. Finance* **2** 61–73.

[10] CONT, R. and TANKOV, P. (2004). *Financial Modelling with Jump Processes*. Chapman & Hall, Boca Raton, FL. MR2042661

[11] CROSBY, J., LE SAUX, N. and MIJATOVIĆ, A. (2010). Approximating Lévy processes with a view to option pricing. *Int. J. Theor. Appl. Finance* **13** 63–91. MR2646974

[12] DONOGHUE, W. F. JR. (1974). The interpolation of Pick functions. *Rocky Mountain J. Math.* **4** 169–174.

[13] FOURATI, S. (2012). Explicit solutions of the exit problem for a class of Lévy processes; applications to the pricing of double-barrier options. *Stochastic Process. Appl.* **122** 1034–1067. MR2891446

[14] GAUTSCHI, W. (1970). On the construction of Gaussian quadrature rules from modified moments. *Math. Comp.* **24** 245–260. MR0285117

[15] GOLUB, G. H. and WELSCH, J. H. (1969). Calculation of Gauss quadrature rules. *Math. Comp.* **23** 221–230. MR0245201

[16] GRADSHTEYN, I. S. and RYZHIK, I. M. (2007). *Table of Integrals, Series, and Products*, 7th ed. Elsevier/Academic Press, Amsterdam. MR2360010

[17] HACKMANN, D. and KUZNETSOV, A. (2014). Asian options and meromorphic Lévy processes. *Finance Stoch.* **18** 825–844. MR3255753

[18] ISERLES, A. (1979). A note on Padé approximations and generalized hypergeometric functions. *BIT Numerical Mathematics* **19** 543–545. MR0559965

[19] JEANNIN, M. and PISTORIUS, M. (2010). A transform approach to compute prices and Greeks of barrier options driven by a class of Lévy processes. *Quant. Finance* **10** 629–644. MR2676789

[20] KALUGIN, G. A., JEFFREY, D. J., CORLESS, R. M. and BORWEIN, P. B. (2012). Stieltjes and other integral representations for functions of Lambert *W*. *Integral Transforms Spec. Funct.* **23** 581–593. MR2959457

[21] KUDRYAVTSEV, O. and LEVENDORSKIĬ, S. (2009). Fast and accurate pricing of barrier options under Lévy processes. *Finance Stoch.* **13** 531–562. MR2519843

[22] KUZNETSOV, A. (2010). Wiener–Hopf factorization and distribution of extrema for a family of Lévy processes. *Ann. Appl. Probab.* **20** 1801–1830. MR2724421

[23] KUZNETSOV, A. (2010). Wiener–Hopf factorization for a family of Lévy processes related to theta functions. *J. Appl. Probab.* **47** 1023–1033. MR2752893

[24] KUZNETSOV, A. (2012). On the distribution of exponential functionals for Lévy processes with jumps of rational transform. *Stochastic Process. Appl.* **122** 654–663. MR2868934

[25] KUZNETSOV, A., KYPRIANOU, A. E. and PARDO, J. C. (2012). Meromorphic Lévy processes and their fluctuation identities. *Ann. Appl. Probab.* **22** 1101–1135. MR2977987

[26] KWAŚNICKI, M. (2013). Rogers functions and fluctuation theory. Available at arXiv:1312.1866.

[27] KYPRIANOU, A. E. (2014). *Fluctuations of Lévy Processes with Applications*: *Introductory Lectures*, 2nd ed. Springer, Heidelberg. MR3155252

[28] LEWIS, A. L. and MORDECKI, E. (2008). Wiener–Hopf factorization for Lévy processes having positive jumps with rational transforms. *J. Appl. Probab.* **45** 118–134. MR2409315

[29] MADAN, D. B., CARR, P. P. and CHANG, E. C. (1998). The Variance Gamma process and option pricing. *European Finance Review* **2** 79–105.

[30] ROGERS, L. C. G. (1983). Wiener–Hopf factorization of diffusions and Lévy processes. *Proc. Lond. Math. Soc.* (3) **47** 177–191. MR0698932

[31] SCHILLING, R. L., SONG, R. and VONDRAČEK, Z. (2012). *Bernstein Functions*: *Theory and Applications*, 2nd ed. *De Gruyter Studies in Mathematics* **37**. de Gruyter, Berlin. MR2978140

[32] SZEGÖ, G. (1975). *Orthogonal Polynomials*, 4th ed. Amer. Math. Soc., Providence, RI. MR0372517

[33] WEIDEMAN, J. A. C. (2005). Padé approximations to the logarithm. I. Derivation via differential equations. *Quaest. Math.* **28** 375–390. MR2164379

DEPARTMENT OF MATHEMATICS AND STATISTICS
YORK UNIVERSITY
4700 KEELE STREET
TORONTO, ON, M3J 1P3
CANADA
E-MAIL: dhackman@mathstat.yorku.ca
kuznetsov@mathstat.yorku.ca
URL: www.danhackmann.com
www.math.yorku.ca/~akuznets