

# Splitting and matrix exponential approach for jump-diffusion models with Inverse Normal Gaussian, Hyperbolic and Meixner jumps

Andrey Itkin\*

*School of Engineering, New York University, Brooklyn, NY, USA*

**Abstract.** This paper is a further extension of the method proposed in Itkin (2014) as applied to another set of jump-diffusion models: Inverse Normal Gaussian, Hyperbolic and Meixner. To solve the corresponding PIDEs we accomplish few steps. First, a second-order operator splitting on financial processes (diffusion and jumps) is applied to these PIDEs. To solve the diffusion equation we use standard finite-difference methods. For the jump part, we transform the jump integral into a pseudo-differential operator and construct its second order approximation on a grid which supersedes the grid used for the diffusion part. The proposed schemes are unconditionally stable in time and preserve positivity of the solution which is computed either via a matrix exponential, or via its Páde approximation. Various numerical experiments are provided to justify these results.

**Keywords:** Jump-diffusion, PIDE, splitting, matrix exponential, unconditionally stable schemes

## 1. Introduction

This paper is a further extension of the method proposed in Itkin (2014) as applied to another set of jump-diffusion models: Inverse Normal Gaussian, Hyperbolic and Meixner. These models have been introduced in mathematical finance within last two decades, see Barndorff-Nielsen (1998), Eberlein and Keller (1995), Eberlein et al. (1998), Schoutens (2001). However, to the best of the author's knowledge, they seem to receive less attention of practitioners as compared, e.g., with the Merton, Kou and CGMY/KoBoL models, see Itkin (2014) for a short survey about the latter models and references therein. At the first glance this looks unfair, because typical distributions

of returns produced by the former models fit the observed market data (and, in particular, fat tails and skew) even better than their more popular counterparts. For instance, Fang and Oosterlee (2009) report that the NIG model returns more consistent measures from time to time and from one company to another as compared with the CGMY model.

One of the possible reasons could be that despite the pdf and the characteristic function (CF) of these models are known in closed form, and, therefore, pricing of plain vanilla and even American options could be done by transform methods (FFT, cosine, adaptive integration in the Fourier space), see Carr and Madan (1999), Fang and Oosterlee (2008), Lewis (2001), Lipton (2001), Lord et al. (2007), the analytic expressions for the pdf and CF are more complicated than that for their counterparts, and sometimes require usage of special functions. However, this doesn't prevent pricing and hedging of simple vanilla instruments to be efficiently done.

---

\*Corresponding author: Andrey Itkin, School of Engineering, New York University, 6 Metro Tech Center, RH 517E, Brooklyn, NY 11201, USA. Tel.: +646 855 3389; Fax: +718 260 3355; E-mail: aitkin@nyu.edu.

The second point is that the models considered in this paper are pure jump models that don't contain a diffusion component. This, however, could be easily relaxed.

On the other hand, nowadays practitioners want to have a model which is capable to simultaneously fit market data on both vanilla and exotic options. For doing that they need to consider some kind of a stochastic local volatility (LSV) model, or even a LSV model with jumps. Under these conditions usually neither the pdf, nor the CF are available in closed form. Therefore, efficient numerical methods should be used to solve the pricing equations which belong to the class of Partial Integro-Differential Equations (PIDE).

A number of methods were proposed to address such a construction, see Itkin (2014) and references therein as well as discussion of problems related to their implementation. In particular, they include a discretization of the PIDE that is implicit in the differential terms and explicit in the integral term (Cont and Voltchkova (2003)), Picard iterations for computing the integral equation (d'Halluin et al. (2005a, 2004)) and a second-order accurate, unconditionally-stable operator splitting (ADI) method that does not require an iterative solution of an algebraic equation at each time step (Andersen and Andreasen (2000)). Various forms of operator splitting technique were also used for this purpose (Itkin (2014), Itkin and Carr (2012)).

Assuming that an efficient discretization of the PIDE in time was properly chosen, the remaining problem is a fast computation of the jump integral, as it was observed to be relatively expensive. A short survey of various methods proposed in the literature including a review of their advantages and disadvantages again could be found in Itkin (2014). Also in that paper the original method proposed in Itkin and Carr (2012) was further elaborated on to exploit the following idea. First, an operator-splitting method on the financial processes is used, thus separating the computation of the diffusion part from the integral part. Then, similar to Itkin and Carr (2012), the jump integral is represented in the form of a pseudo-differential operator. Next thus obtained evolutionary partial pseudo-differential equations are formally solved via a matrix exponential.

In Itkin (2014) it was shown that the matrix exponential can be efficiently computed for Merton's, Kou's and CGMY models, and that the efficiency of this method is not worse than that of the FFT. It was also mentioned that the proposed method is almost universal, i.e., it could potentially be applicable in a

unified form to various jump-diffusion models. It is also important that this method is relatively simple for implementation.

In the present paper we apply this approach to the Inverse Normal Gaussian, Hyperbolic and Meixner models. We construct finite difference schemes to solve the corresponding PIDEs and prove that all the proposed schemes demonstrate the second order convergence in space and time and are unconditionally stable. The suggested approach is new and eliminates some known limitations of the existing methods, see discussion in Itkin (2014). Also, for the first time splitting and matrix exponential method is used as applied to the referenced jump models. This allows an efficient usage of these models in a more complex framework, e.g., the LSV model with jumps. Furthermore, the complexity of solving a pure jump evolutionary equation for the Meixner model using the new method is close to  $O(N)$  which is better than that of the FFT.

Finally, as the distributions underlying the corresponding Lévy processes are capable to fit well the market data, using these jump models together with the efficient numerical method for solving the jump-diffusion PIDEs potentially gives rise to a more efficient pricing and hedging of the derivative instruments.

Also to underline, despite the general idea of the method has been already described in Itkin (2014), Itkin and Carr (2012), constructing a particular discretization of the jump operators could be different for every model. This requires the corresponding Propositions to be proved in every case to demonstrate approximation, stability and complexity of the method. Therefore, these schemes, Propositions and proofs are the main contributions of this paper.

The rest of the paper is organized as follows. In section 2 we briefly discuss a general form of the backward PIDE for the class of Lévy models. In Section 3 we present our general approach to the solution of the PIDE using splitting and the matrix exponential approach. An explicit construction of various finite-difference schemes of the first and second order is presented in the next section. The results presented in that section are new, and to the best of our knowledge have not been discussed in the literature. Our technique utilizes some results from matrix analysis related to the notion of M-matrices, the Metzler matrices and eventually exponentially nonnegative matrices. We also provide results of various numerical tests to demonstrate convergence of our method. The final section concludes.

## 2. Lévy models and backward PIDE

To avoid uncertainty, let us consider the problem of pricing equity options written on a single stock. As we will see, this specification does not cause us to lose any generality, but it makes the description more practical. We assume an underlying asset (stock) price  $S_t$  being driven by an exponential of a Lévy process

$$S_t = S_0 \exp(L_t), \quad 0 \leq t \leq T, \quad (1)$$

where  $t$  is time,  $T$  is the option expiration,  $S_0 = S_t|_{t=0}$ ,  $L_t$  is the Lévy process  $L = (L_t)_{0 \leq t \leq T}$  with a nonzero Brownian (diffusion) part. Under the pricing measure,  $L_t$  is given by

$$L_t = \gamma t + \sigma W_t + Y_t, \quad \gamma, \sigma \in \mathbb{R}, \quad \sigma > 0, \quad (2)$$

with Lévy triplet  $(\gamma, \sigma, \nu)$ , where  $W_t$  is a standard Brownian motion on  $0 \leq t \leq T$  and  $Y_t$  is a pure jump process. We consider this process under the pricing measure, therefore,  $e^{-(r-q)t} S_t$  is a martingale, where  $r$  is the interest rate and  $q$  is a continuous dividend. This allows us to express  $\gamma$  as (Eberlein (2009))

$$\gamma = r - q - \frac{\sigma^2}{2} - \int_{\mathbb{R}} (e^x - 1 - x \mathbf{1}_{|x| < 1}) \nu(dx),$$

where  $\nu(dx)$  is a Lévy measure which satisfies

$$\int_{|x| > 1} e^x \nu(dx) < \infty.$$

The jump measure  $\nu(dx)$  yet is left unspecified, because we are open to consider all types of jumps including those with finite and infinite variation, and finite and infinite activity.<sup>1</sup>

To price options written on the underlying process  $S_t$ , we want to derive a PIDE that describes time evolution of the European option prices  $C(x, t)$ ,  $x \equiv \log(S_t/S_0)$ . Using a standard martingale approach, or by creating a self-financing portfolio, one can derive the corresponding PIDE (Cont and Tankov (2004))

$$\begin{aligned} rC(x, t) &= \frac{\partial C(x, t)}{\partial t} + \left(r - \frac{1}{2}\sigma^2\right) \frac{\partial C(x, t)}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 C(x, t)}{\partial x^2} \\ &+ \int_{\mathbb{R}} \left[ C(x+y, t) - C(x, t) - (e^y - 1) \frac{\partial C(x, t)}{\partial x} \right] \nu(dy) \end{aligned} \quad (3)$$

for all  $(x, t) \in \mathbb{R} \times (0, T)$ , subject to the terminal condition

$$C(x, T) = h(x), \quad (4)$$

where  $h(x)$  is the option payoff, and some boundary conditions which depend on the type of the option. The solutions of this PIDE usually belong to the class of viscosity solutions (Cont and Tankov (2004)).

We now rewrite the integral term using the following idea. It is well known from quantum mechanics (de Lange and Raab (1992)) that a translation (shift) operator in  $L_2$  space could be represented as

$$\mathcal{T}_b = \exp\left(b \frac{\partial}{\partial x}\right), \quad (5)$$

with  $b = \text{const}$ , so

$$\mathcal{T}_b f(x) = f(x + b).$$

Therefore, the integral in Equation (3) can be formally rewritten as

$$\begin{aligned} \int_{\mathbb{R}} \left[ C(x+y, t) - C(x, t) - (e^y - 1) \frac{\partial C(x, t)}{\partial x} \right] \nu(dy) &= \mathcal{J}C(x, t), \end{aligned} \quad (6)$$

$$\mathcal{J} \equiv \int_{\mathbb{R}} \left[ \exp\left(y \frac{\partial}{\partial x}\right) - 1 - (e^y - 1) \frac{\partial}{\partial x} \right] \nu(dy).$$

In the definition of operator  $\mathcal{J}$  (which is actually an infinitesimal generator of the jump process), the integral can be formally computed under some mild assumptions about existence and convergence if one treats the term  $\partial/\partial x$  as a constant. Therefore, operator  $\mathcal{J}$  can be considered as some generalized function of the differential operator  $\partial_x$ . We can also treat  $\mathcal{J}$  as a pseudo-differential operator.

For the future, an important observation is that

$$\mathcal{J} = \phi(-i\partial_x) \quad (7)$$

where  $\phi(u)$  is the characteristic exponent of the jump process. This directly follows from the

<sup>1</sup>We recall that a standard Brownian motion already has paths of infinite variation. Therefore, the Lévy process in Equation (2) has infinite variation since it contains a continuous martingale component. However, here we refer to the infinite variation that comes from the jumps.

Lévy-Khinchine theorem. Note, that thus defined characteristic exponent is computed using the expectation under a risk-neutral measure. In other words, the last term in the definition of  $\mathcal{J}$  is a compensator which is added to make the forward price to be a true martingale under this measure.

With allowance for this representation, the whole PIDE in the Equation (3) can be rewritten in operator form as

$$\partial_\tau C(x, \tau) = [\mathcal{D} + \mathcal{J}]C(x, \tau), \quad (8)$$

where  $\tau = T - t$  and  $\mathcal{D}$  represents a differential (parabolic) operator

$$\mathcal{D} \equiv -r + \left(r - \frac{1}{2}\sigma^2\right) \frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2}, \quad (9)$$

where the operator  $\mathcal{D}$  is an infinitesimal generator of diffusion.

Notice that for jumps with finite variation and finite activity, the last two terms in the definition of the jump integral  $\mathcal{J}$  in Equation (3) could be integrated out and added to the definition of  $\mathcal{D}$ . In the case of jumps with finite variation and infinite activity, the last term could be integrated out. However, here we will leave these terms under the integral for two reasons: i) this transformation (moving some terms under the integral to the diffusion operator) does not affect our method of computation of the integral, and ii) adding these terms to the operator  $\mathcal{D}$  could potentially negatively influence the stability of the finite-difference scheme used to solve the parabolic equation  $\mathcal{D}C(x, t) = 0$ . This equation naturally appears as a part of our splitting method, which is discussed in the next section.

### 3. Operator splitting and solution of jump equations

To solve Equation (8) we use splitting. This technique is also known as the method of fractional steps (see Dyakonov (1964), Samarski (1964), Yanenko (1971)) and sometimes is cited in financial literature as Russian splitting or locally one-dimensional schemes (LOD) (Duffy (2006)).

Below we follow Itkin (2014) to give a short survey of this technique as applied to linear and nonlinear PDEs.

The method of fractional steps reduces the solution of the original  $k$ -dimensional unsteady problem to

the solution of  $k$  one-dimensional equations per time step. For example, consider a two-dimensional diffusion equation with a solution obtained by using some finite-difference method. At every time step, a standard discretization on space variables is applied, such that the finite-difference grid contains  $N_1$  nodes in the first dimension and  $N_2$  nodes in the second dimension. Then the problem is solving a system of  $N_1 \times N_2$  linear equations, and the matrix of this system is block-diagonal. In contrast, utilization of splitting results in, e.g.,  $N_1$  systems of  $N_2$  linear equations, where the matrix of each system is banded (tridiagonal). The latter approach is easy to implement and, more importantly, provides significantly better performance.

The previous procedure uses operator splitting in different dimensions. Marchuk (1975) and then Strang (1968) extended this idea for complex physical processes (for instance, diffusion in the chemically reacting gas, or the advection-diffusion problem). In addition to (or instead of) splitting on spatial coordinates, they also proposed splitting the equation into physical processes that differ in nature, for instance, convection and diffusion. This idea becomes especially efficient if the characteristic times of evolution (relaxation time) of such processes are significantly different.

For the PIDE in Equation (3) we use a version of splitting described in Itkin (2014) which gives rise to the following numerical scheme:

$$\begin{aligned} C^{(1)}(x, \tau) &= e^{\frac{\Delta\tau}{2}\mathcal{D}}C(x, \tau), \\ C^{(2)}(x, \tau) &= e^{\Delta\tau\mathcal{J}}C^{(1)}(x, \tau), \\ C(x, \tau + \Delta\tau) &= e^{\frac{\Delta\tau}{2}\mathcal{D}}C^{(2)}(x, \tau). \end{aligned} \quad (10)$$

Thus, instead of an unsteady PIDE, we obtain one PIDE with no drift and diffusion (the second equation in Equation (10)) and two unsteady PDEs (the first and third ones in Equation (10)).

In what follows, we consider how to efficiently solve the second equation, while assuming that the solution of the first and the third equations can be obtained using any finite-difference method that is sufficiently efficient. To this end, in various examples given in the next sections we will explicitly mention what particular method was used for this purpose.

To solve the second (pure jump) evolutionary equation we again follow the method of Itkin (2014). By definition of the jump generator  $\mathcal{J}$ , under some mild constraints on its existence,  $\mathcal{J}$  could be viewed as a function of the operator  $\partial_x$ . Therefore, solving the

integral (second) equation in Equation (10) requires a few steps.

First, we construct an appropriate discrete grid  $\mathbf{G}(x)$  in the truncated (originally infinite) space domain. This grid could be nonuniform. An important point is that in the space domain where the parabolic equations of Equation (10) are defined, this grid should coincide with the finite-difference grid constructed for the solution of these parabolic equations. In other words, the PIDE grid is a superset of the PDE grid. This is useful to avoid interpolation of the solution that is obtained on the jump grid (the second step of the splitting algorithm) to the diffusion grid that is constructed to obtain solutions at the first and third splitting steps.

For the sake of concreteness let the parabolic equation be solved at the space domain  $[x_0, x_k]$ ,  $x_0 > -\infty$ ,  $x_k < \infty$  using a nonuniform grid with  $k + 1$  nodes  $(x_0, x_1, \dots, x_k)$  and space steps  $h_1 = x_1 - x_0, \dots, h_k = x_k - x_{k-1}$ . The particular choice of  $x_0$  and  $x_k$  is determined by the problem under consideration. We certainly want  $|x_0|$  and  $|x_k|$  not to be too large. The integration limits of  $\mathcal{J}$  in Equation (6) are, however, plus and minus infinity. Truncation of these limits usually is done to fit memory and performance requirements. On the other hand, we want a fine grid close to the option strike  $K$  for better accuracy. Therefore, a reasonable way to construct a jump grid is as follows. For  $x_0 \leq x \leq x_k$ , the jump grid coincides with the grid used for solution of the parabolic PDEs. Outside of this domain, the grid is expanded by adding nonuniform steps; i.e., the entire jump grid is  $x_{-m_1}, x_{1-m_1}, \dots, x_{-1}, x_0, x_1, \dots, x_k, x_{k+1}, \dots, x_{k+m_2}$ . Here  $m_1 > 0$ ,  $m_2 > 0$  are some integer numbers that are chosen based on our preferences. Since contribution to  $\mathcal{J}$  from very large values of  $x$  is negligible, the outer grid points  $x_{-m_1}, x_{1-m_1}, \dots, x_{-1}$  and  $x_{k+1}, \dots, x_{k+m_2}$  can be made highly nonuniform. One possible algorithm could be to have the steps of these grids be a geometric progression. This allows one to cover the truncated infinite interval with a reasonably small number of nodes.

Second, the discretization of  $\partial_x$  should be chosen on  $\mathbf{G}(x)$ . We want this discretization to:

1. Provide the necessary order of approximation of the whole operator  $\mathcal{J}$  in space.
2. Provide unconditional stability of the solution of the second equation in Equation (10).
3. Provide positivity of the solution.

Let  $\Delta_x$  denote a discrete analog of  $\partial_x$  obtained by discretization of  $\partial_x$  on the grid  $\mathbf{G}(x)$ . Accordingly, let us define the matrix  $J(\Delta_x)$  to be the discrete analog of the operator  $\mathcal{J}$  on the grid  $\mathbf{G}(x)$ . In Itkin (2014) the following proposition is proven that translates the above requirements to the conditions on  $J(\Delta_x)$ .

**Proposition 3.1.** *The finite-difference scheme*

$$C(x, \tau + \Delta\tau) = e^{\Delta\tau J(\Delta_x)} C(x, \tau) \quad (11)$$

is unconditionally stable in time  $\tau$  and preserves positivity of the vector  $C(x, \tau)$  if there exists an  $M$ -matrix  $B$  such that  $J(\Delta_x) = -B$ .

This proposition gives us a recipe for the construction of the appropriate discretization of the operator  $\mathcal{J}$ . In the next section, we will give some explicit examples of this approach.

Once the discretization is performed, all we need is to compute a matrix exponential  $e^{\Delta\tau J(\Delta_x)}$ , and then a product of this exponential with  $C(x, \tau)$ . The following facts make this method competitive with those briefly described in the introduction. We need to take into account that:

1. The matrix  $J(\Delta_x)$  can be precomputed once the finite-difference grid  $\mathbf{G}(x)$  has been built.
2. If a constant time step is used for computations, the matrix  $\mathcal{A} = e^{\Delta\tau J(\Delta_x)}$  can also be precomputed.

If the above two statements are true, the second splitting step results in computing a product of a matrix with time-independent entries and a vector. The complexity of this operation is  $O(N^2)$ , assuming the matrix  $\mathcal{A}$  is  $N \times N$ , and the vector is  $N \times 1$ . However,  $N$  in this case is relatively small (see some numerical examples and estimates in Itkin (2014)). Also the product  $\mathcal{A}C(x, \tau)$  can be computed using FFT, if at every time step one re-interpolates values from  $\mathbf{G}(x)$  to the FFT grid, similar to how this was done in d'Halluin et al. (2004). This reduces the total complexity to  $O(N \log_2 N)$ .

In some special cases (Merton's jump model, Kou model) the product  $\mathcal{A}C(x, \tau)$  could be computed with the complexity  $O(N)$  using some tricks proposed in Itkin (2014). We will further exploit this idea for some models described in this paper.

The above consideration is sufficiently general in the sense that it covers any particular jump model where

jumps are modeled as an exponential Lévy process. Below we review three jump models: Inverse Normal Gaussian, Hyperbolic and Meixner. Given a model, our goal is to construct a finite-difference scheme, first for  $\Delta_x$ , and then for  $J(\Delta_x)$ , that satisfies the conditions of Proposition 3. Again we want to emphasize that we discuss these jump models being a part of a more general either LV or LSV model with jumps. Otherwise, as the CF of the considered models are known in closed form, any FFT based method would be more efficient in, e.g., obtaining prices of the European vanilla options.

### 3.1. Normal Inverse Gaussian Model (NIG)

The NIG type Lévy process was introduced by Barndorff-Nielsen (1998) as a model for log returns of stock prices. It is a sub-class of the more general class of hyperbolic Lévy processes that will be discussed in the next section. Barndorff-Nielsen (1998) considered classes of normal variance-mean mixtures and defined the NIG distribution as the case when the mixing distribution is inverse Gaussian with the characteristic exponent

$$\begin{aligned} \phi_{NIG}(\alpha, \beta, \delta, u) \\ = \delta \left( \sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + iu)^2} \right). \end{aligned} \quad (12)$$

Therefore, the CF function reads

$$\begin{aligned} \Phi_{NIG}(\alpha, \beta, \delta, \mu, u) \\ = \exp \left\{ t\delta \left( \sqrt{\alpha^2 - \beta^2} \right. \right. \\ \left. \left. - \sqrt{\alpha^2 - (\beta + iu)^2} \right) + it\mu \right\}, \end{aligned} \quad (13)$$

where  $u \in \mathbb{R}$ ,  $\mu \in \mathbb{R}$ ,  $\delta > 0$ ,  $0 \leq |\beta| \leq \alpha$ . It is known that parameters of the NIG model play the following role for the underlying distribution:  $\alpha$  is responsible for the tail heaviness of steepness,  $\beta$  affects the symmetry,  $\delta$  scales the distribution, and  $\mu$  determines the mean value (location). It is also known that when using the NIG process for option pricing the location parameter of the distribution has no effect on the option value, so one can use  $\mu = 0$ . This, however, is not critical for the below approach and could be easily relaxed.

The linearity of the log of the characteristic function with respect to time shows that it is an infinitely divisible process with stationary independent increments.

The NIG process has no diffusion component making it a pure jump process with the Lévy density

$$\nu(dx) = \frac{2\alpha\delta \exp(\beta x) K_1(\alpha|x|)}{\pi |x|} dx, \quad (14)$$

where  $K_\lambda(z)$  is the modified Bessel function of the third kind.<sup>2</sup>

Next we need to substitute Equation (14) into the definition of the operator  $\mathcal{J}$  in Equation (6) and fulfill a formal integration to obtain the corresponding evolutionary pure jump equation in the explicit form. However, as it was mentioned earlier this step could be formalized by making use of Equation (7). Therefore, we immediately obtain

$$\mathcal{J} = \delta \left( \sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + \nabla)^2} \right), \quad (15)$$

where  $\nabla \equiv \partial/\partial x$ . The corresponding evolutionary pure jump equation to be solved is

$$\begin{aligned} C^{(2)}(x, \tau) &= \mathcal{A}C^{(1)}(x, \tau), \\ \mathcal{A} &= \exp \left\{ \Delta \tau \delta \left( \sqrt{\alpha^2 - \beta^2} \right. \right. \\ &\quad \left. \left. - \sqrt{\alpha^2 - (\beta + \nabla)^2} \right) \right\}. \end{aligned} \quad (16)$$

Before constructing a finite difference scheme to solve this equation we need to introduce some definitions. Define a one-sided *forward* discretization of  $\nabla$ , which we denote as  $A^F$ :  $A^F C(x) = [C(x+h, t) - C(x, t)]/h$ . Also define a one-sided *backward* discretization of  $\nabla$ , denoted as  $A^B$ :  $A^B C(x) = [C(x, t) - C(x-h, t)]/h$ . These discretizations provide first order approximation in  $h$ , e.g.,  $\nabla C(x) = A^F C(x) + O(h)$ . To provide the second order approximations, use the following definitions. Define  $A_2^C = A^F A^B$  - the *central* difference approximation of the second derivative  $\nabla^2$ , and  $A^C = (A^F + A^B)/2$  - the *central* difference approximation of the first derivative  $\nabla$ . Also define a one-sided second order approximations to the first derivatives: *backward* approximation  $A_2^B$ :  $A_2^B C(x) = [3C(x) - 4C(x-h) + C(x-2h)]/(2h)$ , and *forward* approximation  $A_2^F$ :  $A_2^F C(x) = [-3C(x) + 4C(x+h) - C(x+2h)]/(2h)$ . Also let  $I$  denote a unit matrix. All these definitions assume that we work on a uniform grid, however this could be easily generalized for the non-uniform grid as well, see, e.g., In't Hout and

<sup>2</sup>They are also known as the modified Bessel functions of the second kind, or Macdonald functions, see Spanier and Oldham (1987).

Foulon (2010). Below we consider two cases,  $\beta < 0$  and  $\beta \geq 0$ .

**Proposition 3.2.** *If  $\beta < 0$ , then the discrete counterpart  $J$  of the operator  $\mathcal{J}$  is the negative of an EM-matrix if*

$$J = \delta \left( \sqrt{\alpha^2 - \beta^2} I - \left[ (\alpha^2 - \beta^2) I - 2\beta A_2^B - A_2^C \right]^{1/2} \right)$$

The matrix  $J$  is an  $O(h^2)$  approximation of the operator  $\mathcal{J}$ .

**Proof.** See Appendix.

**Proposition 3.3.** *If  $\beta \geq 0$ , then the discrete counterpart  $J$  of the operator  $\mathcal{J}$  is the negative of an EM-matrix if*

$$J = \delta \left( \sqrt{\alpha^2 - \beta^2} I - \left[ (\alpha^2 - \beta^2) I - 2\beta A_2^F - A_2^C \right]^{1/2} \right)$$

The matrix  $J$  is an  $O(h^2)$  approximation of the operator  $\mathcal{J}$ .

**Proof.** See Appendix.

Thus, according to Proposition 3 these finite difference schemes are unconditionally stable starting with some  $N$ , preserve positivity of the solution, and approximate operator  $\mathcal{J}$  with  $O(h^2)$ . In our experiments shown below this positivity was achieved at  $N > 100$ .

To complete the solution we need to compute the matrix exponential  $\mathcal{A} = e^{\Delta\tau\mathcal{J}}$  and the product  $z(x, \tau) = \mathcal{A}C(x)$ . The last step, however, could be further simplified. To see this recall that the diffusion equations in Equation (10) have to be solved up to some order of approximation in time  $\tau$ . Suppose for this purpose we want to use a finite-difference scheme that provides a second order approximation,  $O((\Delta\tau)^2)$ . However, Equation (16) gives an *exact* solution of the corresponding pure jump equation (the second step in Strang's splitting scheme). Since Strang's scheme guarantees only second-order accuracy ( $O((\Delta\tau)^2)$ ) to the exact solution of the full PIDE, the second step could be computed to the same order of accuracy.

To this end we can use the (1,1) Páde approximation of  $e^{\Delta\tau\mathcal{J}}$ ,

$$e^{\Delta\tau\mathcal{J}} \approx \left[ 1 - \frac{1}{2}\Delta\tau\mathcal{J} \right]^{-1} \left[ 1 + \frac{1}{2}\Delta\tau\mathcal{J} \right] + O(\Delta\tau^3). \quad (17)$$

This could be re-written in the form of the fixed point Picard iterations scheme

$$\begin{aligned} C^{(1)}(x, \tau + \Delta\tau) - C^{(1)}(x, \tau) \\ = \frac{1}{2}\Delta\tau\mathcal{J}^* \left[ C^{(1)}(x, \tau + \Delta\tau) + C^{(1)}(x, \tau) \right], \end{aligned}$$

and this equation could be solved iteratively starting with the initial guess  $C^{(1)}(x, \tau + \Delta\tau) = C^{(1)}(x, \tau)$ . Note that at each iteration the vector  $z(x, t)$  should be computed.

**Numerical experiments.** In our numerical experiments we consider the NIG model which also has a diffusion component uncorrelated with the jumps. We compute just one step of the splitting procedure, i.e. the jump integral, and don't consider solution of the diffusion part of the model. We want to price a European call option and take the option model parameters similar to d'Halluin et al. (2005b), i.e.,  $S_0 = K = 100, r = 0.05, \sigma = 0.15$ . The NIG model parameters are  $\delta = 0.2, \alpha = 10, \beta = -5.7, \mu = 0$ . One step in time is computed by taking  $T = \Delta\tau = 0.01$ . As  $C^{(1)}(x, \Delta\tau)$  in the Equation (10) comes after the first step of splitting, we get it by using the Black-Scholes formula with the forward interest rate  $r + c$  where the term  $c = -\log \Phi_{NIG}(\alpha, \beta, \delta, \mu, -i)$  comes from the last term in the jump integral in Equation (6). At the second step the solution of the jump part  $C_j^{(2)}(x, \Delta\tau)$  is produced given the initial condition  $C^{(1)}(x, \Delta\tau)$  from the previous step. We compare our solution for the jump step with that obtained with  $N = 2300$  which is assumed to be close to the exact value<sup>3</sup>. The finite-difference grid was constructed as follows: the diffusion grid was taken from  $x_{min}^D = 10^{-3}$  to  $x_{max}^D = 30 \max(S, K)$ . The jump grid is a superset of the diffusion grid, i.e. it coincides with the diffusion grid at the diffusion domain and then extends this domain up to  $x_{max}^J = \log(10^5)$ . Here to simplify the convergence analysis we use an uniform grid with step  $h$ . However, non-uniform grid can be easily constructed as well, and, moreover, that is exactly what this algorithm was constructed for. The results of such a test are given in Table 1. Here  $C$  is the price in dollars,  $N$  is the number of grid nodes,  $\beta_i$  is the order of convergence of the scheme. The "exact" price obtained at  $N = 2300$  is  $C_{num}(\Delta\tau) = 0.756574$ . It is seen that the convergence order  $\beta_i = \log_2 \frac{C^{(i)} - C_{num}}{C^{(i+1)} - C_{num}}, i = 1, 2, \dots$  of the scheme is asymptotically close to  $O(h^2)$ .

<sup>3</sup>This method is not very accurate. But as the exact solution is not known, it provides a plausible estimate of the convergence.

Table 1

Convergence of the proposed scheme for the NIG model with  $\beta = -5.7 < 0$ ,  $T = \Delta\tau = 0.01$

$C$	$h$	$N$	$\beta_i$
4.0657	0.2763100	51	-
2.1239	0.1381550	101	1.275
1.1914	0.0690776	201	1.653
0.8171	0.0345388	401	2.845
0.7597	0.0172694	801	4.295
0.7565	0.0086347	1601	5.005

Table 2

Convergence of the proposed scheme for the NIG model with  $\beta = 5.7 > 0$ ,  $T = \Delta\tau = 0.01$

$C$	$h$	$N$	$\beta_i$
4.0127	0.2763100	51	-
2.1896	0.1381550	101	1.191
1.1627	0.0690776	201	1.850
0.8393	0.0345388	401	2.480
0.7710	0.0172694	801	4.892
0.7685	0.0086347	1601	4.557

As a sanity check we can compare this value with the reference value obtained by pricing this model (one step) using FFT, which is  $C_{FFT}(\Delta\tau) = 0.757782$ . Note, that  $C_{FFT}(\Delta\tau)$  should not be exactly equal to  $C_{num}(\Delta\tau)$  because we use two steps instead of three as in the Strang algorithm<sup>4</sup> which are equivalent to the splitting scheme of the first order in  $\Delta\tau$ , i.e. it has an error  $O(\Delta\tau)$ . However, still  $C_{FFT}(\Delta\tau)$  is close to  $C_{num}(\Delta\tau)$ . The second experiment uses the same set of parameters, but now  $\beta = 5.7 > 0$ . The results are given in Table 2. The "exact" price obtained at  $N = 2300$  is  $C_{num}(\Delta\tau) = 0.76864$ . Again the convergence order  $\beta_i$  of the scheme is close to  $O(h^2)$ .

In this test  $C_{FFT}(\Delta\tau) = 0.76773$  which is also close to  $C_{num}(\Delta\tau)$ . As a final note, performance wise

$$\nu(dx) = \frac{\exp(\beta x)}{|x|} \left( \int_0^\infty \frac{\exp(-\sqrt{2y + \alpha^2}|x|)}{\pi^2 y (J_{|\lambda|}^2(\delta\sqrt{2y}) + Y_{|\lambda|}^2(\delta\sqrt{2y}))} dy + 1_{\lambda \geq 0} \lambda e^{-\alpha|x|} \right) dx,$$

the principal square root of matrix is better to compute using the product form of the Denman-Beavers iteration, see Denman and Beavers (1976) for the description of the algorithm.

<sup>4</sup>Don't miss this with the accuracy of the whole 3 steps Strang's algorithm which is  $O(\Delta\tau^2)$ . The test validates just the convergence in  $h$ , not in  $\Delta\tau$ .

### 3.2. Generalized hyperbolic models

Generalized hyperbolic process constitutes a broad subclass of Lévy processes which are generated by generalized hyperbolic (GH) distributions. They were introduced in Barndorff-Nielsen (1977). See also Eberlein and Keller (1995) for the detailed survey on how the hyperbolic distributions are used in finance.

The Lebesgue density of the GH distribution is a 5-parameter function

$$f(\lambda, \alpha, \beta, \delta, \mu) = a(\lambda, \alpha, \beta, \delta, \mu) \left[ \delta^2 + (x - \mu)^2 \right]^{\frac{\lambda-1/2}{2}} K_{\lambda-1/2} \left( \alpha \sqrt{\delta^2 + (x - \mu)^2} \right) e^{\beta(x-\mu)}, \quad (18)$$

where the normalization constant reads

$$a(\lambda, \alpha, \beta, \delta, \mu) = \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi} \alpha^{\lambda-1/2} \delta^\lambda K_\lambda(\delta \sqrt{\alpha^2 - \beta^2})}.$$

Here  $\alpha > 0$  determines the shape of the distribution,  $\beta$  determines the skewness, and  $0 \leq |\beta| < \alpha$ ,  $\mu \in \mathbb{R}$  determines the location (mean),  $\delta > 0$  is scaling, and  $\lambda \in \mathbb{R}$  determines the weight of the distribution in the tails. In particular,  $\lambda = -1/2$  corresponds to the NIG distribution considered in the previous sections

The characteristic exponent of the GH process is

$$\phi(u, \alpha, \beta, \delta, \mu, \lambda) = iu\mu + \log \Psi \quad (19)$$

$$\Psi = \left( \frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + iu)^2} \right)^{\lambda/2} \frac{K_\lambda(\delta \sqrt{\alpha^2 - (\beta + iu)^2})}{K_\lambda(\delta \sqrt{\alpha^2 - \beta^2})}.$$

and the Lévy density  $\nu(dx)$  of the GH Lévy motions reads

where  $J, Y$  are the corresponding Bessel functions.

From Equation (7), Equation (19) we immediately obtain that

$$\mathcal{J} = \phi(-i\nabla, \alpha, \beta, \delta, \mu, \lambda),$$

and the jump evolution equation (the second equation in Equation (10)) becomes

$$C(x, \tau + \Delta\tau) = \mathcal{A}C(x, \tau), \quad (20)$$

$$\mathcal{A} = e^{t\mu\nabla} \left( \frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + \nabla)^2} \right)^{\lambda\Delta\tau} \left\{ \frac{K_\lambda(\delta\sqrt{\alpha^2 - (\beta + \nabla)^2})}{K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})} \right\}^{\Delta\tau}.$$

Notice that in our splitting algorithm we have a liberty to re-distribute the terms between the jump and diffusion parts. Therefore, the first term  $e^{t\mu\nabla}$  could be taken out of this expression and moved to the diffusion part of Equation (8). The remaining operator  $\mathcal{B} = e^{-t\mu\nabla}\mathcal{A}$  could be represented in a form of a product of two operators

$$\mathcal{B} = \mathcal{B}_1\mathcal{B}_2, \quad (21)$$

$$\mathcal{B}_1 = \left( \frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + \nabla)^2} \right)^{\lambda\Delta\tau},$$

$$\mathcal{B}_2 = \left\{ \frac{K_\lambda(\delta\sqrt{\alpha^2 - (\beta + \nabla)^2})}{K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})} \right\}^{\Delta\tau}.$$

To construct the approximation of  $\mathcal{B}_2$  take into account that by Abramowitz and Stegun (1964) for the modified Bessel functions of the third kind we have

$$K_\nu(z) = \sqrt{\frac{\pi}{2z}} e^{-z} \sum_{k=0}^{\infty} \frac{a_k(\nu)}{z^k},$$

$$|z| \rightarrow \infty, \quad |\arg z| < \frac{3\pi}{2}, \quad (22)$$

$$a_0(\nu) = 1, \quad a_k(\nu) = \frac{(4\nu^2 - 1^2)\dots(4\nu^2 - (2k - 1)^2)}{k!8^k}.$$

We want to approximate  $K_\nu(z)$  up to  $O(h^2)$ . All operators  $\mathcal{B}_i, i = 1, 2$  in Equation (21) are actually the operator functions of another operator  $z(\nabla)$ , where  $z(x) \equiv \delta\sqrt{\alpha^2 - (\beta + x)^2}$ . Obviously, any order discretization of the operator  $\nabla$  on a grid is proportional to  $1/h$ . Therefore, discretization of  $z$  is also proportional to  $1/h$ . Hence, such discretization applied to the terms  $1/z^k$  will be proportional to  $h^k$ . When this operator affects a discrete vector function also defined at the same grid, the total error will not be worth than  $O(h^2)$ . That means that in the series expansion Equation (22) we can keep only terms with  $k = 0, 1$  while omitting the remaining ones.

With allowance for that and Equation (22) we redefine  $\mathcal{B}_i, i = 1, 2$  as follows.

$$\mathcal{B} = \mathcal{B}_1\mathcal{B}_2, \quad \mathcal{B}_1 = \left( \frac{z(0)}{z(\nabla)} \right)^{(\lambda+1/2)\Delta\tau}, \quad (23)$$

$$\mathcal{B}_2 = \left[ e^{z(0)-z(\nabla)} \left( 1 + \frac{4\lambda^2 - 1}{z(\nabla)} \right) \left( \sum_{k=0}^{\infty} \frac{a_k(\nu)}{z(0)^k} \right)^{-1} \right]^{\Delta\tau}.$$

Thus, now we need to construct an appropriate discretization of the operators  $\mathcal{B}_i, i = 1, 2$ . In doing that consider two cases.

### 3.2.1. $\lambda \geq -1/2$

**Proposition 3.4.** Assume that  $\beta < 0$ . Denote by  $B_1$  the following discrete representation of the operator  $\mathcal{B}_1$  on a given grid  $\mathbf{G}(\mathbf{x})$ :

$$B_1 = \left[ \frac{(\alpha^2 - \beta^2)I - 2\beta A_2^B - A_2^C}{\alpha^2 - \beta^2} \right]^{-\Delta\tau(\lambda+1/2)/2}$$

Then  $B_1$  is a nonnegative matrix with all eigenvalues  $|\lambda_i| < 1, \forall i \in [1, N]$ . The matrix  $B_1$  is an  $O(h^2)$  approximation of the operator  $\mathcal{B}_1$ .

**Proof.** See Appendix.

Note that for a non-uniform grid the prove could be constructed in a similar way, but requires many technical details which we do not consider in this paper.

**Proposition 3.5.** Assume that  $\beta \geq 0$ . Denote by  $B_1$  the following discrete representation of the operator  $\mathcal{B}_1$  on a given grid  $\mathbf{G}(\mathbf{x})$ :

$$B_1 = \left[ \frac{(\alpha^2 - \beta^2)I - 2\beta A_2^F - A_2^C}{\alpha^2 - \beta^2} \right]^{-\Delta\tau(\lambda+1/2)/2}$$

Then  $B_1$  is a nonnegative matrix with all eigenvalues  $|\lambda_i| < 1, \forall i \in [1, N]$ . The matrix  $B_1$  is an  $O(h^2)$  approximation of the operator  $\mathcal{B}_1$ .

**Proof.** The proof is analogous to the previous Proposition with allowance for the fact that matrix  $-2\beta A_2^B$  with  $\beta < 0$  is the transpose of the matrix  $-2\beta A_2^F$  with  $\beta \geq 0$ .  $\square$

Now observe that the first operator  $e^{z(0)-z(\nabla)}$  in the definition of  $\mathcal{B}_2^{\Delta\tau}$  is exactly the operator  $\mathcal{A}$  in Equation (16). Therefore, Propositions 3.1 and 3.1 could be used to construct the corresponding discretizations. The second part (a product of two terms in parentheses) could be represented as

$$\begin{aligned}
C &= \gamma \left( 1 + \frac{4\lambda^2 - 1}{z(\nabla)} \right) \\
&= \gamma \left[ 1 + (4\lambda^2 - 1) \left( \frac{1}{z^2(\nabla)} \right)^{1/2} \right], \gamma = \left( \sum_{k=0}^{\infty} \frac{a_k(v)}{z(0)^k} \right)^{-1}.
\end{aligned} \tag{24}$$

Operator  $1/z^2(\nabla)$  could be discretized using Propositions 3.2.1 and 3.2.1. Coefficient  $\gamma$  guarantees that all eigenvalues of the discrete discretization matrix are less than one. Thus, the proposed scheme is unconditionally stable.

Obviously as each operator  $\mathcal{B}_i$ ,  $i = 1, 2$  has the necessary property, a "product" of three operators (the consecutive application of them one after another) will result into the combined operator having same properties. This finalizes our construction.

Also notice, that at  $\lambda = -1/2$ , which corresponds to the NIG model considered in the previous sections, our scheme exactly translates to the scheme proposed there.

### 3.2.2. $\lambda < -1/2$

At the first glance in this case the previous scheme will not converge as the eigenvalues of the operator  $\mathcal{B}_1$  are greater than one. Nevertheless, as shown below still it can be used. To demonstrate that for the sake of convenience we re-write  $\mathcal{B}$  in Equation (23) in the form

$$\begin{aligned}
\mathcal{B} &= \left\{ e^{\mathcal{M}} \left( 1 + \frac{4\lambda^2 - 1}{z(\nabla)} \right) \left( \sum_{k=0}^{\infty} \frac{a_k(v)}{z(0)^k} \right)^{-1} \right\}^{\Delta\tau}, \\
\mathcal{M} &= z(0) - z(\nabla) - (\lambda + 1/2) \log \frac{z(\nabla)}{z(0)}. \tag{25}
\end{aligned}$$

where now  $-(\lambda + 1/2) > 0$ . So the only difference with the case  $\lambda \geq -1/2$  is that we moved the operator  $\mathcal{B}_1$  into the exponent. Because of that for the terms in parentheses we leave the same discretization as in Propositions 3.2.1, 3.2.1. Then the following Proposition is in order.

**Proposition 3.6.** Assume that  $\beta < 0$ . Denote by  $Z$  the following discrete representation of the operator  $z(\nabla)$  on a given grid  $\mathbf{G}(\mathbf{x})$ :

$$Z = \delta[(\alpha^2 - \beta^2)I - 2\beta A_2^B - A_2^C]^{1/2}$$

Then

$$\begin{aligned}
B &= \left( \sum_{k=0}^{\infty} \frac{a_k(v)}{z(0)^k} \right)^{-\Delta\tau} \left\{ e^{\mathcal{M}} \left[ 1 + (4\lambda^2 - 1)Z^{-1} \right] \right\}^{\Delta\tau}, \\
M &= z(0) - Z - (\lambda + 1/2) \log \frac{Z}{z(0)}.
\end{aligned}$$

is a nonnegative matrix with all eigenvalues  $|\lambda_i| < 1$ ,  $\forall i \in [1, N]$ . The matrix  $B$  is an  $O(h^2)$  approximation of the operator  $\mathcal{B}$  in Equation (25).

**Proof.** See Appendix.

**Proposition 3.7.** Assume that  $\beta \geq 0$ . Denote by  $Z$  the following discrete representation of the operator  $z(\nabla)$  on a given grid  $\mathbf{G}(\mathbf{x})$ :

$$Z = \delta[(\alpha^2 - \beta^2)I - 2\beta A_2^F - A_2^C]^{1/2}$$

Then

$$\begin{aligned}
B &= \left( \sum_{k=0}^{\infty} \frac{a_k(v)}{z(0)^k} \right)^{-\Delta\tau} \left\{ e^{\mathcal{M}} \left[ I + (4\lambda^2 - 1)Z^{-1} \right] \right\}^{\Delta\tau}, \\
M &= z(0) - Z - (\lambda + 1/2) \log \frac{Z}{z(0)}.
\end{aligned}$$

is a nonnegative matrix with all eigenvalues  $|\lambda_i| < 1$ ,  $\forall i \in [1, N]$ . The matrix  $B$  is an  $O(h^2)$  approximation of the operator  $\mathcal{B}$  in Equation (25).

**Proof.** The proof is analogous to the previous Proposition with allowance for the fact that matrix  $-2\beta A_2^B$  with  $\beta < 0$  is the transpose of the matrix  $-2\beta A_2^F$  with  $\beta \geq 0$ . ■

**Numerical experiments.** Our numerical experiments are provided similar to how we did it with the model considered in the previous section. The GH model parameters are  $\alpha = 10$ ,  $\beta = -5.7$ ,  $\delta = 0.2$ ,  $\mu = 0$ . The results for  $\lambda = -1$  are given in Table 3. The "exact" price obtained at  $N = 2100$  is  $C_{num}(\Delta\tau)$

Table 3

Convergence of the proposed scheme for the GH model, $\lambda = -1$			
$C$	$h$	$N$	$\beta_i$
4.0492	0.2763100	51	-
2.1145	0.1381550	101	1.265
1.1701	0.0690776	201	1.666
0.7979	0.0345388	401	2.805
0.7389	0.0172694	801	4.325
0.7357	0.0086347	1601	5.644

Table 4  
Convergence of the proposed scheme for the GH model,  $\lambda = 1$

$C$	$h$	$N$	$\beta_i$
4.1444	0.2763100	51	-
2.1657	0.1381550	101	1.320
1.2827	0.0690776	201	1.592
0.8978	0.0345388	401	3.038
0.8475	0.0172694	801	4.107
0.8443	0.0086347	1601	5.490

= 0.73580. The observed convergence order  $\beta_i$  of the scheme is close to  $O(h^2)$ . Also  $C_{FFT}(\Delta\tau) = 0.73746$  which is also close to  $C_{num}(\Delta\tau)$ .

The second test, which results are given in Table 4, uses  $\lambda = 1$ . The "exact" price obtained at  $N = 2100$  is  $C_{num}(\Delta\tau) = 0.84440$ . The observed convergence order  $\beta_i$  of the scheme is also close to  $O(h^2)$ . The FFT price  $C_{FFT}(\Delta\tau) = 0.846985$  is at the distance of 0.3% from  $C_{num}(\Delta\tau)$  while theoretically the error should be proportional to  $O(\Delta\tau) = 1\%$ . Thus, it seems very reasonable that this error is due to the fact that we actually used the first order approximation in time in this test (see a detailed explanation for doing that in Section 3.1).

### 3.3. Meixner model

Meixner jump model was introduced by Schoutens (2001), Schoutens and Teugels (1998). It is built based on the Meixner distribution which belongs to the class of the infinitely divisible distributions. Therefore, it gives rise to a Lévy process - the Meixner process. The Meixner process is flexible and analytically tractable, i.e. its pdf and CF are known in closed form.

The density of the Meixner distribution  $f(a, b, d, m)$  reads

$$f(x; a, b, d, m) = \frac{[2 \cos(b/2)]^{2d}}{2a\pi\Gamma(2d)} \exp\left[\frac{b(x-m)}{a}\right] \left| \Gamma\left(d + \frac{i(x-m)}{a}\right) \right|^2,$$

where  $a > 0$ ,  $-\pi < b < \pi - a$ ,  $d > 0$ ,  $m \in \mathbb{R}$ .

The characteristic exponent of the Meixner process is

$$\phi(u, a, b, d, m) = 2d \left\{ \log[\cos(b/2)] - \log \left[ \cosh \left( \frac{au - ib}{2} \right) \right] \right\} + imu, \tag{26}$$

and the Lévy density  $\nu(dx)$  of the Meixner process reads

$$\nu(dx) = d \frac{\exp(bx/a)}{x \sinh(\pi x/a)} dx.$$

From Equation (7), Equation (26) we immediately obtain

$$\begin{aligned} \mathcal{J} &= \phi(-i\nabla, a, b, d, m) \\ &= 2d \left\{ \log[\cos(b/2)] - \log \left[ \cos \left( \frac{a\nabla + b}{2} \right) \right] \right\} + m\nabla. \end{aligned} \tag{27}$$

Now observe that the last term  $m\nabla$  could be taken out and moved to the diffusion part of Equation (8). Again, this is because when constructing our splitting algorithm we have a freedom to decide which terms to keep under the jump part and which ones should be moved to the diffusion part. As the term  $m\nabla$  is proportional to  $\nabla$ , i.e. it looks similar to the drift term of the diffusion part, we can naturally add it to the drift and eliminate it from the jump integral assuming that the remaining expression of  $\mathcal{J}$  is well-defined.

Using the remaining part of Equation (26) the operator  $\mathcal{A} = e^{\Delta\tau\mathcal{J}}$  can be represented in the form

$$\mathcal{A} = [\cos(b/2)]^{2d\Delta\tau} \left[ \sec \left( \frac{a\nabla + b}{2} \right) \right]^{2d\Delta\tau} \tag{28}$$

Thus, our goal is to compute the product  $\mathcal{A}\mathcal{C}(x, \tau)$ .

To do that let us use a representation of  $\cos(x)$  in a form of the infinite product, see Abramowitz and Stegun (1964)

$$\cos(x) = \prod_{n=1}^{\infty} \left[ 1 - \frac{x^2}{\pi^2(n-1/2)^2} \right]$$

Then Equation (28) could be re-written as follows

$$\begin{aligned} \mathcal{A} &= [\cos(b/2)]^\kappa \prod_{n=1}^{\infty} \mathcal{A}_n, \\ \mathcal{A}_n &\equiv \left[ 1 - \frac{(a\nabla + b)^2}{4\pi^2(n-1/2)^2} \right]^{-\kappa}, \quad \kappa = 2d\Delta\tau. \end{aligned} \tag{29}$$

The following Proposition now gives the solution of our problem.

**Proposition 3.8.** Assume that  $b < 0$ . Denote by  $M$  the following discrete representation of the operator

$$z(\nabla) = 1 - \frac{(a\nabla + b)^2}{4\pi^2(n - 1/2)^2}$$

on a given grid  $\mathbf{G}(\mathbf{x})$ :

$$M_n = I - \frac{1}{4\pi^2(n - 1/2)^2} \left( a^2 A_2^C + 2abA_2^B + b^2 I \right)$$

Then

$$B = [\cos(b/2)]^\kappa \prod_{n=1}^{\infty} M_n^{-\kappa}$$

is a nonnegative matrix with all eigenvalues  $|\lambda_i| < 1$ ,  $\forall i \in [1, N]$ . The matrix  $B$  is an  $O(h^2)$  approximation of the operator  $\mathcal{A}$  in Equation (28).

**Proof.** See Appendix.

A similar Proposition is in place when  $0 < b < \pi - a$ .

**Proposition 3.9.** Assume that  $0 < b < \pi - a$ . Denote by  $M$  the following discrete representation of the operator  $z(\nabla)$  on a given grid  $\mathbf{G}(\mathbf{x})$ :

$$M_n = I - \frac{1}{4\pi^2(n - 1/2)^2} \left( a^2 A_2^C + 2abA_2^F + b^2 I \right)$$

Then

$$B = [\cos(b/2)]^\kappa \prod_{n=1}^{\infty} M_n^{-\kappa}$$

is a nonnegative matrix with all eigenvalues  $|\lambda_i| < 1$ ,  $\forall i \in [1, N]$ . The matrix  $B$  is an  $O(h^2)$  approximation of the operator  $\mathcal{A}$  in Equation (28).

**Proof.** See Appendix.

Once all matrices  $M_n$  are constructed, a sequential product

$$BC(x, \Delta\tau) = [\cos(b/2)]^\kappa (\dots M_n^{-\kappa} (M_{n-1}^{-\kappa} \dots (M_2^{-\kappa} (M_1^{-\kappa} C(x, \Delta\tau))))))$$

can be obtained by multiplying the first matrix  $M_1^{-\kappa}$  by the vector  $C(x, \Delta\tau)$ , then multiplying the second matrix  $M_2^{-\kappa}$  by the resulting vector, etc. FFT can be used for this purpose, see Wang et al. (2007) for more details.

For practical applications the infinite product in Equation (29) should be truncated. For instance  $p = 5$  already provides a relative accuracy better than 1%.

However, in many situations a more efficient method could be proposed. Assume that  $0 \leq 2d\Delta\tau \leq 2$ . Indeed, this is always the case as  $\Delta\tau \ll 1$  while the values of  $d$  found, e.g., in Schoutens (2003) when calibrating the Meixner model to the option market data, were about  $d = 50$ . Obviously, if necessary, decreasing  $\Delta\tau$  we can always make the above inequality valid. However, this is not an attractive way to proceed, so below we assume that for the reasonable values of  $\Delta\tau$  this inequality is correct.

If so, then a variation of the method proposed in Itkin and Carr (2012) could be applied. The idea of the method is to consider the discrete operator  $B$  as a function of the parameter  $k$

$$B(k) = [\cos(b/2)]^\kappa \prod_{n=1}^{\infty} M_n^{-\kappa}$$

According to our assumption  $0 \leq \kappa \leq 2$ . Therefore, we can compute three vectors  $z_0 = B(0)C(x, \tau)$ ,  $z_1 = B(1)C(x, \tau)$ , and  $z_2 = B(2)C(x, \Delta\tau)$  and then interpolate them point-wise to  $\kappa = 2d\Delta\tau$ . Monotonic spline interpolation could be used for this purpose. Again, see Itkin and Carr (2012) for more details, as well as the Theorem about a continuity of the price in  $\kappa$  space proven there.

It is easy to see that  $z_0 = C(x, \tau)$ . Now observe, that at  $\kappa = 1$  for every  $n = 1, 2, \dots$  vector  $z_{1,n}$  solves the following system of linear equations

$$M_n z_{1,n}(x, \tau + \Delta\tau) = z_{1,n-1}(x, \tau) \quad (30)$$

where  $z_{1,1}(x, \tau) = C(x, \tau)$ . Matrix  $M_n$  by construction is a banded matrix with 4 non-zero diagonals. Also according to the Propositions 3.3, 3.3 it is an EM-matrix, therefore Equation (30) is well-defined. As the complexity of this solution is  $O(N)$ , the total complexity of this step at  $\kappa = 1$  is  $O(pN)$ . This is worse than a pure linear convergence, but still could be better than that of the FFT. After the  $p$  steps are done the final step is to multiply  $z_{1,p}$  by  $[\cos(b/2)]$ .

Similarly, at  $\kappa = 2$  vector  $z_{2,n}$  solves

$$M_n^2 z_{2,n}(x, \tau + \Delta\tau) = z_{2,n-1}(x, \tau) \quad (31)$$

where  $z_{2,1}(x, \tau) = C(x, \tau)$ . Matrix  $M_n^2$  by construction is a banded matrix with 7 non-zero diagonals, and also is an EM-matrix. Therefore, Equation (31) is well-defined and could be solved with the complexity  $O(N)$ . The total complexity of this step is also  $O(pN)$ , hence the total complexity of the entire algorithm is  $O(2pN)$ .

Finally point-wise interpolation of three vectors to the given value of  $\kappa$  has the complexity  $O(N)$  if the interpolation coefficients are precomputed. Therefore, the total complexity of the method is still  $2pO(N)$ .<sup>5</sup>

**Numerical experiments.** In our numerical experiment the values of the Meixner model parameters are taken as suggested in Schoutens (2003), i.e.  $a = 0.04$ ;  $b = -0.32754$ ,  $d = 52$ , but here we use  $m = 0$ . Other parameters are the same as in the previous sections.

The results obtained with the first method at  $p = 10$  are given in Table 5. The "exact" price obtained at  $N = 3200$  is  $C_{num}(\Delta\tau) = 1.0068$ . The observed convergence order  $\beta_i$  of the scheme is close to  $O(h^2)$  up to the point  $N = 1601$  where it drops down. To check what is the problem we ran the second test because further increase of  $N$  when computing the matrix exponential gives rise to the RAM capacity of our PC being exceeded, while for the second method this is not a problem due to the banded structure of all matrices.

In this test  $C_{FFT}(\Delta\tau) = 1.0145$  which is also close to  $C_{num}(\Delta\tau)$ .<sup>6</sup>

In the second test we repeated the previous one but now using our second approach - interpolation in the  $\kappa$  domain. The results are given in Table 6.

The "exact" price obtained at  $N = 6401$  is  $C_{num}(\Delta\tau) = 1.0072$ . Despite the convergence is close to  $O(h^2)$  almost at all  $h$ , there is a spike at  $N = 1601$  which indicates that monotonicity of the price as a function of  $h$  changes close to this point.

The typical time to compute the price in a single point  $x$  using the cosine method with 12 terms in the expansion in our experiments was 4 msec. The time necessary for the interpolation method to compute  $C(x, \tau + \Delta\tau)$  with  $x$  now being a vector containing  $N = 801$  points and  $p = 10$  was 4.2 msec. Observe, that if in this test we change  $p$  to  $p = 5$  the total time accordingly halves the previous one, while the "exact" price obtained at  $N = 6401$  becomes  $C_{num}(\Delta\tau) = 1.0031$ . In other words the difference is just 0.4%.

We also regressed the computational time at  $p = 5$  to the number of grid points  $N$  to check the order of

<sup>5</sup>As in our case the banded matrices have either 4 or 7 diagonals, this complexity is also proportional to some coefficient which in this case could be about 4-7 per one solution for a given  $n$ . Still as was checked in our tests the method is faster than the FFT at relatively small  $p = 3 - 5$ , see below.

<sup>6</sup>The standard FFT method at these values of parameters is very sensitive to the choice of the dumping factor  $\alpha$ , therefore this price was computed using the cosine method of Fang and Oosterlee (2008).

Table 5

Convergence of the proposed scheme for the Meixner model

$C$	$h$	$N$	$\beta_i$
4.1791	0.2763100	51	-
2.1771	0.1381550	101	1.439
1.3661	0.0690776	201	1.7045
1.0328	0.0345388	401	3.789
1.0060	0.0172694	801	5.022
1.0058	0.0086347	1601	0.322

Table 6

Convergence of the proposed scheme for the Meixner model - the "interpolation in  $\kappa$ " method

$C$	$h$	$N$	$\beta_i$
4.1779	0.2763100	51	-
2.1765	0.1381550	101	1.439
1.3657	0.0690776	201	1.705
1.0327	0.0345388	401	3.812
1.0060	0.0172694	801	4.410
1.0058	0.0086347	1601	0.154
1.0068	0.0043174	3201	1.842

Table 7

Regression of the elapsed time  $t_i$  for the interpolation method to the number of grid points  $N_i$

$N$	101	201	401	801	1601
$\beta$	0.529	0.382	1.031	2.017	0.599

convergence of the method. The results are given below in Table 7.

Here  $\beta = \log_2 \left( \frac{t_i - t_{i+1}}{t_{i+1} - t_{i+2}} \right)$ ,  $T_i$  is the elapsed time when  $N = N_i$ . It is known that this method usually is not very accurate in the estimate of the complexity, however we don't have a better one. It is seen that except the point at  $N = 801$  the complexity is about  $O(N)$ .

#### 4. Conclusion

This paper is a further extension of Itkin (2014)) where a new method of solving jump-diffusion PIDEs was proposed. This method exploits a number of ideas, namely:

1. First, we transform a linear non-local integro-differential operator (jump operator) into a local linear but pseudo-differential (fractional) operator. Thus, the whole jump-diffusion operator  $\mathcal{J} + \mathcal{D}$  is represented as a sum of the differential and pseudo-differential parts.
2. Second, operator splitting on financial processes<sup>7</sup> is applied to this operator, namely splitting a space

<sup>7</sup>This is similar to splitting on physical processes, e.g., convection and diffusion, which is well-known in computational physics.

operator into the diffusion and jumps parts. The described splitting scheme provides a second-order approximation of  $\mathcal{J} + \mathcal{D}$  in time.

- At the third step various finite-difference approximations of the pseudo-differential operator  $\mathcal{J}$  are proposed. In Itkin (2014) Merton, Kou and GTSP (aka CGMY or KoBoL) models were considered. In this paper we demonstrated how to construct these approximations for the NIG, Hyperbolic and Meixner models to (i) be unconditionally stable, (ii) be of first- and second-order approximation in  $h$  which is a step size of the grid, and (iii) preserve positivity of the solution. The results are presented as propositions, and the corresponding proofs are given based on modern matrix analysis, including a theory of M-matrices, the Metzler matrices and eventually exponentially nonnegative matrices.

All these results seem to be new. The method is naturally applicable to both uniform and nonuniform grids, and is easy for programming, since the algorithm is similar to all jump models. Also notice that the present approach admits pricing of some exotic, e.g., barrier options as well. In addition, it respects not just vanilla but also digital payoffs. In principle, American and Bermudan options could also be priced by this method, however this requires some more delicate consideration which will be presented elsewhere.

## Acknowledgments

I thank Peter Carr and Peter Forsyth for very fruitful discussions, and Alex Lipton, Igor Halperin and Alexey Polishchuk for useful comments. I assume full responsibility for any remaining errors.

## References

- Abramowitz, M., Stegun, I., 1964. Handbook of Mathematical Functions. Dover Publications, Inc.
- Andersen, L., Andreasen, J., 2000. Jump diffusion processes: Volatility smile fitting and numerical methods for option pricing. Review of Derivatives Research. 4, 231–262.
- Barndorff-Nielsen, O., 1977. Exponentially decreasing distributions for the logarithm of particle size. In Proceedings of the Royal Society of London, volume A, pp. 401–419.
- Barndorff-Nielsen, O.E., 1998. Processes of normal inverse gaussian type. Finance and Stochastics. 2, 41–68.
- Carr, P., Madan, D., 1999. Option valuation using the Fast Fourier Transform. Journal of Computational Finance. 2(4), 61–73.
- Cont, R., Tankov, P., 2004. Financial modelling with jump processes. Financial Mathematics Series, Chapman & Hall/CRCI.
- Cont, R., Voltchkova, E., 2003. A finite difference scheme for option pricing in jump diffusion and exponential Lévy models. Technical Report 513, Rapport Interne CMAP.
- de Lange, O.L., Raab, R.E., 1992. Operator Methods in Quantum Mechanics. Oxford science publications. Chapter 3.
- Denman, E., Beavers, A., 1976. The matrix sign function and computations in systems. Applied Mathematics and Computation. 2(1), 63–94.
- d'Halluin, Y., Forsyth, P.A., Labahn, G., 2005a. A semi-Lagrangian approach for American Asian options under jump diffusion. SIAM Journal on Scientific Computing. 27, 315–345.
- d'Halluin, Y., Forsyth, P.A., Vetzal, K.R., 2004. A penalty method for American options with jump diffusion processes. Numerische Mathematik. 97, 321–352.
- d'Halluin, Y., Forsyth, P.A., Vetzal, K.R., 2005b. Robust numerical methods for contingent claims under jump diffusion processes. IMA J Numerical Analysis. 25, 87–112.
- Duffy, D., 2006. Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach. The Wiley Finance Series.
- Dyakonov, E., 1964. Difference schemes with a separable operator for general second order parabolic equations with variable coefficient. Zhurnal Vychislitelnoi Matematiki i Matematicheskoi Fiziki. 4(2), 278–291.
- Eberlein, E., 2009. Jump-type Lévy processes. In Andersen, T.G., Davis, R.A., Kreiß, J.-P., Mikosch, T., editors, Handbook of Financial Time Series, Springer Verlag, pp. 439–455.
- Eberlein, E., Keller, U., 1995. Hyperbolic distributions in finance. Bernoulli. 1, 281–299.
- Eberlein, E., Keller, U., Prause, K., 1998. New insights into smile, mispricing, and value at risk: The hyperbolic model. Journal of Business. 71(3), 371–406.
- Elhashash, A., Szyld, D., 2008. Generalizations of M-matrices which may not have a nonnegative inverse. Linear Algebra and its Applications. 429, 2435–2450.
- Fang, F., Oosterlee, C., 2008. A novel pricing method for European options based on Fourier-Cosine series expansions. SIAM J Sci Comput. 31(2), 826–848.
- Fang, F., Oosterlee, C., 2009. Pricing early-exercise and discrete barrier options by fourier-cosine series expansions. Numerische Mathematik. 114(1), 27–62.
- Gray, R.M., 2006. Toeplitz and circulant matrices: A review. Foundations and Trends in Communications and Information Theory. 2(3), 155–239.
- In't Hout, K.J., Foulon, S., 2010. ADI finite difference schemes for option pricing in the Heston model with correlation. International Journal of Numerical Analysis and Modeling. 7(2), 303–320.
- Itkin, A., 2014. Efficient solution of backward jump-diffusion PIDEs with splitting and matrix exponentials. Journal of Computational Finance, forthcoming. electronic version is available at <http://arxiv.org/abs/1304.3159>.
- Itkin, A., Carr, P., 2012. Using pseudo-parabolic and fractional equations for option pricing in jump diffusion models. Computational Economics. 40(1), 63–104.
- Le, H.T., McDonald, J.J., 2006. Inverses of M-type matrices created with irreducible eventually nonnegative matrices. Linear Algebra and its Applications. 419, 668–674.

[26] Lewis, A.L., 2001. A simple option formula for general jump-diffusion and other exponential Levy processes. manuscript, Envision Financial Systems and OptionCity.net, Newport Beach, California, USA.

[27] Lipton, A., 2001. Mathematical Methods For Foreign Exchange: A Financial Engineer's Approach. World Scientific.

[28] Lord, R., Fang, F., Bervoets, F., Oosterlee, K., 2007. A fast and accurate FFT-based method for pricing early-exercise options under Levy processes. Technical report, MRPA. available at <http://mpira.ub.uni-muenchen.de/1952/>.

[29] Marchuk, G., 1975. Methods of Numerical Mathematics. Springer-Verlag.

[30] Noutsos, D., Tsatsomeros, M.J., 2008. Reachability and holdability of nonnegative states. SIAM J Matrix Anal Appl, 30(2), 700–712.

[31] Olver, F., Lozier, D., Boisvert, R., Clark, C., 2010. Handbook of Mathematical Functions. Cambridge University Press.

[32] Samarski, A., 1964. Economical difference schemes for parabolic equations with mixed derivatives. Zhurnal Vychislitelnoi Matematiki i Matematicheskoi Fiziki. 4(4), 753–759.

[33] Schoutens, W., 2001. Meixner processes in finance. Technical report, K.U.Leuven Eurandom.

[34] Schoutens, W., 2003. Levy Processes in Finance: Pricing Financial Derivatives. Wiley.

[35] Schoutens, W., Teugels, J., 1998. Lévy processes, polynomials and martingales. Commun Statist- Stochastic Models. 14(1,2), 335–349.

[36] Spanier, J., Oldham, K., 1987. An Atlas of Functions. Washington DC. Hemisphere. ch. 51, The Bessel  $K_\nu(x)$ .

[37] Strang, G., 1968. On the construction and comparison of difference schemes. SIAM J Numerical Analysis. 5, 509–517.

[38] Wang, I., Wan, J., Forsyth, P., 2007. Robust numerical valuation of European and American options under the CGMY process. J Comp Finance. 4, 31–70.

[39] Yanenko, N., (1971). The method of fractional steps. Springer-Verlag.

## Appendices

### A Proof of Proposition 3.2

To prove this proposition we need technique that we used in Itkin (2014). It is closely related to the concept of an “eventually positive matrix”, see Noutsos and Tsatsomeros (2008). Below we reproduce some definitions from this paper necessary for our further analysis.

**Definition** An  $N \times N$  matrix  $A = [a_{ij}]$  is called

- *eventually nonnegative*, denoted by  $A \stackrel{v}{\geq} 0$ , if there exists a positive integer  $k_0$  such that  $A^k \geq 0$  for all  $k > k_0$ ; we denote the smallest such positive integer by  $k_0 = k_0(A)$  and refer to  $k_0(A)$  as the power index of  $A$ ;

- *exponentially nonnegative* if for all  $t > 0$ ,  $e^{tA} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \geq 0$ ;
- *eventually exponentially nonnegative* if there exists  $t_0 \in [0, \infty)$  such that  $e^{tA} \geq 0$  for all  $t > t_0$ . We denote the smallest such nonnegative number by  $t_0 = t_0(A)$  and refer to it  $t_0(A)$  as the exponential index of  $A$ .

We also need the following Lemma from Noutsos and Tsatsomeros (2008):

**Lemma A.1.** *Let  $A \in \mathbb{R}^{N \times N}$ . The following are equivalent:*

1.  $A$  is eventually exponentially nonnegative.
2.  $A + bI$  is eventually nonnegative for some  $b \geq 0$ .
3.  $A^T + bI$  is eventually nonnegative for some  $b \geq 0$ .

We also introduce a definition of an EM-matrix, see Elhashash and Szyld (2008).

**Definition.** An  $N \times N$  matrix  $A = [a_{ij}]$  is called an EM-Matrix if it can be represented as  $A = sI - B$  with  $0 < \rho(B) < s, s > 0$  is some constant,  $\rho(B)$  is the spectral radius of  $B$ , and  $B$  is an eventually nonnegative matrix.

For the following we need two Lemmas.

**Lemma A.2.** *Let  $A \in \mathbb{R}^{N \times N}$ , and  $A = (\alpha^2 - \beta^2)I - 2\beta A_2^B$ . Then  $A$  is an EM-matrix.*

**Proof.** Denote  $d_i$  the  $i$ -th upper diagonal of  $A$ . So  $d_0$  means the main diagonal, etc.

1. First, show that  $2\beta A_2^B$  is an eventually exponentially nonnegative matrix. To see this use representation  $e^{t\beta A_2^B} \equiv [e^{tB}]^{1/(2h)}$  where  $B$  is a lower tridiagonal matrix with all  $d_0$  elements equal to  $3\beta < 0$ , all  $d_{-1}$  elements equal to  $-4\beta > 0$ , and all  $d_{-2}$  elements equal to  $\beta < 0$ . Positivity of  $e^{tB}$  can be verified explicitly at  $t > N$ . The intuition behind that is that the elements on  $d_2$  are small in absolute values as compared with that of  $d_1$ . Taking the square of  $B$  propagates large positive values on  $d_1$  to the diagonal  $d_2$ . Taking the square of  $B^2$  propagates them to  $d_3$ , etc.

From  $h > 0$  it follows that  $e^{2t\beta A_2^B} \geq 0$ , i.e.  $2\beta A_2^B$  is eventually exponentially nonnegative.

According to Lemma A.1, the eventual exponential non-negativity of  $\beta A_2^B$  means that there exists  $b \geq 0$  such that  $\beta A_2^B + bI = \frac{1}{2h}(B + 2hbI)$  is eventually nonnegative for some  $b \geq 0$ . Let us denote  $B_1 = B + 2hbI$  and chose  $b = 3/(2h) + \epsilon$ , where  $\epsilon \ll 1$ .

In practical examples we can choose  $\epsilon = 10^{-6}$ . Then  $d_0(B_1) = \epsilon$ ,  $d_1(B_1) = 2$ ,  $d_2(B_1) = -1$ . It is easy to check that  $B_1^{N+3} \geq 0$ . Again that is because  $d_{-1}(B_1) > 0$ ,  $|d_{-1}(B_1)| > |d_{-2}(B_1)|$ , so taking the square of  $B_1$  propagates large positive values on  $d_{-1}$  to the diagonal  $d_{-2}$ , etc. Thus,  $\beta A_2^B + bI$  with  $b = 3/(2h) + \epsilon$  is the eventually nonnegative matrix.

2. Represent  $A$  as  $A = (\alpha^2 - \beta^2 + b)I - (2\beta A_2^B + bI)$ . Observe, that  $\rho(2\beta A_2^B + bI) = \epsilon$  and  $s = (\alpha^2 - \beta^2 + b) > \epsilon$  as  $|\beta| < \alpha$ . Thus, by definition,  $A$  is an EM-matrix.  $\square$

For the later we will also need this Lemma:

**Lemma A.3.** *The inverse of the matrix  $A \equiv sI - P$  is a nonnegative matrix.*

**Proof.** Observe that all eigenvalues of  $P$  are  $\lambda_i = \epsilon$ ,  $\forall i \in [1, N]$ . Therefore  $\rho(P) = \epsilon$ . Following Le and McDonald (2006) denote  $index_\lambda(A)$  to be the degree of  $\lambda$  as a root of the minimal polynomial of  $A$ . As matrix  $P$  doesn't have zero eigenvalues in its spectrum  $index_0(P) = 0 < 1$ .

Non-negativity of  $A^{-1}$  then follows from the Theorem

**Theorem A.4. [Theorem 4.2 in Le and McDonald (2006)]** *Let  $P$  be an  $N \times N$  irreducible eventually nonnegative matrix with  $index_0(P) \leq 1$ , then there exists  $\mu > \rho(P)$  such that if  $\mu > s > \rho(P)$ , then  $(sI - P)^{-1} \geq 0$ .*

To apply this Theorem choose any positive  $\mu > s$ .

Now we are ready to prove the Proposition 3.1.

**Proof of Proposition 3.2** Recall, that in the Proposition 3.1 the following scheme is proposed

$$J = \delta \left( \sqrt{\alpha^2 - \beta^2} I - \left[ (\alpha^2 - \beta^2) I - 2\beta A_2^B - A_2^C \right]^{1/2} \right)$$

We prove separately each statement of the proposition, namely:

1. The above scheme is  $O(h^2)$  approximation of the operator  $L_R$ ;
2. Matrix  $J$  is the negative of an EM-matrix.

**Proof of (1):** This follows from the fact that  $A_2^C$  is a central difference approximation of the operator  $\nabla^2$  to second order in  $h$ , while  $A_2^B$  is the one-sided second order approximation.

**Proof of (2):** Matrix  $M_1 = (\alpha^2 - \beta^2)I - 2\beta A_2^B$  is an EM-matrix. Matrix  $-A_2^C$  is an M-matrix. The sum

of an EM-matrix and M-matrix is an EM-matrix. Therefore  $M_2 = (\alpha^2 - \beta^2)I - 2\beta A_2^B - A_2^C$  is an EM-matrix. According to properties of M-matrices  $M_2^{1/2}$  is also an EM-matrix. Then,  $-M_2$  is the negative of an EM-matrix. Now adding a diagonal matrix  $M_3 = \sqrt{\alpha^2 - \beta^2}I$  to  $-M_2$  we still obtain the resulting matrix to be the negative of an EM-matrix. That is because  $\beta < 0$ , and thus diagonal elements of  $d_0(M_3) < d_0(M_2)$ . In other words, diagonal elements of  $M_3 - M_2$  are negative. Finally, as  $\delta > 0$  the entire matrix  $J$  is the negative of an EM-matrix. That means that starting with some  $N$  matrix  $e^{\Delta\tau J}$  is positive, and all eigenvalues of  $J$  are negative. Therefore, the spectral norm of the operator  $\mathcal{A} = e^{\Delta\tau J}$  (which is  $\lambda = \max_i |\lambda_i|$ ,  $\forall i \in [1, N]$ ,  $\lambda_i$  are the eigenvalues of  $\mathcal{A}$ ) obeys  $\lambda < 1$ . That means that the proposed scheme is unconditionally stable starting from some  $N$ . That finalizes the proof.  $\square$

### B Proof of Proposition 3.3

This proposition could be proven in a similar way as we did it with Proposition 3.1. The main observation here is that. Suppose we proved Proposition 3.1 with  $\beta = \beta_1 < 0$ . Now choose  $\beta_2 = -\beta_1 > 0$ . Then matrix  $M_2 = (\alpha^2 - \beta_2^2)I - 2\beta_2 A_2^F$  is the transpose of the matrix  $M_1 = (\alpha^2 - \beta_1^2)I - 2\beta_1 A_2^B$  in the previous proof. So this is an upper triangular EM matrix. The remaining proof follows the exactly same steps as in the previous Appendix.  $\square$

### C Proof of Proposition 3.4

By Lemma 4 matrix  $(\alpha^2 - \beta^2)I - 2\beta A_2^B$  is an EM-matrix. So is  $-A_2^C$ . A sum of two EM-matrices is an EM-matrix. Therefore,  $M = [(\alpha^2 - \beta^2)I - 2\beta A_2^B - A_2^C]^{-1}$  is an EM-matrix. Non-negativity of  $M^{-1}$  then follows from Lemma 4. Therefore,  $B_1$  is a nonnegative matrix.

Second order approximation follows from the fact that  $A_2^C$  is the second order central difference approximation of the second derivative, and  $A_2^B$  is the second order one-sided approximation of the first derivative.

The last point to prove is that all eigenvalues  $\lambda_i$  of an  $B_1$  have positive real parts and obey the condition  $|\lambda_i| < 1$ ,  $\forall i \in [1, N]$ . First, argue some intuition behind this. Consider matrices  $M_1 = -2\beta A_2^B$ ,  $M_2 = (\alpha^2 - \beta^2)I - A_2^C$ . On a uniform grid they both are

Toeplitz matrices. It is known that asymptotically at  $N \rightarrow \infty$  the Toeplitz matrices commute, see Gray (2006). For commuting matrices the eigenvalues of the sum are the sum of the eigenvalues. Now,  $M_1$  is a lower triangular matrix with the eigenvalues being the values at the main diagonal, i.e.  $\lambda_i = 3|\beta|/h > 0$ ,  $i \in [1, N]$ . The eigenvalues of  $M_2$  could be represented as  $\lambda_i = (\alpha^2 - \beta^2) + 2/h^2 + \Lambda_i$ ,  $i \in [1, N]$ , where  $\Lambda_i$  are the eigenvalues of the matrix constructed of the first lower and upper diagonals of  $M_2$  while all the other elements vanish. It is known that the eigenvalues of such a matrix are  $\Lambda_i = -\frac{2}{h^2} \cos \frac{i\pi}{N+1}$ . Therefore, the eigenvalues of  $M$  are

$$\begin{aligned} \lambda_i &= (\alpha^2 - \beta^2) + 3\frac{|\beta|}{h} + \frac{4}{h^2} \sin^2 \frac{i\pi}{2(N+1)} \\ &> \alpha^2 - \beta^2 > 0, \quad i \in [1, N]. \end{aligned} \quad (32)$$

Thus, they are positive. Also, based on this inequality the eigenvalues of  $B_1$  are

$$\left( \frac{\alpha^2 - \beta^2}{\lambda_i} \right)^{\Delta\tau(\lambda+1/2)/2} < 1.$$

Thus, the latest statement of the Proposition is asymptotically correct at large  $N$  for an uniform grid. For smaller  $N$  this could not be the case. Note however, that in our numerical experiments  $N = 100$  was sufficient for  $B_1$  to acquire this property.  $\square$

### D Proof of Proposition 3.6

First, we need the following Lemma:

**Lemma D.1.** *Let  $A \in \mathbb{R}^{N \times N}$  be an  $M$ -matrix. By definition an  $M$ -Matrix can be represented as  $A = sI - B$  with  $0 < \rho(B) < s$ ,  $s > 0$  is some constant,  $\rho(B)$  is the spectral radius of  $B$ , and  $B$  is a nonnegative matrix. Then  $\log A$  is an  $M$ -matrix if  $s - \rho(B) > 1$ .*

**Proof.** Represent  $A$  in the form

$$\log A = \log s + \log(I - B/s) = \log s - \sum_{k=1}^{\infty} \frac{1}{is^k} B^k$$

As  $B$  is a nonnegative matrix, the sum in the right hand side of the above equality is also a nonnegative matrix. Hence, all elements of  $\log A$  are nonpositive, except might be those on the main diagonal. Actually, all  $d_0(\log A)$  elements are positive.

To see that observe that

$$\rho \left( \sum_{k=1}^{\infty} \frac{1}{is^k} B^k \right) = \sum_{k=1}^{\infty} \frac{1}{is^k} \rho(B)^k = -\log \left[ 1 - \frac{\rho(B)}{s} \right]$$

and, therefore,

$$\rho(\log A) = \log [s - \rho(B)] > 0,$$

if  $s - \rho(B) > 1$ .  $\square$

**Corollary D.2.** *Let  $A \in \mathbb{R}^{N \times N}$  be an EM-matrix. Then  $\log A$  is an EM-matrix if  $s - \rho(B) > 1$ .*

**Proof.** The proof directly follows the steps in the Proof of the above Lemma with allowance for the definition and properties of an EM-matrix.  $\square$

Now we can prove the Proposition 3.2.2. As same discretization of  $z(\nabla)$  as in Propositions 3.1 is used all we need is to prove two statements:

1. The discretization  $D$  of  $\log \frac{z(\nabla)}{z(0)}$  should be an EM-matrix, if  $D[z(\nabla)]$  is an EM-matrix.
2. The real parts of the eigenvalues  $\lambda_i(D[M])$ ,  $i \in [1, N]$  should be negative, where  $D[M]$  is the discretization of  $M$ .

**Proof of 1.** Consider the eigenvalues of  $Z$  which could be found using Equation (32)

$$\begin{aligned} \lambda_i(Z) &= \delta \left[ (\alpha^2 - \beta^2) + 3\frac{|\beta|}{h} \right. \\ &\quad \left. + \frac{4}{h^2} \sin^2 \frac{i\pi}{2(N+1)} \right]^{1/2}, \quad i \in [1, N]. \end{aligned}$$

From here

$$\rho(Z) = \max_i \lambda_i > \delta \frac{\pi}{h} > 1$$

if  $h < \delta\pi$ .

Matrix  $Z$  by construction is an EM-matrix, see Proof to Proposition 3.2.1. Therefore, it can be represented in the form  $Z = sI - B$ . Then

$$1 < \rho(Z) = s - \rho(B).$$

Now we are under assumptions of the Corollary D.2, therefore  $\log Z$  is an EM-matrix.

As  $\log z(0)I$  is a non-negative diagonal matrix, matrix  $M_z = \log [Z(z(0)I)^{-1}]$  is an EM-matrix. That is because  $Z$  is an EM-matrix, and multiplication of  $Z$  by  $(z(0)I)^{-1}$  (which is a diagonal matrix with the diagonal elements  $1/(\alpha^2 - \beta^2)$ ) changes only the diagonal

elements of  $Z$ . In other words, the diagonal elements of  $M_z$  are  $\lambda_i(Z)/(\alpha^2 - \beta^2) > 1$ .  $\square$

**Proof of 2.** The second property follows from the fact that the eigenvalues of  $D[z(\nabla)]$  are positive, see Proposition 3.2.1. If so, the principal matrix logarithm exists and is well-defined. Based on asymptotic properties of the Toeplitz matrices the eigenvalues of  $D[M]$  are asymptotically equal to the sum of the eigenvalues of  $D[z(0) - z(\nabla)]$  and  $D[-(\lambda + 1/2) \log(z(\nabla)/z(0))]$ , see Proof to Proposition 3.2.1 in Appendix. Also, based on Equation (32) the eigenvalues  $\Lambda_i$  of  $D[z(\nabla)]$  in the leading term are proportional to  $1/h$ . Therefore, in the leading term

$$\lambda_i(D[M]) = -\delta \left( \frac{\bar{\Lambda}_i}{h} + (\lambda + 1/2) \log \frac{\bar{\Lambda}_i}{h(\alpha^2 - \beta^2)} \right)$$

where  $\bar{\Lambda}_i > 0$  is a part of  $\Lambda_i$  which in the leading term doesn't depend on  $h$ . Now observe that the inequality

$$\frac{1}{h} \bar{\Lambda}_i + (\lambda + 1/2) \log \frac{\bar{\Lambda}_i}{h(\alpha^2 - \beta^2)} > 0$$

with  $\kappa = -(\lambda + 1/2) > 0$  could be transformed to

$$\frac{1}{h} \bar{\Lambda}_i - \kappa \log \frac{\bar{\Lambda}_i}{h} > 0$$

if  $b = \log(\alpha^2 - \beta^2) > 0$ . It is always valid if  $\kappa < e$ , valid at  $h < \bar{\Lambda}_i/e$  if  $\kappa = e$ , and at  $h < \bar{\Lambda}_i e^{W(-1/\kappa)}$  at  $\kappa > e$ , where  $W_k(y)$  is the Lambert W-function, see Olver et al. (2010).

If  $b < 0$  we need to consider the inequality

$$\frac{1}{h} \bar{\Lambda}_i - \kappa \log \frac{\bar{\Lambda}_i}{h} > -\kappa b > 0$$

which solution reads

$$h < -\frac{\bar{\Lambda}_i}{\kappa W\left(-\frac{e^{-b^2/\kappa}}{\kappa}\right)} \quad \square$$

Thus in both cases  $b > 0$  and  $b < 0$  there exists an upper boundary on the space step  $h$  which, however, doesn't depend on step in time  $\Delta\tau$ . Therefore, in this sense the proposed scheme is unconditionally stable in  $h$  starting from some  $h$  given in the solutions of the above inequalities. Numerical calculations show that this upper limit is not very restrictive unless we consider an extreme case when  $\alpha \approx |\beta|$ .

Combining all the above we conclude that  $\lambda_i(D[M]) < 0$ ,  $i \in [1, N]$ . Therefore, the eigenvalues

of the operator  $e^M$  are nonnegative and  $\lambda_i(D[e^M]) < 1$ ,  $i \in [1, N]$ .

The last point to prove is that matrix  $M_2 = I + (4\lambda^2 - 1)Z^{-1}$  is a nonnegative matrix with all positive eigenvalues less than one in the absolute value. This follows from the fact that: i)  $Z$  is an EM-matrix, ii) an inverse of the EM-matrix is an eventually nonnegative matrix, iii) all eigenvalues of  $Z$  are positive, therefore so are the eigenvalues of  $Z^{-1}$ ; iv) the eigenvalues of  $Z$  are less than one,  $\lambda(Z)_i > 1$ ,  $\forall i \in [1, N]$ , therefore  $\lambda(Z^{-1})_i < 1$ ,  $\forall i \in [1, N]$ . All this properties were already proven in Proposition 3.4.

The entire statement of the Proposition now follows because the product of two nonnegative matrices is a nonnegative matrix. Also as eigenvalues of both matrices  $M$  and  $M_2$  are positive and less than one, consecutive application of them produces a convergent transformation with the same properties of the eigenvalues of the operator product. This finalizes the prove.

### E Proof of Proposition 3.8

First observe that the operators  $M_n$ ,  $n = 1, 2, \dots$  has the same structure as the operator  $M$  in the beginning of the proof of Proposition 3.2.1. Therefore,  $M_n$  is an EM-matrix, and  $M_n^{-\kappa}$  with  $\kappa > 0$  is the nonnegative matrix.

Second order approximation follows from the fact that  $A_2^C$  is the second order central difference approximation of the second derivative, and  $A_2^B$  is the second order one-sided approximation of the first derivative.

The next point to prove is that all eigenvalues  $\lambda_{n,i}$  of an  $M_n^{-\kappa}$  have positive real parts and obey the condition  $|\lambda_{n,i}| < 1$ ,  $\forall i \in [1, N]$ . This also directly follows from the proof of Proposition 3.4.

Thus, at every  $n$  the map  $M_n^{-\kappa} : z_{n-1}(x, \tau) \rightarrow z_n(x, \tau) = M_n^{-\kappa} z_{n-1}(x, \tau)$  preserves positivity of the vector  $z_n(x, \tau)$  and is convergent in the spectral norm as all  $|\lambda_i|_n < 1$ ,  $\forall i \in [1, N]$ .

### F Proof of Proposition 3.9

The proof is analogous to the previous Proposition with allowance for the fact that matrix  $2abA_2^B$  with  $\beta < 0$  is the transpose of the matrix  $2abA_2^F$  with  $\beta \geq 0$ .