FAST NUMERICAL SOLUTION OF PARABOLIC INTEGRO-DIFFERENTIAL EQUATIONS WITH APPLICATIONS IN FINANCE

By

Ana-Maria Matache

Christoph Schwab

and

Thomas P. Wihler

IMA Preprint Series # 1954

(January 2004)

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

514 Vincent Hall
206 Church Street S.E.
Minneapolis, Minnesota 55455–0436

Phone: 612/624-6066 Fax: 612/626-7370

URL: http://www.ima.umn.edu

FAST NUMERICAL SOLUTION OF PARABOLIC INTEGRO-DIFFERENTIAL EQUATIONS WITH APPLICATIONS IN FINANCE

ANA-MARIA MATACHE, CHRISTOPH SCHWAB, AND THOMAS P. WIHLER

ABSTRACT. We numerically solve parabolic problems $u_t + \mathcal{A}u = 0$ in $(0,T) \times \Omega$, $T < \infty$, where $\Omega \subset \mathbb{R}$ is a bounded interval and $\mathcal{A}$ is a strongly elliptic integro-differential operator of order $\rho \in [0,2]$. A discontinuous Galerkin (dG) discretization in time and a wavelet discretization in space are used. The densely populated matrices in the corresponding linear systems of equations are replaced by sparse ones using appropriate wavelet compression techniques. The linear systems in each time step are solved by an incomplete GMRES iteration. Under these conditions, we show that the complexity of our algorithm is linear (up to some logarithmic terms) in the number of spatial degrees of freedom and present error estimates. Applications to purely discontinuous Lévy processes arising in finance are given.

1. INTRODUCTION

Parabolic, possibly degenerate, evolution problems arise as Fokker-Planck equations for moments of Markovian diffusion processes [15]. We mention here only micro models of polymers [27] and asset pricing problems in mathematical finance: the celebrated Black-Scholes equation is a degenerate diffusion equation.

If the Markov process is not a diffusion, but a jump process, then the infinitesimal generator of the process appearing in the Fokker-Planck equation is in general nonlocal. More precisely, if the jump process is of infinite intensity, the associated infinitesimal generators are pseudo-differential operators of order $\rho \in [0,2]$; see e.g. [16, 17].

The discretization of such operators with standard, Galerkin type schemes, entails densely populated, ill-conditioned stiffness matrices which unduly increase the computational complexity of an implicit time stepping scheme.

Often, the infinitesimal generators of the Markov processes generate analytic evolution semi-groups, even in the case of degenerate spatial operators, so that the solutions are analytic functions of time in the open interval $(0,T]$ ([16, Chapter 4]). This (nonuniform) time analyticity of the solution is not exploited in the classical, low order schemes such as backward differencing or implicit Runge-Kutta methods of fixed order.

In the present paper, we analyze and implement a numerical scheme for the efficient numerical solution of Fokker-Planck equations for Markov processes with jumps. The scheme is based on hp-discontinuous Galerkin (dG, for short) time-stepping to exploit the time analyticity of the solution for $t > 0$ and to cope with the loss of this analyticity at $t = 0$, and on piecewise polynomial functions (of fixed degree) in space. As in [34], the

2000 *Mathematics Subject Classification.* 45K05, 65N30, 60J75, 65M12, 65M60, 65T60.

Key words and phrases. Parabolic equations, integro-differential operators, discontinuous Galerkin methods, wavelets, matrix compression, GMRES, Lévy processes, Markov processes.

Supported in part IHP Network Breaking Complexity of the EC (contract number HPRN-CT-2002-00286) with support by the Swiss Federal Office for Science and Education under grant No. BBW 02.0418.

Supported by the Swiss National Science Foundation.

matrices corresponding to the nonlocal infinitesimal generators are compressed by means of a wavelet basis, thereby reducing complexity of a matrix vector multiplication from N^2 to $O(N \log N)$, where N denotes the number of spatial degrees of freedom. At the same time, the condition numbers of the matrices are shown to be bounded independently of the spatial mesh width and of the time step. We prove that this algorithm allows to compute the whole price vector for European Vanillas on Lévy (and even certain classes of Markov) processes to a mean square accuracy of $O(N^{-1})$ with N degrees of freedom in at most $O(N(\log N)^8)$ operations.

We discuss in detail some aspects of the implementation, based in part on [33], and numerical examples from the pricing of European contracts.

The outline of the paper is as follows: In Section 2, we introduce the parabolic problems under consideration. In Section 3, our applications to finance are given. Section 4 deals with the dG and wavelet discretizations in time and space, respectively. The solution of the linear systems is studied in Section 5. Section 6 contains the numerical experiments. Finally, in Section 7, we summarize our work with some concluding remarks.

2. Abstract Parabolic Evolution Problems

In this section, the parabolic problems studied in this paper are introduced. Moreover, an appropriate functional setting is established.

Let $\Omega = (a, b)$, $-\infty < a < b < \infty$, be a bounded interval in $\mathbb{R}$. For $0 \le \rho \le 2$, define the spaces

$$V = \widetilde{H}^{\rho/2}(\Omega) = \begin{cases} H^{\rho/2}(\Omega) & \text{if } 0 \le \rho < 1 \\ H^{1/2}_{00}(\Omega) & \text{if } \rho = 1 \\ H^{\rho/2}_0(\Omega) & \text{if } 1 < \rho \le 2 \end{cases},$$

where $H^{\rho/2}$, $H^{1/2}_{00}$, and $H^{\rho/2}_0$ are the usual Sobolev spaces; see [1, 22], for example. We have the dense injection $V \overset{d}{\hookrightarrow} L^2(\Omega)$, where $L^2(\Omega)$ is the space of square integrable functions equipped with the usual inner product $(\cdot, \cdot)$, as well as the Gelfand triple $V \overset{d}{\hookrightarrow} L^2(\Omega) \overset{d}{\hookrightarrow} V^*$, where $V^* = H^{-\rho/2}(\Omega)$ denotes the dual space of V. Furthermore, let $\langle \cdot, \cdot \rangle_{V^* \times V}$ be the duality pairing in $V^* \times V$, and $\| \cdot \|$, $\| \cdot \|_V$, $\| \cdot \|_{V^*}$ the norms in $L^2(\Omega)$, V, V^*, respectively.

In the time interval $J = (0, T)$, $0 < T < \infty$, consider the parabolic evolution problem

$$u'(t) + \mathcal{A}u(t) = 0 \qquad \text{in } Q = J \times \Omega \tag{2.1}$$

$$u(0) = u_0 \tag{2.2}$$

with an initial condition $u_0 \in L^2(\Omega)$, and a possibly nonlocal operator $\mathcal{A} \in \mathcal{L}(V, V^*)$ of order ρ, given in the weak form

$$\langle \mathcal{A}u, v \rangle_{V^* \times V} = a(u, v), \qquad \mathcal{A}u \in V^*,\ u, v \in V.$$

Here, $a : V \times V \to \mathbb{R}$ is a continuous bilinear form, i.e. there exists a constant $\alpha > 0$ such that

$$|a(u, v)| \le \alpha \|u\|_V \|v\|_V, \qquad \forall u, v \in V. \tag{2.3}$$

Moreover, we assume that, for some constants $\beta, \gamma > 0$, the form a satisfies a Gårding inequality, i.e.

$$a(u, u) \ge \beta \|u\|^2_V - \gamma \|u\|^2, \qquad \forall u \in V. \tag{2.4}$$

Henceforth, without loss of generality, the bilinear form a is supposed to be coercive on $V \times V$. Indeed, substituting $\widetilde{u} = e^{-\gamma t}u$, results in the equation

$$\widetilde{u}'(t) + (\mathcal{A} + \gamma \mathbb{I})\widetilde{u} = 0 \quad \text{in } J \times \Omega \tag{2.5}$$

$$\widetilde{u}(0) = u_0, \tag{2.6}$$

and by (2.4) the operator $\mathcal{A} + \gamma\mathbb{I}$ is coercive with coercivity constant β.
Problem (2.1)–(2.2) has a unique (weak) solution $u(t)$, and there holds the a priori estimate (see [22])

$$\|u\|_{\mathcal{C}(\overline{J}, L^2(\Omega))} + \|u\|_{L^2(J,V)} + \|u'\|_{L^2(J,V^*)} \leq C\|u_0\|_{L^2(\Omega)}.$$

In this paper, the operator $\mathcal{A}$ is assumed to be the sum of a second order differential operator and of an integro-differential operator with a distributional kernel $k(x, x-y) \in \mathcal{D}'(\Omega \times \Omega)$, with associated bilinear form given by

$$a(u,v) = \langle \mathcal{A}u, v\rangle_{V^* \times V} = \int_\Omega \int_\Omega k(x, x-y)u(x)v(y)\,dx\,dy.$$

The kernels $k(x,z) \in \mathcal{C}^\infty(\Omega \times \mathbb{R} \setminus \{x = y\})$ are assumed to satisfy *Caldéron-Zygmund type estimates*, i.e. $\forall \alpha \in \mathbb{N}_0$, there holds

$$|D_z^\alpha k(x,z)| \leq C|z|^{-1-\rho-\alpha}, \qquad z \neq 0, \tag{2.7}$$

uniformly with respect to $x \in \Omega$.

3. Pricing European Vanillas under Lévy Processes

Parabolic problems of type (2.1)–(2.2) arise in finance in the problem of option pricing. For a general Lévy process X_t [3, 28], rather than the standard Wiener process W_t as price process for the risky underlying, the infinitesimal generator $\mathcal{L}^X$ of the transition semi-group generated by X_t is an integro-differential operator.

Specifically, we consider the problem of pricing a European option with asset price S_t given by the exponential law

$$S_t = S_0 e^{\mu t + X_t},$$

with $\mu \in \mathbb{R}$ and X_t a Lévy process.

For a given pay-off $f(S)$, the value of the European option $V(t,S)$ at $S = S_t$ with maturity T is expressed by the following discounted conditional expectation,

$$V(t,S) = \mathbf{E}^{\mathbb{Q}}[e^{-r(T-t)}f(S_T)|S_t = S], \tag{3.1}$$

where $\mathbb{Q}$ is an equivalent martingale measure and r represents the riskless mean-rate of return of S_t under the measure $\mathbb{Q}$; see [5, 18, 21].

3.1. Parabolic Integro-Differential Equation.

Itô's formula for semi-martingales and the principle of no-arbitrage imply that V defined in (3.1) solves the parabolic integro-differential equation (cf. [26], see also [10, Proposition 1])

$$\begin{aligned} &\frac{\partial V}{\partial t}(t,S) + \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2}(t,S) + rS\frac{\partial V}{\partial S}(t,S) - rV(t,S) \\ &\quad + \int_{\mathbb{R}} \Big(V(t, Se^y) - V(t,S) - S(e^y - 1)\frac{\partial V}{\partial S}(t,S)\Big)\nu_{\mathbb{Q}}(dy) = 0 \end{aligned} \tag{3.2}$$

in $J \times (0, \infty)$, and

$$V(T,S) = f(S) \tag{3.3}$$

in $(0, \infty)$. Here, by $\sigma \geq 0$, we denote the diffusion component of the Lévy process, i.e., X_t can be uniquely decomposed as $X_t = \sigma W_t + Y_t$ with W_t being a standard Brownian motion and Y_t a quadratic pure jump process independent of W_t. Furthermore, $\nu_{\mathbb{Q}}(dy)$ represents the Lévy measure satisfying

$$\int_{\mathbb{R}} \min(1, y^2)\nu_{\mathbb{Q}}(dy) < \infty. \tag{3.4}$$

The Fourier transform of a Lévy process can be expressed explicitly,

$$E_{\mathbb{Q}}[e^{-iuX_t}] = e^{-t\psi(u)},$$

in terms of the Lévy-Khintchine exponent

$$\psi(u) = \frac{\sigma^2}{2}u^2 + i\gamma u + \int_{|x|<1}(1 - e^{-iux} - iux)\nu_{\mathbb{Q}}(dx) + \int_{|x|\geq 1}(1 - e^{-iux})\nu_{\mathbb{Q}}(dx). \tag{3.5}$$

A Lévy process is characterized by the *Lévy triple* $\sigma, \gamma \in \mathbb{R}$ and the Lévy-measure $\nu_{\mathbb{Q}}$.

Henceforth, we assume that $\nu_{\mathbb{Q}}(dy)$ has a density $k_{\mathbb{Q}}$, i.e., $\nu_{\mathbb{Q}}(dy) = k_{\mathbb{Q}}(y)dy$. The Lévy density $k_{\mathbb{Q}}(y)$ describes the activity of jumps of size y in X_t. We say that the Lévy process X_t is of finite activity if $k_{\mathbb{Q}}(y)$ is integrable; otherwise X_t is said to be of infinite activity.

We drop in (3.2) the subscript $\mathbb{Q}$, and switch to logarithmic price $x = \log(S)$ and to time to maturity $t \to T - t$, thereby obtaining

$$\frac{\partial u}{\partial t}(t, x) + \mathcal{L}^X[u](t, x) = 0 \qquad \text{in } J \times \mathbb{R} \tag{3.6}$$

$$u(0, x) = g(x) \quad \text{in } \mathbb{R}, \tag{3.7}$$

where $g(x) := f(e^x)$ and $\mathcal{L}^X = \mathcal{L}^{\text{BS}} + \mathcal{L}^{\text{jump}}$ with

$$\mathcal{L}^{\text{BS}}[v] := -\frac{\sigma^2}{2}v'' + \left(\frac{\sigma^2}{2} - r\right)v' + rv, \tag{3.8}$$

$$\mathcal{L}^{\text{jump}}[v](x) := -\int_{\mathbb{R}}\{v(x+y) - v(x) - (e^y - 1)v'(x)\}k(y)dy. \tag{3.9}$$

We shall use some or all of the following assumptions on the Lévy density k [24]

(A1) *Activity of small jumps*: the characteristic function $\psi_0(u)$ of the pure jump part Y_t of the Lévy process X_t satisfies: there exist constants $c_1, C_+ > 0$ and $\alpha < 2$ such that

$$|\psi_0(\xi) - ic_1\xi| \leq C_+(1 + |\xi|^2)^{\alpha/2}, \qquad \forall \xi \in \mathbb{R}. \tag{3.10}$$

(A2) *Semi-heavy tails*: there are constants $C > 0$, $\eta_- > 0$ and $\eta_+ > 1$ such that

$$\forall |z| > 1: \quad k(z) \leq C \begin{cases} e^{-\eta_-|z|} & \text{if } z < 0, \\ e^{-\eta_+|z|} & \text{if } z > 0. \end{cases} \tag{3.11}$$

(A3) *Smoothness*: for all $\beta \in \mathbb{N}_0$, there exists $C(\beta)$ such that

$$|k^{(\beta)}(z)| \leq C(\alpha)|z|^{-(1+\alpha+\beta)_+}, \qquad \forall z \neq 0. \tag{3.12}$$

In addition, if $\sigma = 0$, we assume that $0 < \alpha < 2$.

(A4) *Boundedness from below of* $k(z)$: there is $C_- > 0$ such that

$$\forall 0 < |z| < 1: \quad \frac{1}{2}(k(-z) + k(z)) \geq \frac{C_-}{|z|^{1+\alpha}}. \tag{3.13}$$

Remark 3.1.

(i) Since $E_{\mathbb{Q}}[S_t] < \infty$, we have that $\int_{|z|>1} e^z k(z)dz < \infty$. This holds for all k satisfying (A1)–(A2), due to $\eta_+ > 1$ which, from now on, is assumed to be fulfilled.
(ii) Assumption (A2) implies that X_t has finite moments of all orders.
(iii) Assumption (A3) is required in the analysis of the wavelet compression of the moment matrix of $k(z)$; however, it only needs to be satisfied for a finite range of β.
(iv) A Lévy process X_t with $\sigma = 0$ is called *pure jump process*. If X_t is a pure jump process, we assume that it is of infinite activity, i.e.

$$k(z) \text{ satisfies (A1)–(A4) with } 0 < \alpha < 2 \qquad (\text{if } \sigma = 0). \tag{3.14}$$

(v) If the Lévy process is of finite activity, we assume $\sigma > 0$ and that $k(z)$ satisfies (A1)–(A3) with $\alpha < 0$.

3.2. **Examples of Lévy Processes.** Different parameterizations of the Lévy measure have been proposed in the literature. We mention here the Variance Gamma (cf. [23]), or shortly VG process, with

$$k(y) = C \begin{cases} |y|^{-1} e^{-\eta_- |y|} & \text{if} \quad y < 0 \\ |y|^{-1} e^{-\eta_+ |y|} & \text{if} \quad y > 0, \end{cases}$$

the so-called CGMY model [6, 8, 9, 19] (also referred to as tempered or truncated tempered stable processes or 'truncated Lévy flights' in physics), with

$$k(y) = C \begin{cases} \dfrac{e^{-\eta_- |y|}}{|y|^{1+\alpha}} & \text{if} \quad y < 0 \\ \dfrac{e^{-\eta_+ |y|}}{|y|^{1+\alpha}} & \text{if} \quad y > 0, \end{cases} \tag{3.15}$$

$0 < \alpha < 2$, the normal inverse Gaussian process (NIG) [2], and hyperbolic and generalized hyperbolic processes [12, 13, 14]. These are all examples of Lévy processes of infinite activity.

The classical Merton Model [25] is based on a drifted Brownian motion with finitely many jumps, i.e., $X_t = \sigma W_t + \sum_{i=1}^{N_t} Y_i$. Here, $\{Y_i\}$ are independent, identically distributed random variables with distribution function $f(z)$, and $\{N_t\}$ is a Poisson process with intensity λ. The Lévy measure is given by $\nu(dz) = k(z)dz$ with $k(z) = \lambda f(z)$. With $f_M(z) = (\sqrt{2\pi}\sigma_M)^{-1} \exp(-(z-\mu_M)^2/(2\sigma_M^2))$, i.e, f_M is the normal distribution with mean μ_M and standard deviation σ_M, Merton's model is a finite intensity Lévy process which satisfies (A1)–(A3) with $\alpha = -\infty$.

In addition, we mention here the model proposed by Kou [20] which assumes that $f_{Kou}(z) = p_+ M \exp(-Mz)\chi_{\mathbb{R}_+}(z) + p_- G \exp(Gz)\chi_{\mathbb{R}_-}(z)$, $p_+ + p_- = 1$. Here, X_t is a finite activity Lévy process with k satisfying (A1)–(A3) for $\alpha = -1$.

3.3. **Variational Setting.** Here, without loss of generality, we only elaborate on the problem of pricing European vanillas. More precisely, we consider European call, $f(S) = \max(S-K, 0)$, and put, $f(S) = \max(K-S, 0)$, options, with $K > 0$ being the so-called strike price. More general pay-offs in (3.3) (with polynomial growth at infinity, for example) could be treated similarly.

Since $(e^x - K)_+$, $(K - e^x)_+ \notin L^2(\mathbb{R})$, we consider weighted Sobolev spaces accounting for the growth of solutions at infinity in order to obtain an appropriate variational formulation of (3.6), (3.7). To this end, denote, for $w \in L^1_{\text{loc}}(\mathbb{R})$, $w' \in L^\infty(\mathbb{R})$, by $H^1_{\pm w}(\mathbb{R})$ the weighted Sobolev space

$$H^1_{\pm w}(\mathbb{R}) := \{v \in L^1_{\text{loc}}(\mathbb{R}) \mid e^{\pm w} v, e^{\pm w} v' \in L^2(\mathbb{R})\}. \tag{3.16}$$

Similarly, let $L^2_{\pm w}(\mathbb{R}) := \{v \in L^1_{\text{loc}}(\mathbb{R}) \mid e^{\pm w}v \in L^2(\mathbb{R})\}$. To the operator $\mathcal{L}^X$ the following bilinear form is associated:

$$a^{\pm w}(\varphi,\psi) := \int_{\mathbb{R}} \mathcal{L}^X[\varphi](x)\psi(x)e^{\pm 2w(x)}\,dx, \qquad \varphi,\psi \in C_0^\infty(\mathbb{R}).$$

Set $g(x) = (e^x - K)_+$ (the case of a put can be treated analogously), and let

$$w(x) = \begin{cases} \omega_-|x| & \text{if } x < 0 \\ \omega_+|x| & \text{if } x > 0. \end{cases} \tag{3.17}$$

Furthermore, assume that $0 \le \omega_- < \eta_-$, $1 < \omega_+ < \eta_+$ and that $\sigma > 0$. Then, by [24, Theorem 3.4], the problem, find $u \in L^2(J; H^1_{-w}(\mathbb{R})) \cap H^1(J; (H^1_{-w}(\mathbb{R}))^*)$ such that

$$\begin{aligned} \frac{\partial}{\partial t}(u(t),v)_{L^2_{-w}(\mathbb{R})} + a^{-w}(u(t),v) &= 0, \quad \forall\, v \in H^1_{-w}(\mathbb{R}) \\ u(0) &= g, \end{aligned}$$

admits a unique solution.

3.4. **Localization.** The numerical solution of (3.6), (3.7) requires the truncation of $\mathbb{R}$ to a bounded computational domain $\Omega_R = (-R, R)$ (note that barrier-like contracts directly lead to the PIDE on the finite interval). In the classical Black-Scholes model [4], i.e., in the absence of $\mathcal{L}^{\text{jump}}$, the localization can be effected by a simple restriction to Ω_R and the localization error may be estimated by local considerations. However, in the presence of the non-local operator $\mathcal{L}^{\text{jump}}$ corresponding to the jump part of the price process $(X_t)_t$, such local arguments do not apply, since the computation of the solution requires information on the pay-off on the whole real line $\mathbb{R}$.

As the transformation $u(t,\cdot) \to e^{-rt}u(t,\cdot + rt)$ reduces (3.6), (3.7) to the case where $r = 0$, we henceforth assume, without loss of generality, that $r = 0$.

Based on the observation that the excess to pay-off value $U = u - g$ decays exponentially as $|x| \to \infty$, we restrict the PIDE for the excess to pay-off value U to Ω_R. Specifically, the excess to pay-off function U solves

$$\frac{\partial U}{\partial t}(t,x) + \mathcal{L}^X[U](t,x) = -\mathcal{L}^X[g] \quad \text{in } J \times \mathbb{R} \tag{3.18}$$

$$U(0,x) = 0 \qquad \text{in } \mathbb{R}. \tag{3.19}$$

Now, rather then solving (3.18), (3.19), we solve the truncated problem in $(0,T) \times \Omega_R$

$$\frac{\partial U_R}{\partial t} + \mathcal{L}_R[U_R] = -\mathcal{L}^X[g]|_{\Omega_R} \quad \text{in} \quad J \times \Omega_R \tag{3.20}$$

$$U_R(t,x) = 0 \quad \text{in} \quad J \times \mathbb{R}\backslash\Omega_R \tag{3.21}$$

$$U_R(0) = 0 \quad \text{in} \quad \Omega_R, \tag{3.22}$$

with $\mathcal{L}_R$ being the restriction of $\mathcal{L}^X$ to Ω_R. Note that, in contrast to the standard Black-Scholes case ($\mathcal{L}^{\text{jump}} = 0$), the non-local operator $\mathcal{L}^{\text{jump}}$ requires the pay-off function g to be specified also outside the domain Ω_R.

By [24, Proposition 3.6], the right-hand side $-\mathcal{L}^X[g]$ is given by

$$-\mathcal{L}^X[g] = \sigma^2/2K\delta_{\ln(K)} - \mathcal{L}^{\text{jump}}[g]$$

with $\mathcal{L}^{\text{jump}}[g] \in C^\infty(\mathbb{R}\backslash \ln(K)) \cap L^1_{\text{loc}}(\mathbb{R})$ decaying exponentially at infinity:

$$-\mathcal{L}^{\text{jump}}[g] \le Ce^{-\eta_\mp|x|} \text{ as } \pm x \to \infty.$$

Choosing the exponent w as in (3.17) with $0 < \omega_+ < \eta_-$ and $0 < \omega_- < \eta_+$, implies that $-\mathcal{L}^X[g] \in (H^1_w(\mathbb{R}))^*$, and by [24, Theorem 3.4], the variational problem, find $U \in L^2(J; H^1_w(\mathbb{R})) \cap H^1(J; (H^1_w(\mathbb{R}))^*)$ such that

$$\frac{d}{dt}(U(t), v)_{L^2_w(\mathbb{R})} + a^w(U(t), v) = -\langle \mathcal{L}^X[g], v\rangle_{(H^1_w(\mathbb{R}))^* \times H^1_w(\mathbb{R})}, \quad \forall\, v \in H^1_w(\mathbb{R}),$$

with

$$U(0) = 0,$$

is uniquely solvable. Consequently, the excess to pay-off function U decays exponentially at infinity, and by [24, Theorem 4.1], the error $U_R - U$ caused by the localization decays exponentially as $R \to \infty$.

3.5. **Parabolic Setting.** We now verify that the localized problem (3.20)–(3.22) can be cast in the variational setting (2.3), (2.4). Let $\mathcal{A} := \mathcal{L}_R$ and $a(\cdot,\cdot)$ denote the bilinear form induced by $\mathcal{A}$:

$$a(\varphi, \psi) := \int_{\Omega_R} \mathcal{A}[\varphi](x)\psi(x)dx, \qquad \forall\, \varphi, \psi \in C_0^\infty(\Omega_R).$$

If $\sigma > 0$, the price process contains a diffusion component and the order of the operator $\mathcal{L}^X$ is 2. Then $a(\cdot,\cdot)$ can be extended continuously to $V \times V$ with $V := H^1_0(\Omega_R)$ and the following Gårding inequality holds: there exist $c_1, c_2 > 0$ such that

$$a(v, v) \geq c_1\|v\|^2_{H^1(\Omega_R)} - c_2\|v\|^2_{L^2(\Omega_R)}, \qquad \forall\, v \in V.$$

In the pure jump case, $\sigma = 0$, the order of $\mathcal{L}^X$ is in general $\max\{1, \alpha\}$ due to the presence of the drift term in $\mathcal{L}^X$.

Proposition 3.2. *Assume that X_t is a pure jump Lévy process (i.e. $\sigma = 0$) with Lévy density k satisfying* (A1)–(A4) *for some $0 < \alpha < 2$. Then, there exist two positive constants $c_1 = c_1(R) > 0$ and $c_2 = c_2(R) > 0$ such that*

$$\forall\, u \in \tilde{H}^{\alpha/2}(\Omega_R) : \; a(u, u) \geq c_1\|u\|^2_{\tilde{H}^{\alpha/2}(\Omega_R)} - c_2\|u\|^2_{L^2(\Omega_R)}, \tag{3.23}$$

i.e., on $\tilde{H}^{\alpha/2}(\Omega_R)$, the bilinear form $a(\cdot,\cdot)$ satisfies a Gårding inequality as in (2.4)*.*

Proof. By density, we may assume $u \in C_0^\infty(\Omega_R)$. Then $(u', u) = 0$. By $\tilde{u}$, we denote the extension of u by zero to $\mathbb{R}$, and by $a^{\text{sym}}(\varphi, \psi) = (a(\varphi, \psi) + a(\psi, \varphi))/2$ the bilinear form corresponding to the symmetric part $\mathcal{L}^{\text{sym}} := \frac{1}{2}(\mathcal{L}^X + (\mathcal{L}^X)^*) = \frac{1}{2}(\mathcal{L}^{\text{jump}} + (\mathcal{L}^{\text{jump}})^*)$ of $\mathcal{L}^X = \mathcal{L}^{\text{jump}}$. Then there holds

$$a(u, u) = a^{\text{sym}}(u, u) = \langle \mathcal{L}^{\text{sym}}\tilde{u}, \tilde{u}\rangle = \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}} (k(x-y) + k(y-x))(\tilde{u}(x) - \tilde{u}(y))^2 dxdy.$$

Due to (A4), we obtain

$$a(u, u) \geq C_- \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{(\tilde{u}(x) - \tilde{u}(y))^2}{|x - y|^{1+\alpha}} dydx = C_- \left(\|u\|^2_{\tilde{H}^{\alpha/2}(\Omega_R)} - \|u\|^2_{L^2(\Omega_R)}\right).$$

□

Remark 3.3. Proposition 3.2 states a Gårding inequality (2.4) in the pure jump case $\sigma = 0$ for all jump intensities $\alpha \in (0, 2)$. However, if the term (u', v) is present, the continuity (2.3) on $V = \tilde{H}^{\alpha/2}$ holds only if $1 \leq \alpha \leq 2$ [24]. Nevertheless, our framework is applicable to (3.15), for all $0 < \alpha < 2$, since the term (u', v) can be removed.

In order to prove (2.3) in the case $\sigma = 0$ and $0 < \alpha < 1$, it is important to remove the ‘drift’ term so that the transformed equation satisfies the Gårding inequality (2.4) in

$\tilde{H}^{\alpha/2}(\Omega_R)$ as well as the continuity estimate (2.3). Let $\mathcal{A}^{\rm jump} := \mathcal{L}^{\rm jump}|_{\Omega_R}$ be the restriction of the integro-differential operator $\mathcal{L}^{\rm jump}$ to Ω_R. By reduction to the 'transformed' variables, we assume again that $r = 0$, together with $0 < \alpha < 1$. Then,

$$\mathcal{A}[u] = -\frac{\sigma^2}{2}u'' + \frac{\sigma^2}{2}u' + \mathcal{A}^{\rm jump}[u]. \tag{3.24}$$

Next, the operator $\mathcal{A}^{\rm jump}$ can be written as a sum of a first order term and an integro-differential operator $\mathcal{C}$ which is continuous and coercive on $\tilde{H}^{\alpha/2}(\Omega_R)$:

$$\langle \mathcal{A}^{\rm jump}\varphi, \psi\rangle = c_1\langle \varphi', \psi\rangle + \langle \mathcal{C}[\varphi], \psi\rangle, \quad \forall\, \varphi, \psi \in C_0^\infty(\Omega_R), \tag{3.25}$$

with

$$c_1 := \int_{\mathbb{R}} (e^y - 1)k(y)dy.$$

The operator $\mathcal{C}$ is defined by

$$\langle \mathcal{C}u, v\rangle := -\int_{\mathbb{R}}\int_{\mathbb{R}} k(y-x)(\tilde{u}(y) - \tilde{u}(x))\tilde{v}(x)dydx, \quad \forall\, u, v \in \tilde{H}^{\alpha/2}(\Omega_R),$$

where, by $\tilde{v}$, we denote the extension by zero of v to $\mathbb{R}$. Then, there exist constants $C, C_1, C_2 > 0$ such that

$$\begin{aligned} \langle \mathcal{C}u, v\rangle &\le C\|u\|_{\tilde{H}^{\alpha/2}(\Omega_R)}\|v\|_{\tilde{H}^{\alpha/2}(\Omega_R)}, && \forall\, u, v \in \tilde{H}^{\alpha/2}(\Omega_R),\\ \langle \mathcal{C}u, u\rangle &\ge C_1\|u\|^2_{\tilde{H}^{\alpha/2}(\Omega_R)} - C_2\|u\|^2_{L^2(\Omega_R)}, && \forall\, u \in \tilde{H}^{\alpha/2}(\Omega_R). \end{aligned}$$

We have seen that the presence of the drift term $c_1\varphi'$ obstructs the continuity of the form $\langle \mathcal{A}^{\rm jump}[\varphi], \psi\rangle$. We remove this term by the transformation

$$U(t,x) = V(t, x - (\sigma^2/2 + c_1)t), \tag{3.26}$$

which yields the parabolic problem

$$\begin{aligned} \frac{\partial V}{\partial t}(t,x) - \frac{\sigma^2}{2}\frac{\partial^2 V}{\partial x^2} + \mathcal{C}[V](t,x) &= \left(-\mathcal{L}^X[g]|_{\Omega_R}\right)(x + (\sigma^2/2 + c_1)t)\\ V(0,x) &= 0. \end{aligned}$$

Remark 3.4. We note that the numerical scheme proposed in this paper exploits the parabolic nature of (3.6) and, furthermore, that a dominant first order term may be a source of instabilities even in the case $\sigma \neq 0$. Therefore, the removal of the 'drift' is also of interest from the point of view of numerical approximation.

Due to our previous considerations, $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ satisfies (2.3), (2.4) (cf. Proposition 3.2) in the triple $V \overset{d}{\hookrightarrow} L^2(\Omega_R) \overset{d}{\hookrightarrow} V^*$ with

$$V := \tilde{H}^{\rho/2}(\Omega_R),$$

where the order ρ is given by

$$\rho = \begin{cases} 2 & \text{if } \sigma > 0\\ \alpha & \text{if } \sigma = 0. \end{cases} \tag{3.27}$$

For any $v \in H^{\rho/2}$, let $P_a^V v$ denote the V-projection of v with respect to $a(\cdot,\cdot)$, i.e. $a(v - P_a^V v, w) = 0$, for all $w \in V$. By the continuity of $a(\cdot,\cdot)$, it holds that

$$\|P_a^V v\|_V \le \|v\|_{H^{\rho/2}(\Omega_R)}, \qquad \forall\, v \in H^{\rho/2}(\Omega_R). \tag{3.28}$$

The variational formulation of (3.20)–(3.22) reads: find $U_R \in L^2(J,V) \cap H^1(J,V^*)$ such that

$$\left(\frac{\partial U_R}{\partial t}(t,\cdot), v\right)_{L^2(\Omega_R)} + a(U_R(t,\cdot),v) = -\langle \mathcal{L}^X[g], v\rangle_{V^*\times V}, \qquad \forall\, v \in V, \tag{3.29}$$

$$U_R(0,\cdot) = 0. \tag{3.30}$$

Obviously, (3.29), (3.30) is equivalent to solving

$$\left(\frac{\partial u_R}{\partial t}(t,\cdot), v\right)_{L^2(\Omega_R)} + a(u_R(t,\cdot),v) = 0 \qquad \forall\, v \in V \tag{3.31}$$

$$u_R(0,\cdot) = u_0 := P_a^V g, \tag{3.32}$$

where, evidently, $u_R = P_a^V g + U_R$.

Hence, dropping the subscript R to simplify notations and referring to (3.31), (3.32), we are now in the setting of the previous Section 2 with $u_0 \in V$.

3.6. **Non-Translation-Invariant Operators.** The numerical method presented in this paper (see Sections 4–6) applies also to the more general case where the log-price process X_t is a Markov process with increments $X_t - X_s$ that are no longer independent of X_s for $t > s$. In particular, we allow for volatilities $\sigma(x)$ and singular jump measures that depend on x. Instead of the translation-invariant operator $\mathcal{L}^X$ from (3.8)–(3.9) we now consider a more general operator $\mathcal{L}^X$

$$\mathcal{L}^X[\varphi](x) := -\frac{\sigma(x)^2}{2}\varphi''(x) + \left(\frac{\sigma(x)^2}{2} - r\right)\varphi'(x) + r\varphi + \mathcal{L}^{\text{jump}}[\varphi](x), \tag{3.33}$$

with the integral operator $\mathcal{L}^{\text{jump}}$ given by

$$\mathcal{L}^{\text{jump}}[\phi](x) := -\int_{\mathbb{R}} \{\phi(y) - \phi(x) - \left(e^{y-x} - 1\right)\phi'(x)\} k(x, y-x)\, dy. \tag{3.34}$$

We assume

$$0 < \sigma_0 \le \sigma(x) \le \sigma_1 \quad \forall\, x \in \Omega_R,$$

and that $k(x,z)$ satisfies the assumptions (A1)–(A2) uniformly with respect to $x \in \Omega_R$. Instead of (A3), we require the Calderón-Zygmund estimates: for all $\beta,\gamma \in \mathbb{N}_0$, there holds

$$|\partial_x^\gamma \partial_z^\beta k(x,z)| \le C(\gamma,\beta)|z|^{-(1+\alpha+\beta+\gamma)}, \qquad z \neq 0. \tag{3.35}$$

In the case $\sigma = 0$, we require (A4) uniformly with respect to $x \in \Omega_R$. If $\sigma > 0$ or $\alpha \ge 1$ the corresponding bilinear form a_R satisfies (2.3), (2.4). If $\sigma = 0$ and $\alpha < 1$, we need to assume that we can transform the problem such that a_R satisfies (2.3).

4. Numerical Solution

The aim of this part is to discretize the parabolic equation (2.1), (2.2) in time and space using a discontinuous Galerkin (dG) scheme on $J = (0,T)$, and a wavelet finite element method on $\Omega = (a,b)$. Again, $\mathcal{A} : V \to V^*$ in (2.1) is a possibly nonlocal operator.

In the ensuing analysis, we shall need to consider functions in V with additional regularity. For this reason, we introduce the spaces $\mathcal{H}^s(\Omega)$, $s \ge \rho/2$, given by

$$\mathcal{H}^s(\Omega) = \begin{cases} V = \widetilde{H}^{\rho/2}(\Omega) & \text{for } s = \rho/2 \\ V \cap H^s(\Omega) & \text{for } s > \rho/2. \end{cases}$$

From now on, in order to cope with the assumptions in the previous Section 3, the initial condition is supposed to fulfill $u_0 \in V$.

4.1. **Spatial Semi-Discretization.** Let $\mathcal{T}^0$ be a fixed coarse partition of Ω. Furthermore, define the mesh $\mathcal{T}^l$, for $l > 0$, by bisection of each interval in $\mathcal{T}^{l-1}$. We assume that the mesh $\mathcal{T}_h$ is obtained in this way as $\mathcal{T}^L$, for some $L > 0$, such that $h = C2^{-L}$.

For $0 \leq \rho < 1$, the space V_h is defined as the space of piecewise, possibly discontinuous polynomials of degree $p \geq 0$ on $\mathcal{T}_h$. For $1 \leq \rho \leq 2$, V_h is the space of all continuous piecewise polynomials of degree $p \geq 1$ on the triangulation $\mathcal{T}_h$ which vanish on the boundary $\partial\Omega$.

In the same way, we define the spaces V^l corresponding to the triangulation $\mathcal{T}^l$, so that we have $V^0 \subset V^1 \subset \cdots \subset V^L = V_h$. Let $N^l = \dim V^l$ and $N = \dim V_h = N^L = C2^L$.

The semi-discrete problem corresponding to (2.1), (2.2) reads: Given $u_0 \in V$, find $u_h \in H^1(J, V_h)$ such that

$$\frac{d}{dt}(u_h, v_h) + a(u_h, v_h) = 0, \qquad \forall\, v_h \in V_h, \tag{4.1}$$

and

$$u_h(0) = P_a^{V_h} u_0. \tag{4.2}$$

Here, $P_a^{V_h}$ is the projection onto V_h with respect to the bilinear form a, i.e. for $v \in V$, $P_a^{V_h} v \in V_h$ is defined by

$$a(v - P_a^{V_h} v, w) = 0, \qquad \forall w \in V_h. \tag{4.3}$$

4.1.1. *Wavelet basis.* In order to compute a fully discrete approximation (in time and space) to the parabolic problem (2.1), (2.2), systems of linear equations have to be solved in each time step. Due to the nonlocal character of the operator $\mathcal{A}$ the matrices corresponding to these linear systems are typically fully populated. Therefore, matrix compression techniques, transforming the dense matrices into sparse ones, have to be employed. The compressed systems may then be manipulated in linear complexity.

Using an appropriate basis for the finite dimensional space V_h, we will be able to represent the bilinear form $a(\cdot,\cdot)$ as a matrix whose entries are mostly negligible, thereby allowing for matrix compression techniques.

More precisely, we consider so-called biorthogonal wavelets

$$\{\psi_j^l\}_{j,l}, \qquad l = 0, 1, \ldots, \quad j = 1, 2, \ldots, M^l, \tag{4.4}$$

which allow for optimal preconditioning. They satisfy the following properties (see e.g. [7]):

1. $V^L = \operatorname{span}\{\psi_j^l \mid 0 \leq l \leq L, 1 \leq j \leq M^l\}$.
2. The function ψ_j^l has support $S_j^l = \operatorname{supp} \psi_j^l$ of diameter bounded by $C2^{-l}$.
3. Wavelets ψ_j^l with $\overline{S_j^l} \cap \Omega = \emptyset$ have vanishing moments up to order p, i.e., $(\psi_j^l, q) = 0$ for all polynomials q of degree p or less.
4. The functions ψ_j^l for $l \geq l_0$ are obtained by scaling and translation of the functions $\psi_j^{l_0}$.
5. For all $v_h = \sum_{l=0}^{L} \sum_{j=1}^{M^l} v_j^l \psi_j^l \in V_h$, there holds the norm equivalence

$$|||v_h|||_s^2 := \sum_{l=0}^{L} \sum_{j=1}^{M^l} |v_j^l|^2 2^{2ls} \sim \|v\|_{\widetilde{H}^s(\Omega)}^2, \tag{4.5}$$

for $0 \leq s \leq 1$.

Remark 4.1. Our bi-orthogonal wavelets in the case $p = 1$ are continuous, piecewise linear spline wavelets vanishing outside $\Omega = (0, 1)$ (for more general domains $\Omega = (a, b)$, they are obtained by simple scalings). The interior wavelets have two vanishing moments and are obtained from the mother wavelet $\psi(x)$ which takes the values $(0, -1/2, 1, -1/2, 0)$ at $(0, 1/4, 1/2, 3/4, 1)$ by scaling and translations: $\psi_j^l(x) := \psi(2^{l-1}x - (2j-1)2^{-2})$ for $1 \leq j \leq 2^l - 2$ and $l \geq 2$. The boundary wavelets are constructed from the continuous, piecewise linear functions ψ_*, with values $(0, 1, -1/2, 0)$ at $(0, 1/4, 1/2, 3/4)$, and ψ^*, taking values $(0, -1/2, 1, 0)$ at $(1/4, 1/2, 3/4, 1)$: $\psi_0^l = \psi_*(2^{l-1}x)$ and $\psi_{2^l-1}^l = \psi^*(2^{l-1}x - 2^{l-1} + 1)$. Higher order spline wavelets with analogous properties have been given in [11].

4.1.2. *Matrix Compression.* The bilinear form a on $V_h \times V_h$ corresponds to a matrix $\mathbf{A}$ with entries $A_{(l,j),(l',j')} = a(\psi_{j'}^{l'}, \psi_j^l)$ which, due to the Caldéron-Zygmund estimates (2.7), decay with increasing distance of their supports. Hence, we can define a compressed matrix $\tilde{\mathbf{A}}$ and a corresponding bilinear form $\tilde{a}$ by replacing some of the small entries in $\mathbf{A}$ with zero:

$$\tilde{A}_{(j,l),(j',l')} := \begin{cases} A_{(j,l),(j',l')} & \text{if } \operatorname{dist}(S_j^l, S_{j'}^{l'}) \leq \delta_{l,l'} \text{ or } S_j^l \cap \partial\Omega \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{4.6}$$

Here, the truncation parameters $\delta_{l,l'}$ are given by

$$\delta_{l,l'} := c \max\{2^{-L+\hat{\alpha}(2L-l-l')}, 2^{-l}, 2^{-l'}\}, \tag{4.7}$$

with some parameters $c > 0$ and $1 \geq \hat{\alpha} > 0$, and $S_j^l = \operatorname{supp} \psi_j^l$.

By continuity and coercivity of the bilinear form a and using (4.5), we have the norm equivalence

$$\|u\|_a := (a(u, u))^{1/2} \sim \|u\|_V.$$

In [34, Proposition 3.2], the following was proved:

Proposition 4.2. *Assume* c *in* (4.7) *is sufficiently large. Then there exists* $0 < \widetilde{\alpha} < \infty$ *independent of* L *such that*

$$|\widetilde{a}(u_h, v_h)| \leq \widetilde{\alpha}\|u_h\|_a\|v_h\|_a, \tag{4.8}$$

and there exists $0 < \widetilde{\beta} \leq \widetilde{\alpha}$ *such that*

$$\widetilde{a}(u_h, u_h) \geq \widetilde{\beta}\|u_h\|_a^2, \tag{4.9}$$

for all $u_h, v_h \in V_h$.

Moreover, we have [32, 34]:

Proposition 4.3. *There exists* $0 < \delta < 1$ *independent of* L *such that for all* $L > 0$ *the following consistency condition is satisfied:*

$$|a(u_h, v_h) - \widetilde{a}(u_h, v_h)| \leq \delta\, \|u_h\|_a\, \|v_h\|_a, \qquad \forall u_h, v_h \in V_h. \tag{4.10}$$

In addition, if

$$\hat{\alpha} \geq \frac{2p+2}{2p+2+\rho}, \tag{4.11}$$

then, for all $u_h, v_h \in V_h$*, there holds*

$$|a(u_h, v_h) - \widetilde{a}(u_h, v_h)| \leq Ch^{p+1-\rho/2}|\log h|^\nu |||u_h|||_{p+1}\, \|v_h\|_V, \tag{4.12}$$

with $\nu = 1$ *if equality holds in* (4.11)*, and* $\nu = 0$ *otherwise.*

The matrix compression (4.6) reduces the number of nonzero elements from N^2 in $\mathbf{A}$ to N times a logarithmic term in $\tilde{\mathbf{A}}$; see [29].

Proposition 4.4. *$\widehat{\alpha}$ in* (4.11) *can be chosen in such a way that $\nu = 0$ in* (4.12)*. For $\widehat{\alpha} < 1$, the number of nonzero elements in $\tilde{\mathbf{A}}$ is $\mathcal{O}(N \log N)$. If $\widehat{\alpha} = 1$, then the number of nonzero elements in $\tilde{\mathbf{A}}$ is $\mathcal{O}(N(\log N)^2)$.*

4.1.3. *Perturbed Spatial Semi-Discretization.* In accordance with the matrix compression from the previous section, we consider, instead of (4.1), (4.2), the perturbed spatial semi-discretization of (2.1), (2.2): find $\widetilde{u}_h \in V_h$ such that

$$\frac{d}{dt}(\widetilde{u}_h, v_h) + \widetilde{a}(\widetilde{u}_h, v_h) = 0, \qquad \forall\, v_h \in V_h. \tag{4.13}$$

and

$$\widetilde{u}_h(0) = P_a^{V_h} u_0. \tag{4.14}$$

Assumption 4.5. *We assume that the error between the exact solution u of* (2.1), (2.2) *and the approximate solution $\widetilde{u}_h$ obtained by* (4.13), (4.14) *at $t = T$ is of a certain algebraic order, i.e. there exist constants C, β (depending on p) such that*

$$\|u(T) - \widetilde{u}_h(T)\| \leq Ch^{\beta}. \tag{4.15}$$

Here, $h = C2^{-L}$ is the mesh width.

4.2. **Discontinuous Galerkin Time Discretization.** We discretize (4.13), (4.14) in time using a discontinuous Galerkin method following [31]. For $0 < T < \infty$ and $M \in \mathbb{N}$, let $\mathcal{M} = \{I_m\}_{m=1}^M$ be a partition of $J = (0, T)$ into M subintervals $I_m = (t_{m-1}, t_m)$, $m = 1, 2, \ldots, M$ with $0 = t_0 < t_1 < t_2 < \ldots < t_M = T$. Moreover, denote by $k_m = t_m - t_{m-1}$ the length of I_m.

For $u \in H^1(\mathcal{M}, V_h) := \{v \in L^2(J, V_h) : v|_{I_m} \in H^1(I_m, V_h), m = 1, 2, \ldots, M\}$, define the one-sided limits

$$u_m^+ := \lim_{s \to 0^+} u(t_m + s), \qquad m = 0, 1, \ldots, M-1,$$
$$u_m^- := \lim_{s \to 0^+} u(t_m - s), \qquad m = 1, 2, \ldots, M,$$

and the jumps

$$[u]_m := u_m^+ - u_m^-, \qquad m = 1, 2, \ldots, M-1.$$

In addition, to each time interval I_m, a polynomial degree (approximation order) $r_m \geq 0$ is associated. These numbers are stored in the degree vector $\underline{r} = \{r_m\}_{m=1}^M$. Then, the following space being used for the discontinuous Galerkin (dG) method is introduced:

$$\mathcal{S}^{\underline{r}}(\mathcal{M}, V_h) := \{u \in L^2(J, V_h) : u|_{I_m} \in \mathcal{P}_{r_m}(I_m, V_h), m = 1, 2, \ldots, M\},$$

where $\mathcal{P}_{r_m}(I_m)$ denotes the space of polynomials of degree at most r_m on I_m.

With these definitions, the fully discrete dG scheme for the solution of problem (2.1), (2.2), respectively (4.13), (4.14) reads as follows: find $\widetilde{U}_h^{dG} \in \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h)$ such that for all $W \in \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h)$, there holds

$$\widetilde{B}_{dG}(\widetilde{U}_h^{dG}, W) = F_{dG}(W), \tag{4.16}$$

where

$$\widetilde{B}_{dG}(U, W) := \sum_{m=1}^{M} \int_{I_m} \big((U', W) + \widetilde{a}(U, W)\big)\, dt + \sum_{m=1}^{M-1} ([U]_m, W_m^+) + (U_0^+, W_0^+), \tag{4.17}$$

and

$$F_{dG}(W) := (P_a^{V_h} u_0, W_0^+). \tag{4.18}$$

Proposition 4.6. *(see* [30]*)*

(a) *The solution* $\widetilde{U}_h^{dG} \in \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h)$ *of the dG method* (4.16) *is uniquely defined.*
(b) *The solution* $\widetilde{u}_h$ *of* (4.13), (4.14) *satisfies the dG formulation* (4.16)*, i.e. we have the Galerkin orthogonality*

$$\widetilde{B}_{dG}(\widetilde{u}_h - \widetilde{U}_h^{dG}, W) = 0, \tag{4.19}$$

for all $W \in \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h)$.

Remark 4.7. The dG method (4.16) can be interpreted as a time stepping scheme of variable step size k_m and orders r_m. Indeed, assuming that $\widetilde{U}_h^{dG}$ is known on the time intervals $I_m = (t_{m-1}, t_m)$, $m = 1, 2, \ldots, n-1$, we may find $\widetilde{U}_h^{dG} \in \mathcal{P}_{r_n}(I_n, V_h)$, $1 \le n \le M$, by solving the variational problem

$$\int_{I_n} \left((\partial_t \widetilde{U}_h^{dG}, W) + \widetilde{a}(\widetilde{U}_h^{dG}, W) \right) dt + (\widetilde{U}_{n-1}^{dG+}, W_{n-1}^+) = (\widetilde{U}_{n-1}^{dG-}, W_{n-1}^+), \tag{4.20}$$

for all $W \in \mathcal{P}_{r_n}(I_n, V_h)$. Here, we set $\widetilde{U}_0^{dG-} = P_a^{V_h} u_0$.

4.2.1. *Stability.* For $v_h \in V_h$ and $U \in \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h)$, we introduce the norms

$$\|v_h\|_{\widetilde{a}} = \left(\widetilde{a}(v_h, v_h)\right)^{1/2},$$

and

$$\|U\|_{dG}^2 = \sum_{m=1}^{M} \int_{I_m} \|U\|_{\widetilde{a}}^2 dt + \frac{1}{2}\Big(\|U_0^+\|^2 + \sum_{m=1}^{M-1} \|[U]_m\|^2 + \|U_M^-\|^2\Big).$$

The dG time-stepping is stable:

Proposition 4.8. *The solution* $\widetilde{U}_h^{dG} \in \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h)$ *of the dG method* (4.16) *satisfies the stability estimate*

$$\|\widetilde{U}_h^{dG}\|_{dG} \le C \|u_0\|_V,$$

with a constant $C > 0$ *independent of* T *and of* L.

Proof. Let $U \in \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h)$ be arbitrary. Then, using that

$$\begin{aligned}\sum_{m=1}^{M} \int_{I_m} (U', U)\, dt &= \frac{1}{2} \sum_{m=1}^{M} \int_{I_m} \frac{d}{dt} \|U\|^2\, dt \\ &= \frac{1}{2}\Big(\|U_M^-\|^2 + \sum_{m=1}^{M-1} \left(\|U_m^-\|^2 - \|U_m^+\|^2\right) - \|U_0^+\|^2\Big),\end{aligned}$$

yields

$$\begin{aligned}\widetilde{B}_{dG}(U, U) = \sum_{m=1}^{M} \int_{I_m} \widetilde{a}(U, U)\, dt &+ \frac{1}{2}\left(\|U_0^+\|^2 + \|U_M^-\|^2\right) \\ &+ \frac{1}{2} \sum_{m=1}^{M-1} \left(2([U]_m, U_m^+) + \|U_m^-\|^2 - \|U_m^+\|^2\right).\end{aligned}$$

Since, for $m = 1, 2, \ldots, M-1$, there holds

$$2([U]_m, U_m^+) + \|U_m^-\|^2 - \|U_m^+\|^2 = \|[U]_m\|^2,$$

we obtain

$$\widetilde{B}_{dG}(U, U) = \|U\|_{dG}^2, \qquad \forall U \in \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h). \tag{4.21}$$

Moreover,

$$|F_{dG}(U)| \le |(P_a^{V_h} u_0, U_0^+)| \le \|P_a^{V_h} u_0\| \|U_0^+\|.$$

Using the stability of $P_a^{V_h}$, that is

$$\|P_a^{V_h} v\|_V \le \|v\|_V, \qquad \forall v \in V,$$

results in

$$|F_{dG}(U)| \le C\|u_0\|_V \|U_0^+\| \le C\|u_0\|_V \|U\|_{dG}. \tag{4.22}$$

Now, putting $U := \widetilde{U}_h^{dG}$ and using (4.21), (4.22) and the dG formulation (4.16), we finally get

$$\|\widetilde{U}_h^{dG}\|_{dG}^2 \le |\widetilde{B}_{dG}(\widetilde{U}_h^{dG}, \widetilde{U}_h^{dG})| = |F_{dG}(\widetilde{U}_h^{dG})| \le C\|u_0\|_V \|\widetilde{U}_h^{dG}\|_{dG}.$$

Dividing both sides by $\|\widetilde{U}_h^{dG}\|_{dG}$ completes the proof. ☐

4.3. **Convergence of the Fully Discrete Scheme.** The solution operator of the parabolic problem (2.1), (2.2) generates an analytic semi-group (see e.g. [30]), and therefore, the solution $u(t)$ becomes analytic with respect to t for $t > 0$. However, due to the non-smoothness of the initial data, singularities may arise at $t = 0$. In this section, we show how, by the use of so-called geometric time partitions and linearly increasing polynomial degrees in the time discretization, the low regularity of the solution at $t = 0$ can be resolved.

Definition 4.9. A partition $\mathcal{M}_{M,\gamma} = \{I_m\}_{m=1}^M$ in $J = (0, T)$, $T > 0$, is called geometric with M time steps $I_m = (t_{m-1}, t_m)$, $m = 1, 2, \ldots, M$, and grading factor $\gamma \in (0, 1)$, if

$$t_0 = 0, \qquad t_m = T\gamma^{M-m}, \quad 1 \le m \le M.$$

Definition 4.10. A polynomial vector $\underline{r}$ is called linear with slope $\mu > 0$ on $\mathcal{M}_{M,\gamma}$, if $r_1 = 0$ and $r_m = \lfloor \mu m \rfloor$, $m = 2, \ldots, M$, where $\lfloor \mu m \rfloor = \max\{q \in \mathbb{N}_0 : q \le \mu m\}$.

We have the following a priori error estimate on the fully discrete dG method. Its proof is worked out in the Appendix.

Theorem 4.11. *Let $\rho \in [0, 2]$, and u be the solution of the parabolic problem* (2.1), (2.2) *on $J \times \Omega = (0, T) \times \Omega$, with $u_0 \in V = \widetilde{H}^{\rho/2}(\Omega)$. Moreover, assume* (4.15) *and consider the geometric partition $\mathcal{M}_{M,\gamma}$ from Definition 4.9 with $0 < \gamma < 1$ and $M = \mathcal{O}(|\log h|)$, and the linear polynomial degree vector $\underline{r}$ with slope $\mu > 0$ from Definition 4.10. Then, the fully discrete dG method* (4.16) *on the finite element space $\mathcal{S}^{\underline{r}}(\mathcal{M}_{M,\gamma}, V_h)$ satisfies the a priori error estimate*

$$\|u(T) - \widetilde{U}_h^{dG}(T)\| \le Ch^\beta. \tag{4.23}$$

Here, h is the mesh size of the spatial mesh $\mathcal{T}_h$, β is the constant from (4.15)*, and $C > 0$ is a constant independent of h.*

5. Solution Algorithm

We will now study the linear systems resulting from the dG method (4.16). We show that they may be solved iteratively, without causing a loss in the rates of convergence in the error estimate (4.23), by the use of an incomplete GMRES method. Furthermore, we prove, that the overall complexity of the fully discrete dG method is linear (up to some logarithmic terms). Since these matters have already been studied in [33, Section 5], we restrict ourselves to a brief presentation of the results there.

5.1. **Derivation of the Linear Systems.** The dG time stepping scheme (4.20) corresponds to a linear system of size $(r_m+1)N^L$ to be solved in each time step $m = 1, 2, \dots, M$. Here and in what follows, in order to clarify the dependence on the refinement level L explicitly, we denote by N^L the dimension of the finite element space $V_h = V^L$.

For $1 \le m \le M$, let $\{\phi_j = \sqrt{j+1/2}L_j\}_{j=0}^{r_m}$, where L_j is the j-th Legendre polynomial on $(-1,1)$ (normalized such that $L_j(1) = 1$), be a basis of the polynomial space $\mathcal{P}_{r_m}(-1,1)$. Then, the temporal shape functions on the time interval I_m are given by $\phi_j \circ F_m^{-1}$, where the mapping $F_m : (-1,1) \to I_m$ is given by

$$t = F_m(\hat{t}) = \frac{1}{2}(t_{m-1}+t_m) + \frac{1}{2}k_m\hat{t}, \qquad k_m = t_m - t_{m-1}, \qquad \hat{t} \in (-1,1).$$

Writing $\widetilde{U}^{dG}_{h,m}(x,t) = \widetilde{U}^{dG}_h|_{I_m}(x,t)$ and $W_m = W|_{I_m}$ in (4.20) as

$$\widetilde{U}^{dG}_{h,m}(x,t) = \sum_{j=0}^{r_m} \widetilde{U}^{dG}_{h,m,j}(x)(\phi_j \circ F_m^{-1})(t),$$
$$W_m(x,t) = \sum_{j=0}^{r_m} W_{m,j}(x)(\phi_j \circ F_m^{-1})(t),$$

the variational formulation (4.20) reads: find $(\widetilde{U}^{dG}_{h,m,j})_{j=0}^{r_m} \in (V_h)^{\underline{r}}$ such that there holds for all $(W_{m,i})_{i=0}^{r_m} \in (V_h)^{\underline{r}}$,

$$\sum_{i,j=0}^{r_m} C_{ij}(\widetilde{U}^{dG}_{h,m,j}, W_{m,i})_{L^2(\Omega)} + \frac{k_m}{2}\sum_{i=0}^{r_m} \widetilde{a}(\widetilde{U}^{dG}_{h,m,j}, W_{m,i}) = \sum_{i=0}^{r_m} f_{m,i}(W_{m,i}), \tag{5.1}$$

where, for $i, j = 0, 1, \dots, r_m$,

$$C_{ij} = \sigma_{ij}\sqrt{(i+1/2)(j+1/2)}, \quad \sigma_{ij} = \begin{cases} (-1)^{i+j} & \text{if } j > i \\ 1 & \text{else} \end{cases}, \tag{5.2}$$

and $f_{m,i}(v) = \phi_i(-1)\big(\widetilde{U}^{dG-}_{h,m-1}(t_{m-1}), v\big)$.

Henceforth, for the sake of readability, we will drop the subscript m. Denoting by $\mathbf{M}$ and $\widetilde{\mathbf{A}}$ the mass and (compressed) stiffness matrix on $V^h = V^L$ with respect to $(\cdot,\cdot)$ and $\widetilde{a}(\cdot,\cdot)$, respectively, (4.20) takes the matrix form:

$$\mathbf{R}\underline{u} = \underline{f} \quad \text{with} \quad \mathbf{R} = \mathbf{C} \otimes \mathbf{M} + \frac{k}{2}\mathbf{I} \otimes \widetilde{\mathbf{A}}, \tag{5.3}$$

where $\underline{u}$ denotes the coefficient vector of $\widetilde{U}^{dG}_{h,m} = \widetilde{U}^{dG}_h|_{I_m} \in \mathcal{P}_{r_m}(I_m, V_h)$.

5.2. **Decoupling.** In [30] it has been found that the system (5.3) of size $(r+1)N^L$ can be reduced to solving $r+1$ linear systems of size N^L. To this end, let $\mathbf{C} = \mathbf{Q}\mathbf{T}\mathbf{Q}^{\mathsf{H}}$ be the Schur decomposition of the $(r+1)\times(r+1)$ matrix $\mathbf{C}$ with a unitary matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{T}$. Note that the diagonal of $\mathbf{T}$ contains the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_{r+1}$ of $\mathbf{C}$. Then, multiplying (5.3) by $\mathbf{Q}^{\mathsf{H}} \otimes \mathbf{I}$ from the left results in the linear system

$$\big(\mathbf{T} \otimes \mathbf{M} + \frac{k}{2}\mathbf{I} \otimes \widetilde{\mathbf{A}}\big)\underline{w} = \underline{g} \quad \text{with} \quad \underline{w} = (\mathbf{Q}^{\mathsf{H}} \otimes \mathbf{I})\underline{u}, \quad \underline{g} = (\mathbf{Q}^{\mathsf{H}} \otimes \mathbf{I})\underline{f}.$$

This system is block-upper-triangular. With $\underline{w} = (\underline{w}_0, \underline{w}_1, \dots, \underline{w}_r)$, $\underline{w}_j \in \mathbb{C}^{N^L}$, we obtain its solution by solving

$$\big(\lambda_{j+1}\mathbf{M} + \frac{k}{2}\widetilde{\mathbf{A}}\big)\underline{w}_j = \underline{s}_j \tag{5.4}$$

for $j = r, r-1, \ldots, 0$, where $\underline{s}_j = \underline{g}_j - \sum_{l=j+1}^{r} \mathbf{T}_{j+1,l+1}\mathbf{M}\underline{w}_l$.

5.3. **Iterative Solution of the Linear Equations.** By (5.4), a dG-time step of order r amounts to solving $r+1$ linear systems with a matrix of the form

$$\mathbf{B} = \lambda\mathbf{M} + \frac{k}{2}\widetilde{\mathbf{A}}. \tag{5.5}$$

Here, λ is an eigenvalue of $\mathbf{C}$ in (5.2). The linear equations (5.4) are solved approximately with incomplete GMRES iteration, causing an additional error in the overall scheme which we are going to analyze next.

5.3.1. *Eigenvalues of* $\mathbf{C}$. The eigenvalues λ_j of $\mathbf{C}$ have a major impact on the choice of the preconditioner (see Section 5.3.2) and thus on the convergence of the GMRES method for (5.4). It can be shown that (see [33, Lemma 5.5]) for $r = 1, 2, 3, \ldots$, the eigenvalues $\lambda_j^{(r)}$ of the matrix $\mathbf{C}$ from (5.2) satisfy

$$\mathrm{Re}\lambda_j^{(r)} \geq C_1|\lambda_j^{(r)}|r^{-\alpha}, \quad |\lambda_j^{(r)}| \geq C_2 r^{\widetilde{\alpha}}, \qquad j = 1, 2, \ldots, r+1, \tag{5.6}$$

with $\alpha = 2$, $\widetilde{\alpha} = 0$ and constants $C_1, C_2 > 0$ independent of r. Furthermore, the matrix $\mathbf{T}$ of the Schur decomposition $\mathbf{C} = \mathbf{QTQ}^{\mathsf{H}}$ satisfies $\|\mathbf{T}\|_2 \leq Cr^2$.

Remark 5.1. The estimates (5.6) do not seem to be sharp. Numerical results in [33] suggest that (5.6) holds with $\alpha = 2/3$ and $\widetilde{\alpha} = 1$.

5.3.2. *Preconditioning.* For the preconditioning of the linear system (5.5), we define the matrix $\mathbf{S}$ and the scaled matrix $\widehat{\mathbf{B}} \in \mathbb{R}^{N^L} \times \mathbb{R}^{N^L}$ by

$$\mathbf{S} := \left(\mathrm{Re}(\lambda)\mathbf{I} + \frac{k}{2}\mathbf{D_A}\right)^{1/2}, \qquad \widehat{\mathbf{B}} := \mathbf{S}^{-1}\mathbf{B}\mathbf{S}^{-1},$$

where $\mathbf{D_A}$ is the diagonal matrix with entries $D_{(l,j),(l,j)} = 2^{l\rho/2}$ corresponding to the basis functions $\psi_j^l \in V_h$; cf. (4.4). Moreover, we introduce the norm

$$\|\underline{w}\|_{\mathbf{D_A}^{-1}} = (\underline{w}^{\mathsf{H}}\mathbf{D_A}^{-1}\underline{w})^{1/2}, \qquad \underline{w} \in \mathbb{R}^{N^L}.$$

Let GMRES(m_0) denote the GMRES method with restart every $m_0 \geq 1$ iterations. For the linear system $\widehat{\mathbf{B}}\widehat{\underline{x}} = \widehat{\underline{b}}$, let $\widehat{\underline{x}}_j$ be the iterates obtained by restarted GMRES(m_0) with initial guess $\widehat{\underline{x}}_0$. Then, there exists $C > 0$ such that (cf. [33, Lemma 5.7])

$$\|\widehat{\underline{b}} - \widehat{\mathbf{B}}\widehat{\underline{x}}_j\| \leq Cq^j\|\widehat{\underline{b}} - \widehat{\mathbf{B}}\widehat{\underline{x}}_0\|, \quad j = 0, 1, 2, \ldots,$$

where $q = 1 - cr^{-2\alpha}$. Moreover, let $\underline{x}_j = \mathbf{S}^{-1}\widehat{\underline{x}}_j$, $\underline{b} = \mathbf{S}\widehat{\underline{b}}$. Then

$$\|\underline{b} - \mathbf{B}\underline{x}_j\|_{\mathbf{D_A}^{-1}} \leq Ch^{-\rho/2}(1 + kr^{\alpha-\widetilde{\alpha}})^{1/2}(1 - cr^{-2\alpha})^j\|\underline{b} - \mathbf{B}\underline{x}_0\|_{\mathbf{D_A}^{-1}}, \tag{5.7}$$

with $\alpha, \widetilde{\alpha}$ from (5.6), and $C, c > 0$ independent of L, k, r.

5.3.3. *Convergence of Incomplete GMRES.* The fully discrete dG scheme with incomplete GMRES yields approximations $\widehat{U}_{h,1}^{dG}, \widehat{U}_{h,2}^{dG}, \ldots, \widehat{U}_{h,M}^{dG}$ to $\widetilde{U}_{h,1}^{dG}, \widetilde{U}_{h,2}^{dG}, \ldots, \widetilde{U}_{h,M}^{dG}$ (cf. (5.1)). By (5.7) and proceeding as in [33, Sections 5.5.3 and 5.5.4], we obtain that

$$\|\widehat{U}_{h,M}^{dG}(T) - \widetilde{U}_{h,M}^{dG}(T)\| \leq Ch^{-3\rho/2}k^{-1}r^3q^{n_G}. \tag{5.8}$$

Here, n_G denotes the number of GMRES(m_0) iterations which has to be chosen in such a way that the bound in (5.8) is smaller than the right-hand side of (4.23) in the convergence Theorem 4.11, i.e.

$$h^{-3\rho/2}k^{-1}r^3q^{n_G} \leq Ch^{\beta}.$$

In order to fulfill this condition, we note that $q = 1 - cr^{-2\alpha}$, and use, in accordance with the assumptions in Theorem 4.11, that $r = \mathcal{O}(|\log h|)$ and $|\log k| \le Cr$. This results in

$$n_G \ge Cr^{2\alpha}\mathcal{O}(|\log h|) = \mathcal{O}(|\log h|)^{1+2\alpha} \tag{5.9}$$

GMRES iterations per time step. Combining the previous results, yields:

Theorem 5.2. *Let the assumptions of Theorem 4.11 hold. Then, choosing the number and order of time steps such that* $M = r = \mathcal{O}(|\log h|)$ *and in each time step* $n_G = \mathcal{O}(|\log h|)^{1+2\alpha}$ *GMRES iterations, implies*

$$\|u(T) - \widehat{U}_h^{dG}(T)\| \le Ch^{\beta}, \tag{5.10}$$

where $\widehat{U}_h^{dG}$ *denotes the dG approximation of the exact solution* u *obtained with the incomplete GMRES method.*

5.4. **Complexity.** Applying the matrix compression techniques from Section 4.1.2, the judicious combination of geometric mesh refinement and linear increase of polynomial degrees in the dG time-stepping scheme (Theorem 4.11), and an appropriate number of GMRES iterations (Theorem 5.2), results in an linear (up to some logarithmic terms) overall complexity of the fully discrete scheme (4.16) for the solution of the parabolic problem (2.1), (2.2).

Theorem 5.3. *Under the assumptions of Theorems 4.11, 5.2, the fully discrete scheme* (4.16) *with* n_G *GMRES iterations per time step satisfying* (5.9) *yields* $\widehat{U}_h^{dG}(T)$ *in at most* $\mathcal{O}(N(\log N)^{4+2\alpha})$ *operations, where* $N = N^L = \dim V_h = \mathcal{O}(h^{-1})$ *is the number of spatial degrees of freedom and* α *is as in* (5.6)*.*

Proof. There are $\mathcal{O}(r) = \mathcal{O}(|\log h|) = \mathcal{O}(\log N)$ time steps. In each time step $\mathcal{O}(r)$ linear systems of type (5.5) have to be solved by the incomplete GMRES method with

$$n_G = \mathcal{O}(|\log h|^{1+2\alpha}) = \mathcal{O}((\log N)^{1+2\alpha})$$

iterations. We note that the number of entries in the mass matrix $\mathbf{M}$ (with respect to the wavelet basis) is of order $\mathcal{O}(N \log N)$. Hence, recalling Proposition 4.4, each GMRES iteration involves a matrix-vector multiplication of complexity $\mathcal{O}(N \log N)$. This implies the desired result. □

6. Numerical Examples

In this section, we study the performance of the dG time-stepping scheme. Our wavelet spatial discretization is based on the biorthogonal spline wavelets (see Remark 4.1) with wavelet compression according to (4.6). The bandwidth of the lower-right block in the compressed wavelet matrix is 5 and the growth factor is taken to be 0.8, which corresponds to $c = 1$ and $\hat{\alpha} = 0.8$ in (4.6).

Our experiments address the numerical solution of (3.31), (3.32) with $\Omega = \Omega_R = (-6, 6)$, $g(x) = (e^x - K)_+$, $K = 1$ and $T = 1$. The integral kernel k of the non-local operator $\mathcal{L}^{\mathrm{jump}}$ in (3.9) is of CGMY-type (3.15) with $\alpha \in [0, 2)$. We recall that, for $\sigma > 0$, the operator $\mathcal{A} : V \to V^*$ with $V = H_0^1(\Omega)$ is given by $\mathcal{A} = (\mathcal{L}^{\mathrm{BS}} + \mathcal{L}^{\mathrm{jump}})|_{\Omega}$ as in (3.8), (3.9). In the pure jump case, i.e. $\sigma = 0$, there holds $\mathcal{A} = \mathcal{L}^{\mathrm{jump}}|_{\Omega}$.

Our aim is to illustrate the effect of the choice of the mesh size $h = 2R2^{-L-1}$, the number M of time steps and the polynomial degree r in the dG method on the error of the numerical approximation at $t = T$.

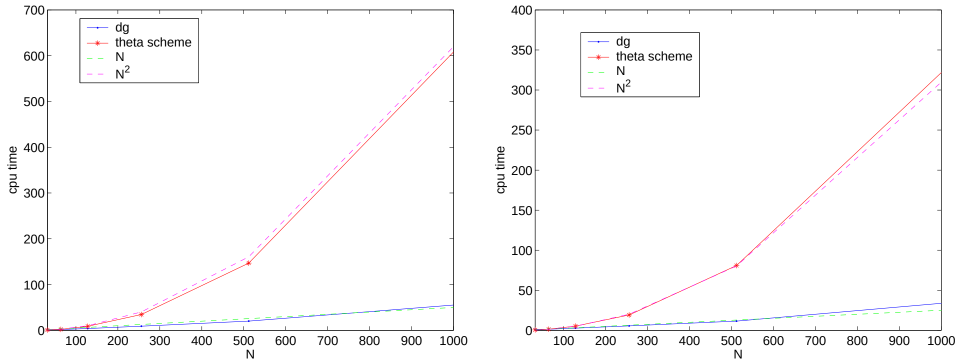

FIGURE 1. Performance comparison for the dG time-stepping vs. θ-scheme (both with incomplete GMRES) to advance from $t = 0$ to $t = T$; left: $\sigma > 0$, right: $\sigma = 0$.

To do so, let $\widehat{U}_h^{dG}(T)$ be the numerical price obtained by the dG-time stepping scheme using the incomplete GMRES method with wavelet compression and wavelet preconditioning. Furthermore, denote by $\widehat{U}_h^{\theta}(T)$ the corresponding solution of the well-known θ-scheme (also with incomplete GMRES and wavelet preconditioning). In Figure 1 we compare the cpu times for the computation of $\widehat{U}_h^{dG}(T)$ and $\widehat{U}_h^{\theta}(T)$ (with $\theta = 1$, i.e., implicit Euler scheme) for $\sigma > 0$ (left) and $\sigma = 0$ (right). The CGMY parameters are $C = 1$, $\eta_- = 1.8$, $\eta_+ = 2.5$ and $\alpha = 1.2$.

In addition, we study the convergence of the numerical prices obtained by the dG time-stepping and by the θ time-stepping scheme, respectively, expressed in the error estimate (5.10) as $h \to 0$. We set $\sigma = 0$ and $\alpha = 1.8$. The other CGMY parameters are $C = 1$, $\eta_- = 1.8$ and $\eta_+ = 2.5$. We have chosen a large fixed value for the number n_G of GMRES iterations, so that all the errors become insensitive with respect to the iteration error. Again, by $U_h^{dG}(T,x)$ and $U_h^{\theta}(T,x)$, we denote the solutions of the fully discrete dG time-stepping and the θ time-stepping scheme, respectively. The number M of time steps is chosen proportionally to L; more precisely, $M = 2|\log(h)|$. Our 'exact' solution $U_{\rm ex}(T,x)$ is computed as the numerical solution of the dG method with $L = 11$, $M = 2|\log(h)|$ and $r = M$ for all time steps. In Figure 2, we plot the relative error with respect to the L^2 norms

$$\frac{\|U_{\rm ex}(T,x)\|^2_{L^2(\Omega_R)} - \|U_h^{dG}(T,x)\|^2_{L^2(\Omega_R)}}{\|U_{\rm ex}(T,x)\|^2_{L^2(\Omega_R)}} \tag{6.1}$$

for the dG time-stepping scheme, and compare it to the relative L^2 error of the standard (low order) θ-scheme

$$\frac{\|U_{\rm ex}(T,x)\|^2_{L^2(\Omega_R)} - \|U_h^{\theta}(T,x)\|^2_{L^2(\Omega_R)}}{\|U_{\rm ex}(T,x)\|^2_{L^2(\Omega_R)}}.$$

As above, we take $\theta = 1$. We clearly observe linear convergence for the implicit Euler time stepping and an experimental convergence rate of order 1.69 for the dG time-stepping scheme.

In Figures 3 and 4, we plot the relative L^2 errors and the convergence rates obtained by least square fitting for the dG time stepping and varying values of α and σ. Again, the number M of time steps is proportional to L, $M = 2|\log(h)|$. The geometric time

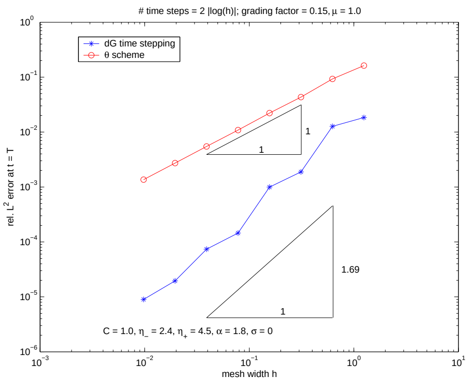

FIGURE 2. Comparison of the L^2 errors at $t = T$.

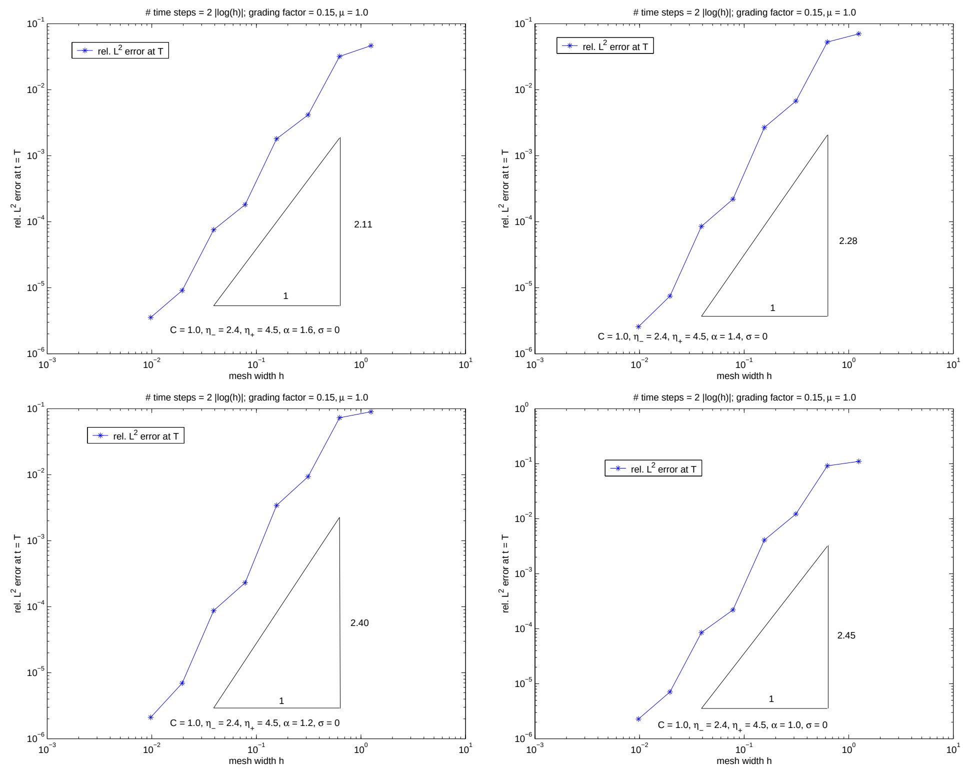

FIGURE 3. Relative L^2 errors (6.1) for pure jump CGMY Lévy processes ($\sigma = 0$), $\alpha \in \{1.6, 1.4, 1.2, 1.0\}$.

mesh has a grading factor $\gamma = 0.15$ and linearly increasing polynomial degrees with slope $\mu = 1.0$. The 'exact' solution is obtained as before.

Table 1 presents the values of the numerical approximation $U_h^{dG}(T)$ in the pure jump case ($\sigma = 0$) with CGMY parameters $C = 1.0$, $\eta_- = 2.4$, $\eta_+ = 4.5$ and $\alpha = 1.8$. The grading factor of the geometric time mesh has been chosen to be $\gamma = 0.2$ and the slope

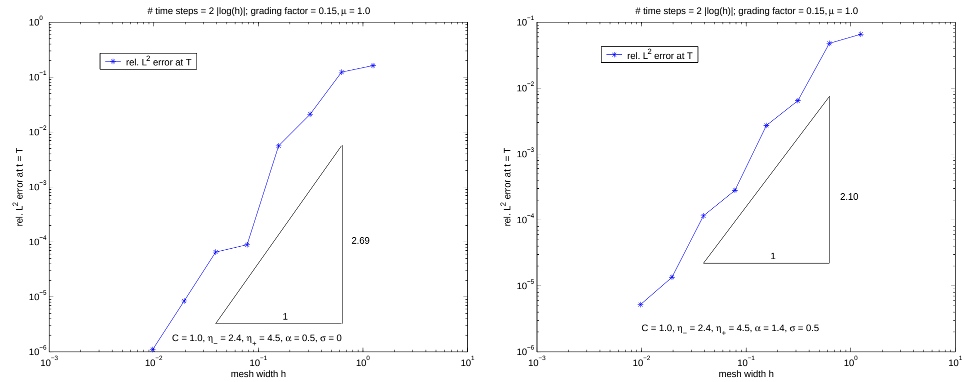

FIGURE 4. Relative L^2 errors (6.1) for $\sigma = 0$ and $\alpha = 0.5 < 1.0$ (left) and $\sigma > 0$ and $\alpha = 1.4$ (right).

of the linear polynomial degree vector is given by $\mu = 1.0$. As we increase the number of time steps (for fixed L), i.e., $M \to \infty$, we observe rapidly converging error levels. This indicates that the spatial discretization errors are dominant and that the time stepping scheme renders exponential convergence.

7. CONCLUDING REMARKS

In this work, we have studied the (fast) numerical approximation of (possibly degenerate) parabolic equations for general Markov processes arising in finance. Our scheme is based on hp-discontinuous Galerkin discretizations in time and wavelet discretizations in space. The densely populated matrices corresponding to the nonlocal infinitesimal generators are compressed using an appropriate wavelet basis. The linear systems are solved by an incomplete GMRES method. Under these conditions, we have shown that the overall complexity of the algorithm is linear up to some logarithmic terms. Moreover, we have proved that if the wavelet spatial semi-discretization error is of order h^β, then our algorithm gives $\widehat{U}_h^{dG}(T)$ with same error bound in log-linear complexity. Finally, numerical experiments illustrate the performance of the proposed scheme.

APPENDIX A. PROOF OF THEOREM 4.11

In this section, for simplicity, we assume that $T = 1$. A generalization of the following results to arbitrary time intervals $J = (0, T)$, $T > 0$, may then be obtained straightforwardly using a scaling argument.

For the proof of Theorem 4.11, we proceed in several steps.

A.1. **Time Regularity.**

We first discuss the time regularity of the semi-discrete (perturbed) parabolic problem (4.13), (4.14). Referring to [30], we obtain the following result:

Lemma A.1. *There exist two constants $C, d > 0$ (depending on the initial data u_0) such that for all $l \in \mathbb{N}_0$ and $0 < t \leq 1$, there holds*

$$\|\widetilde{u}_h^{(l)}(t)\|_V^2 \leq C d^{2l} \Gamma(2l + 2) t^{-2l}.$$

Here, $\Gamma(.)$ is the well-known Gamma-function.

Relative L^2 error for $\alpha = 1.8$, $\sigma = 0.0$							
$L \backslash M$	3	4	5	6	7	8	9
2	0.0197202	0.0196413	0.0196417	0.0196431	0.0196427	0.0196428	0.0196428
3	0.0143566	0.0142750	0.0142687	0.0142715	0.0142710	0.0142711	0.0142711
4	0.0019381	0.0018671	0.0018557	0.0018595	0.0018589	0.0018590	0.0018590
5	0.0010473	0.0009786	0.0009663	0.0009702	0.0009696	0.0009697	0.0009697
6	0.0002205	0.0001526	0.0001401	0.0001440	0.0001434	0.0001435	0.0001435
7	0.0001507	0.0000828	0.0000703	0.0000742	0.0000736	0.0000737	0.0000737
8	0.0000964	0.0000287	0.0000161	0.0000201	0.0000195	0.0000195	0.0000195
9	0.0000858	0.0000181	0.0000055	0.0000095	0.0000089	0.0000089	0.0000089
10	0.0000793	0.0000116	0.0000009	0.0000030	0.0000024	0.0000024	0.0000024

TABLE 1. Convergence of $\|U_h^{dG}(T)\|_{L^2(\Omega_R)}$ as $h \to 0$; $\alpha = 1.8$.

A.2. ***hp*-Approximations.** As before, we denote by $\widetilde{U}_h^{dG}$ the solution of the fully discrete (perturbed) dG method (4.16) in $\mathcal{S}^{\underline{r}}(\mathcal{M}, V_h)$. Moreover, let $\widetilde{u}_h$ be the solution of the parabolic problem (4.13), (4.14).

We consider an interpolant

$$\Pi^{\underline{r}}: \; H^1(J, V_h) \to \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h),$$

with $\Pi^{\underline{r}}|_{I_m} = \Pi^{r_m}$ satisfying the conditions

$$\begin{aligned} &\int_{I_m} (u_h - \Pi^{r_m} u_h, q)\, dt = 0, \qquad \forall q \in \mathcal{P}^{r_m - 1}(I_m, V_h), \\ &\Pi^{r_m} u_h(t_m^-) = u_h(t_m^-), \end{aligned} \tag{A.1}$$

for all $1 \leq m \leq M$, $u_h \in H^1(J, V_h)$. If, for some m, $r_m = 0$ then the first condition is void on I_m. According to [30], the interpolant $\Pi^{\underline{r}}$ is uniquely defined by the above conditions.

We have the following approximation result:

Lemma A.2. *Let $\mathcal{M}_{M,\sigma}$ with $0 < \sigma < 1$ be the geometric partition from Definition 4.9 and $\underline{r}$ the linear polynomial degree vector with slope $\mu > 0$ from Definition 4.10. Then, there exist constants $C, b > 0$ independent of M such that there holds the estimate*

$$\|\widetilde{u}_h - \Pi^{\underline{r}} \widetilde{u}_h\|^2_{L^2(J,V)} \leq C e^{-b\sqrt{N}},$$

where $N = \dim \mathcal{S}^{\underline{r}}(\mathcal{M}_{M,\sigma}, V_h)$.

Proof. Applying Lemma A.1 and referring to the definition of $\Pi^{\underline{r}}$, i.e. (since $r_1 = 0$) $\Pi^{\underline{r}} \widetilde{u}_h|_{I_1} = \widetilde{u}_h(t_1^-)$, yields

$$\begin{aligned} \int_{I_1} \|\widetilde{u}_h - \Pi^{r_1} \widetilde{u}_h\|_V^2 \, dt &= \int_0^{t_1} \|\widetilde{u}_h(t) - \widetilde{u}_h(t_1)\|_V^2 \, dt \\ &\leq 2 \int_0^{t_1} \|\widetilde{u}_h(t)\|_V^2 \, dt + 2t_1 \|\widetilde{u}_h(t_1)\|_V^2 \leq C t_1 \leq C \sigma^{(M-1)}. \end{aligned} \tag{A.2}$$

In addition, proceeding as in [30], we obtain for $2 \leq m \leq M$,

$$\|\widetilde{u}_h - \Pi^{r_m} \widetilde{u}_h\|^2_{L^2(I_m, V_h)} \leq C \sigma^{M-1} q^m,$$

for a real number $q \in (0,1)$ (independent of m). Hence, summing up the above bound over all $2 \leq m \leq M$ and using (A.2) implies

$$\|\eta\|^2_{L^2(J,V_h)} = \sum_{m=1}^{M} \|\widetilde{u}_h - \Pi^{r_m} \widetilde{u}_h\|^2_{L^2(I_m,V_h)} \leq C \sigma^{M-1} \Big(1 + \sum_{m=2}^{M} q^m \Big).$$

Noticing that $\sum_{m=2}^{M} q^m < \sum_{m=2}^{\infty} q^m < \infty$ and $N = \dim \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h) \leq C \mu M^2$ (as $M \to \infty$) completes the proof. □

A.3. **Convergence of the Fully Discrete dG Scheme.** The error of the dG method (4.16) with respect to the semi-discretization (4.13), (4.14) may be split as follows:

$$\widetilde{u}_h - \widetilde{U}_h^{dG} = \eta + \xi, \tag{A.3}$$

where

$$\eta = \widetilde{u}_h - \Pi^{\underline{r}} \widetilde{u}_h, \qquad \xi = \Pi^{\underline{r}} \widetilde{u}_h - \widetilde{U}_h^{dG}. \tag{A.4}$$

In the following Lemma A.3 we will prove that the dG error (A.3) at $t = T$ may be estimated in terms of η only.

Lemma A.3. *For the dG method* (4.16), *there holds the a priori estimate*

$$\|\widetilde{u}_h(T) - \widetilde{U}_h^{dG}(T)\|^2 \leq C \int_0^T \|\eta\|_V^2 \, dt.$$

Here, η is the interpolation error from (A.4).

Proof. The main idea is to bound ξ in terms of η. To do so, we first note that, by the Galerkin orthogonality (4.19), we have

$$\widetilde{B}_{dG}(\xi, \xi) = -\widetilde{B}_{dG}(\eta, \xi). \tag{A.5}$$

Moreover, as in (4.21), there holds that

$$\widetilde{B}_{dG}(\xi, \xi) = \|\xi\|_{dG}^2 \geq \int_0^T \widetilde{a}(\xi, \xi) \, dt. \tag{A.6}$$

Furthermore, making use of the properties (A.1) of the interpolant Π^r_-, we obtain for all $1 \leq m \leq M$ that $\int_{I_m} (\eta, \xi') \, dt = 0$, and $\eta(t_m^-) = 0$. Hence, integrating by parts leads to

$$\begin{aligned} \widetilde{B}_{dG}(\eta, \xi) &= \sum_{m=1}^{M} \int_{I_m} \big((\eta', \xi) + \widetilde{a}(\eta, \xi)\big) \, dt + \sum_{m=1}^{M} ([\eta]_m, \xi_m^+) + (\eta_0^+, \xi_0^+) \\ &= \int_0^T \widetilde{a}(\eta, \xi) \, dt \leq C \int_0^T \|\eta\|_V \|\xi\|_V \, dt. \end{aligned}$$

In the last step, the continuity of $\widetilde{a}$ was applied. Additionally, using a weighted Cauchy-Schwarz inequality and the coercivity of the bilinear form $\widetilde{a}$, results in

$$|\widetilde{B}_{dG}(\eta, \xi)| \leq C \int_0^T \|\eta\|_V^2 \, dt + \frac{1}{2} \int_0^T \widetilde{a}(\xi, \xi) \, dt.$$

Thus, recalling (A.5) and inserting (A.6), we obtain

$$\widetilde{B}_{dG}(\xi, \xi) \leq |\widetilde{B}_{dG}(\xi, \xi)| \leq |\widetilde{B}_{dG}(\eta, \xi)| \leq C \int_0^T \|\eta\|_V^2 \, dt + \frac{1}{2} \widetilde{B}_{dG}(\xi, \xi).$$

Finally, subtracting $\frac{1}{2}\widetilde{B}_{dG}(\xi, \xi)$ on both sides of the above inequality and noticing that $\eta_M^- = 0$, yields

$$\|\widetilde{u}_h(T) - \widetilde{U}_h^{dG}(T)\|^2 = \|\xi_M^-\|^2 \overset{(4.21)}{\leq} 2\widetilde{B}_{dG}(\xi, \xi) \leq C \int_0^T \|\eta\|_V^2 \, dt.$$

This is the desired estimate. $\square$

Combining the previous Lemmas A.2 and A.3 and the assumption (4.15) on the convergence of the spatial semi-discrete problem (4.13), (4.14), we obtain the following

Proposition A.4. *There exist two constants $C, \widetilde{b} > 0$ such that the error $u - \widetilde{U}_h^{dG}$ of the fully discrete dG scheme* (4.16) *satisfies the a priori estimate*

$$\|u(T) - \widetilde{U}_h^{dG}(T)\| \leq C\big(h^\beta + e^{-\widetilde{b}\sqrt{N}}\big),$$

where u and $\widetilde{U}_h^{dG}$ are the solutions of (2.1), (2.2) *and* (4.16), *respectively.*

Proof. There holds

$$\|u(T) - \widetilde{U}_h^{dG}(T)\|^2 \leq 2\|u(T) - \widetilde{u}_h(T)\|^2 + 2\|\widetilde{u}_h(T) - \widetilde{U}_h^{dG}(T)\|^2,$$

where $\widetilde{u}_h$ is the solution of the semi-discrete problem (4.13), (4.14). Then, referring to the Lemmas A.2 and A.3 and (4.15), implies

$$\|u(T) - \widetilde{U}_h^{dG}(T)\|^2 \leq C\big(h^\beta + e^{-b\sqrt{N}}\big)\|u_0\|_V^2,$$

which completes the proof. □

Now, the proof of Theorem 4.11 follows easily from the fact that $\mathcal{O}(\sqrt{N}) = \mathcal{O}(M) = \mathcal{O}(|\log h|)$.

REFERENCES

[1] R. A. Adams. *Sobolev Spaces*. Academic Press, 1978. New York.

[2] O. Barndorff-Nielsen. Processes of normal inverse gaussian type. *Finance and Stochastics*, 2:41–68, 1998.

[3] J. Bertoin. *Lévy processes*. Cambridge Univ. Press, 1996.

[4] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[5] S.I. Boyarchenko and S.Z. Levendorskii. *Non-Gaussian Merton-Black-Scholes Theory*. Advanced Series on Statistical Science & Applied Probability **9**, World Scientific, 2002.

[6] P. Carr, H. Geman, and D. B. Madan. The fine structure of asset returns: An empirical investigation. *Journal of Business*, 2002.

[7] A. Cohen. Wavelet methods for operator equations. In *Handbook of Numerical Analysis*. Elsevier, 2001. To appear.

[8] R. Cont, J. P. Bouchaud, and M. Potters. Scaling in financial data: stable laws and beyond. In B. Dubrulle, F. Graner, and D. Sornette, editors, *Scale invariance and beyond*. Springer Berlin, 1997.

[9] R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman and Hall/CRC Press, 2003.

[10] R. Cont and E. Voltchkova. A finite-difference scheme for option pricing in jump-diffusion and exponential lévy models. *Rapport Interne CMAP*, 513, 2003.

[11] W. Dahmen, A. Kunoth, and K. Urban. Biorthogonal spline wavelets on the interval—stability and moment conditions. *Appl. Comput. Harmon. Anal.*, 6(2):132–196, 1999.

[12] E. Eberlein. Application of generalized hyperbolic Lévy motions to finance. In O. E. Barndorff-Nielsen, T. Mikosch, and S. Resnick, editors, *Lévy Processes: Theory and Applications*, pages 319–337. Birkhäuser, 2001.

[13] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1:281–299, 1995.

[14] E. Eberlein, U. Keller, and K. Prause. New insights into smile, mispricing and value at risk: the hyperbolic model. *Journal of Business*, 71(3):371–405, 1998.

[15] M. Freidlin. *Functional integration and partial differential equations*. Annals of Mathematics Studies. Princeton University Press, 1985.

[16] N. Jacob. *Pseudo-Differential Operators and Markov Processes. Vol. 1: Fourier Analysis and Semigroups*. Imperial College Press, London, 2001.

[17] N. Jacob. *Pseudo-Differential Operators and Markov Processes. Vol. 2: Generators and their Potential Theory*. Imperial College Press, London, 2002.

[18] I. Karatzas and S. E. Shreve. *Methods of Mathematical Finance*. Springer, 1999.

[19] I. Koponen. Analytic approach to the problem of convergence of truncated lévy flights towards the gaussian stochastic process. *Phys. Rev. E*, 52:1197–1199, 1995.

[20] G. Kou. A jump diffusion model for option pricing. *Mangement Science*, 48:1086–1101, 2002.

[21] D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance*. Chapman & Hall, 1997.

[22] J. L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications*. Springer Berlin, 1972.

[23] D. B. Madan, P. Carr, and E. Chang. The variance gamma process and option pricing. *European Finance Review*, 2(79–105), 1998.

[24] A.-M. Matache, T. von Petersdorff, and Ch. Schwab. Fast deterministic pricing of options on Lévy driven assets. Technical Report 2002-11, in press in R.A.I.R.O. M^2AN, available also as Research Report Seminar for Applied Mathematics, ETH Zürich, 2002.

[25] R.C. Merton. Option pricing when the underlying stocks are discontinuous. *J. Financ. Econ.*, 5:125–144, 1976.

[26] D. Nualart and W. Schoutens. Backward stochastic differential equations and Feynman-Kac formula for Lévy processes, with applications in finance. *Bernoulli*, 7(5):761–776, 2001.

[27] H.C. Öttinger. *Stochastic Processes in Polymeric Fluids*. Springer, 1998.

[28] K.-I. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.

[29] R. Schneider. *Multiskalen- und Wavelet-Matrixkompression. Analysisbasierte Methoden zur effizienten Lösung groser vollbesetzter Gleichungssysteme*. Advances in Numerical Mathematics. B. G. Teubner, Stuttgart, 1998.

[30] D. Schötzau and C. Schwab. Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method. *SIAM J. Num. Anal.*, 38:837–875, 2000.

[31] D. Schötzau and C. Schwab. hp-discontinuous Galerkin time-stepping for parabolic problems. *C.R.Acad. Sci. Paris*, Série I:1121–1126, 2001.

[32] T. von Petersdorff and C. Schwab. Fully discrete multiscale galerkin bem. In *Multiresolution Analysis and Partial Differential Equations, W. Dahmen, P. Kurdila and P. Oswald (Eds.)*, volume 6 of *Wavelet Analysis and Its Applications*, pages 287–346. Academic Press, New York, 1997.

[33] T. von Petersdorff and C. Schwab. Numerical solution of parabolic equations in high dimensions. *To appear in R.A.I.R.O. M^2AN, available also as Technical Report NI03013-CPD Isaac Newton Institute for Mathematical Sciences, Cambridge.* `http://www.newton.cam.ac.uk/preprints/NI03013.pdf`, 2002.

[34] T. von Petersdorff and C. Schwab. Wavelet discretizations of parabolic integro-differential equations. *SIAM J. Numer. Anal.*, 41(1):159–180, 2003.

SEMINAR FOR APPLIED MATHEMATICS AND RISKLAB, ETH ZÜRICH, 8092 ZÜRICH, SWITZERLAND.
E-mail address: `amatache@math.ethz.ch`

SEMINAR FOR APPLIED MATHEMATICS, ETH ZÜRICH, 8092 ZÜRICH, SWITZERLAND.
E-mail address: `schwab@math.ethz.ch`

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA, MINNEAPOLIS MN, 55455, USA.
E-mail address: `wihler@math.umn.edu`