Wavelet solution of variable order pseudodifferential equations [*]

R. Schneider[†] and C. Schwab

Research Report No. 2005-05
May 2005

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

[†]Scientific Computing, CAU Kiel, D-24098 Kiel, Germany

Wavelet solution of variable order pseudodifferential equations[*]

R. Schneider
Scientific Computing
CAU Kiel
D-24098 Kiel
Germany

and

C. Schwab
Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Abstract

Sobolev spaces $H^{m(x)}(I)$ of variable order $0 < m(x) < 1$ on an interval $I \subset \mathbb{R}$ arise as domains of Dirichlet forms for certain quadratic, pure jump Feller processes $X_t \in \mathbb{R}$ with unbounded, state-dependent intensity of small jumps. For spline wavelets with complementary boundary conditions, we establish multilevel norm equivalences in $H^{m(x)}(I)$ and prove preconditioning and wavelet matrix compression results for the variable order pseudodifferential generators A of X.

Sufficient conditions on A to satisfy a Gårding inequality in $H^{m(x)}(I)$ and time-analyticity of the semigroup T_t associated with the Feller process X_t are established.

As application, wavelet-based algorithms of log-linear complexity are obtained for the valuation of contingent claims on pure jump Feller-Lévy processes X_t with state-dependent jump intensity by numerical solution of the corresponding Kolmogoroff equations.

Keywords: Feller processes, Wavelets, Pseudodifferential Operators,
Analytic Semigroups, Option Pricing

Subject Classification: 35K15, 45K05, 65N30

1 Introduction

Arbitrage-free values $v(x, \bar{t})$ of contingent claims on assets whose log-returns are modelled by a strong Markov process X_t with state space $\mathbb{R}$ can be expressed as (conditional) expectations of prices at time of maturity $\bar{t} > 0$, i.e.

$$v(x, \bar{t}) = \mathbb{E}^x(g(X_{\bar{t}})) \tag{1.1}$$

(e.g. [10]). Here $\mathbb{E}^x$ denotes expectation with respect to a martingale measure of X_t, conditional to $X_0 = x$, and $g(x)$ is a given, so-called *payoff function* of the contingent claim. The classical example for X_t is geometric Brownian motion in the Black-Scholes-Merton model [4, 24].

Definition (1.1) suggests to estimate $v(x, \bar{t})$ by Monte-Carlo simulation, i.e. by averaging over M realizations $\{X_t(\omega_j)\}_{j=1}^M$ of sample-paths. The statistical error in the estimate $v^M(x, \bar{t})$ thus obtained tends to zero not faster than $O(1/\sqrt{M})$, in general.

Deterministic methods to compute $v(x, \bar{t})$ are based on the semigroup T_t of X_t defined by

$$v(x, t) = (T_t g)(x) = \mathbb{E}^x(g(X_t)), \qquad t > 0. \tag{1.2}$$

The process X_t and its Semigroup T_t are called *Feller* if

i) T_t maps $C_0(\mathbb{R})$, the continuous functions on $\mathbb{R}$ vanishing at infinity, into itself:

$$T_t : C_0(\mathbb{R}) \to C_0(\mathbb{R}) \quad \text{boundedly}$$

and

ii) if T_t is strongly continuous, i.e. $\lim_{t \to 0^+} \|u - T_t u\|_{L^\infty(\mathbb{R})} = 0$ for all $u \in C_0(\mathbb{R})$.

The *infinitesimal generator* A_X with domain $\mathcal{D}(A_X)$ of a Feller process X_t with semigroup T_t is defined by the strong limit

$$A_X u := \lim_{t \to 0^+} \frac{1}{t}(T_t u - u) \tag{1.3}$$

on all functions $u \in \mathcal{D}(A_X) \subset C_0(\mathbb{R})$ for which the limit (1.3) exists w.r. to the sup-norm. We call $(A_X, \mathcal{D}(A_X))$ *Feller generator* of X. Feller generators admit the *positive maximum principle*, i.e.

$$\text{if} \quad u \in \mathcal{D}(A_X) \quad \text{and} \quad \sup_{x \in \mathbb{R}} u(x) = u(x_0) > 0, \quad \text{then} \quad (A_X u)(x_0) \leq 0 \tag{1.4}$$

and admit a pseudodifferential representation (e.g. [8, 14, 15]):

Theorem 1.1. *Let $(A, \mathcal{D}(A))$ be a Feller generator with $C_0^\infty(\mathbb{R}) \subset \mathcal{D}(A)$. Then $A|_{C_0^\infty(\mathbb{R})}$ is a pseudodifferential operator,*

$$(Au)(x) = -a(x, D)u(x) = -(2\pi)^{-1/2} \int_{\mathbb{R}} a(x, \xi)\hat{u}(\xi)e^{ix\xi}d\xi \quad u \in C_0^\infty(\mathbb{R}) \tag{1.5}$$

with symbol $a(x, \xi) : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ which is measurable and locally bounded in (x, ξ) and which admits the Lévy-Khintchine *representation*

$$a(x, \xi) = c(x) - i\gamma(x)\xi + (\sigma(x))^2\xi^2 + \int_{0 \neq y \in \mathbb{R}} \left(1 - e^{iy\xi} + \frac{iy\xi}{1 + y^2}\right) N(x, dy) \tag{1.6}$$

1

The parameters $c(x), \gamma(x), \sigma(x), N(x, dy)$ in (1.6) are called *characteristics* of the Feller process X_t. Spatially and temporally homogeneous Feller processes X_t are Lévy-processes (e.g.[2, 27]). Their characteristics, the *Lévy characteristics*, do not depend on x explicitly.

Symbols $a(x, \xi)$ of the form (1.6) are called *negative definite* symbols. In (1.6), $N(x, dy)$ is the compensated jump measure of the Feller process $X_t^x = \{X_t | X_0 = x\}$ which satisfies

$$\sup_{x \in \mathbb{R}} \int_{\mathbb{R}} \min(1, y^2) N(x, dy) < \infty. \tag{1.7}$$

Fast *deterministic* computation of the conditional expectation $v(x, t)$ in (1.2) across all maturities $0 < t < \bar{t}$ is based on the numerical solution of the backward Kolmogoroff equation

$$v_t + Av = 0, \qquad v|_{t=T} = g. \tag{1.8}$$

In the Black-Scholes model [4], X_t is a diffusion for which $N(x, dy) = 0$ in (1.6). The generator A in (1.8) is a diffusion with possibly spatially inhomogeneous characteristics $c(x), \gamma(x), \sigma(x)$, resp. killing, drift and volatility, corresponding to local volatility models.

Exploiting the analyticity of the semigroup T_t and a spline wavelet discretization of degree $p \geq 1$ for generators A in (1.5) which are classical pseudodifferential operators of constant order $2m$, $0 \leq m \leq 1$, it has been shown in [21, 22, 23] that $v(x, t)$ can be computed for $0 < t \leq T$ with essentially $O(N)$ work and memory essentially [*] to accuracy $O(N^{-p-1})$. Key ingredient in the numerical analysis in [22, 23] were i) wavelet norm equivalences in the "energy" space of A (which are also crucial ingredients for adaptive solution methods), and ii) the wavelets' vanishing moment property to compress the N^2 entries in the stiffness matrices of A to $O(N \log N)$ "essential" elements without loss in accuracy.

Here, we address the efficient solution of (1.8) for pure jump processes X with state dependent jump intensity. In this case, $\sigma = 0$ in (1.6) and the domain V of the Dirichlet form associated to A_X is a Sobolev space of variable order $m(x) \in (0, 1)$. Development of linear-complexity solvers for the Kolmogoroff equation (1.8) along the lines of [22, 23] and also for the pseudodifferential inequality arising from expectations over the stopped processes X (e.g. [20]) requires a) verification of wavelet norm equivalences in the "energy" space $H^{m(x)}$ and b) wavelet compression of the generator A_X or, equivalently, of the jump measure of the Feller process X and c) establishing time-analyticity of the semigroup T_t. Proving a) - c) is the main purpose of the present paper.

Its outline is as follows: in Section 2, we introduce a class of pseudodifferential operators of variable order containing, in particular, certain generators A_X of the form (1.6). To analyze wavelet-based solution algorithms for (1.8), we derive, based on the calculus of [18, 12], estimates for the Schwartz kernels of A_X, resp. the densities of the compensated jump measure $N_X(x, dy)$ of X_t with respect to the Lebesgue measure dy which are interesting in their own right. In Section 3, we define the variable order Sobolev spaces $H^{m(x)}$ for $0 \leq m(x) < 1$ which are the domains of Dirichlet forms of A_X. They also form the basis of Galerkin discretization of A_X. In Section 4, we use the bounds on the Schwartz kernels to establish our main results: multilevel norm equivalences in the variable order spaces $H^{m(x)}$ for $0 \leq m(x) < 1$ and compression estimates for the moment matrices of A_X in the wavelet basis. In Section 5, we prove the Gårding inequality for the generator (1.6) in variable order Sobolev spaces and deduce time-analyticity of the Semigroup T_t. The final section 6 addresses the generalization of the fast pricing algorithms from [22, 23] to variable order generators.

---

[*]Throughout, "essentially" means up to powers of $\log N$ or $\log h$ which will be clear from the context

2 Generators of Variable Intensity Feller-Lévy Processes

The infinitesimal generator A_X of the Feller process X is the pseudodifferential operator $a(x, D)$ provided by (1.5) in Theorem 1.1. Denote by $K_X(x, y)$ the Schwartz kernel of the nonlocal part of A_X. The distribution $K_X(x, y)$ is the density of the jump measure $N_X(x, dy)$ of X with respect to the Lebesgue measure dy, i.e.

$$N_X(x, dy) = K_X(x, y)dy. \tag{2.1}$$

For pure jump Feller processes X with state-dependent jump-intensity, domains $\mathcal{D}(A_X)$ of their generators A_X are Sobolev spaces of variable order $m(x)$ which, by (1.7), satisfies $m(x) < 1$. Wavelet solution of the derivative pricing pseudodifferential equation (1.8) for variable intensity processes X with generators of type (1.6), (1.7) requires multilevel norm equivalences. To prove these, we use a calculus for variable order pseudodifferential operators. We start by identifying classes of variable intensity pure jump Feller processes X through conditions on the Lévy symbols $a(x, \xi)$ in (1.6) of their infinitesimal generators A_X.

Throughout we use the notation $\langle \xi \rangle := (1 + |\xi|^2)^{1/2}$. We define a class of quadratic, pure jump Feller processes X with variable jump intensity through their generators A_X which are variable order pseudodifferential operators. They, in turn, are given in terms of their symbols. The following symbol classes of variable order have been introduced in [16].

Definition 2.1. *Let $0 \leq \delta < \rho \leq 1$ and let $m(x) \in C^\infty(\mathbb{R})$ be a real-valued function all of which derivatives are bounded on $\mathbb{R}$. The symbol $a(x, \xi)$ belongs to the class $S_{\rho,\delta}^{m(x)}$ of symbols of variable order $m(x)$ if $a(x, \xi) \in C^\infty(\mathbb{R} \times \mathbb{R})$ and $m(x) = s + \tilde{m}(x)$ with $\tilde{m} \in \mathcal{S}(\mathbb{R})$ is a tempered function, and if, for every $\alpha, \beta \in \mathbb{N}_0$ there is a constant $c_{\alpha,\beta}$ such that*

$$\forall x, \xi \in \mathbb{R} : \quad |D_x^\beta D_\xi^\alpha a(x, \xi)| \leq c_{\alpha,\beta} \langle \xi \rangle^{m(x) - \rho|\alpha| + \delta|\beta|}. \tag{2.2}$$

The variable order pseudo-differential operators $A(x, D) \in \Psi_{\rho,\delta}^{m(x)}$ correspond to symbols $a(x, \xi) \in S_{\rho,\delta}^{m(x)}$ by

$$A(x, D)u(x) := \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{i(x-y)\cdot\xi} a(x, \xi) u(y) dy d\xi, \qquad u \in C_0^\infty(\mathbb{R}). \tag{2.3}$$

Given a variable order $m(x)$, we define $\overline{m} := \sup_{x \in \mathbb{R}} m(x)$ and $\underline{m} := \inf_{x \in \mathbb{R}} m(x)$ and assume $0 \leq \delta < \rho \leq 1$ throughout.

For $a(x, \xi) \in S_{\rho,\delta}^{m(x)}$, the *Martingale Problem* is the problem of existence of a Feller Process X with given generator $A(x, D)$. We have ([16, 12], [15], Ch. 2.10 and [28]).

Proposition 2.2. *For every negative definite symbol $a(x, \xi) \in S_{\rho,\delta}^{m(x)}$ of variable order $m(x)$ with $1 \geq \rho > \delta > 0$ exists a unique Feller process X_t with generator A_X as in (1.5), (1.6) and jump measure $N_X(x, dy)$ as in (2.1).*

By $FP_{\rho,\delta}^{m(x)}$ we denote the set of all Feller processes with generator $A_X \in \Psi_{\rho,\delta}^{m(x)}$ given by a negative definite characteristic function $a(x, \xi) \in S_{\rho,\delta}^{m(x)}$. Their domains $\mathcal{D}(A_X)$ are Sobolev spaces of variable order. Our definition and characterization of these variable order Sobolev spaces will be based on variable order Riesz potentials $\Lambda^{m(x)}$ with symbol $a(x, \xi) = \langle \xi \rangle^{m(x)}$. The operators $\Lambda^{m(x)}$ are the generators of so-called 'stable-like' Feller processes whose existence has been established in [1].

3

More general pseudodifferential generators and their associated Feller processes have been considered for example in [16, 12, 28]. Since $|D_x^\beta D_\xi^\alpha \langle \xi \rangle^{m(x)}| \lesssim \langle \xi \rangle^{m(x)-|\alpha|} |\log\langle \xi \rangle|^{|\beta|}$, we have $a_\Lambda \in S_{\rho,\delta}^{m(x)} \subset S_{\rho,\delta}^{\overline{m}}$ for all $0 < \delta < \rho \leq 1$. The following lemma from [16] is the variable order analogue to a classical result for constant order pseudo-differential operators (see, e.g., Theorems 3.1 to 3.4 in [31] or Chap. VI.6 in [32]).

Lemma 2.3. *Consider compound symbols of the form $a(x,\xi,y) \in C^\infty(\mathbb{R} \times \mathbb{R} \times \mathbb{R})$ satisfying*

$$|D_x^\alpha D_\xi^\beta D_y^\gamma a(x,\xi,y)| \leq c_{\alpha,\beta,\gamma} \langle \xi \rangle^{m(x)+\delta(|\alpha|+|\gamma|)-\rho|\beta|},$$

together with the corresponding operators

$$(A(x,D,y)u)(x) := \frac{1}{2\pi} \int_\mathbb{R} \int_\mathbb{R} e^{i(x-y)\xi} a(x,\xi,y) u(y) dy d\xi \ , \ \ u \in C_0^\infty(\mathbb{R}).$$

If $0 \leq m(x)$ for all $x \in \mathbb{R}$ and $\|m\|_{L^\infty(\mathbb{R})} < 1$, and $A(x,D,y)$ is compactly supported (see e.g. [17, 31]), there exists a pseudo-differential operator $B \in \Psi_{\rho,\delta}^{m(x)}$ with symbol $b \in S_{\rho,\delta}^{m(x)}$ together with a pseudo-differential operator $C \in \Psi_{\rho,\delta}^{m(x)-\rho+\delta}$, such that

$$A(x,D,y) = B(x,D) + C \tag{2.4}$$

holds.

Furthermore there exist pseudo-differential operators $T_1, T_2 \in \Psi_{1,\delta}^{-\epsilon}$, $\epsilon, \delta > 0$, and an $\epsilon > 0$, such that, for $a(x,\xi) := \langle \xi \rangle^{m(x)} \in S_{1,\delta}^{m(x)}$, $b(x,\xi) := \langle \xi \rangle^{-m(x)} \in S_{1,\delta}^{-m(x)}$ and corresponding operators $A(x,D) \in \Psi_{1,\delta}^{m(x)}$, $B(x,D) \in \Psi_{1,\delta}^{-m(x)}$ there holds

$$A(x,D)B(x,D) = I + T_1, \qquad B(x,D)A(x,D) = I + T_2. \tag{2.5}$$

A calculus for variable order pseudodifferential operators $\Psi_{\rho,\delta}^{m(x)}$ was developed in [16], see also [12].

Define the operators L by $(Lu)(x) := (\Lambda^{m(x)}u)(x) := \frac{1}{2\pi} \int \int e^{i\xi(x-y)} \langle \xi \rangle^{m(y)} u(y) dy d\xi$. Then the calculus [16] shows that $L^\top := \Lambda^{m(x)} \in \Psi_{1,\delta}^{m(x)}$ and the operators $A = L$, $A = L^\top \in \Psi_{1,\delta}^{m(x)}$, $A = L^{-1} \in \Psi_{1,\delta}^{-m(x)}$ and $A = LL^\top = (\Lambda^{m(x)})^\top (\Lambda^{m(x)}) \in \Psi_{1,\delta}^{2m(x)}$, as well as $A^{(-1)} = (LL^\top)^{(-1)} = (\Lambda^{-m(x)})^\top (\Lambda^{-m(x)}) \in \Psi_{1,\delta}^{-2m(x)}$ can be expressed as singular integrals with respect to their Schwartz kernels $K_A(x,y)$ (related to the jump measure $N_X(x,dy)$ of the Markov process X with generator A by (2.1))

$$Au(x) = \int_\mathbb{R} K_A(x,y) u(y) dy.$$

Formally, these kernels are given by oscillatory integrals

$$K_A(x,y) = \frac{1}{2\pi} \int_\mathbb{R} e^{i(x-y)\cdot\xi} a_A(x,\xi,y) d\xi,$$

with

$$a_L(x,\xi) = \langle \xi \rangle^{m(x)}, \quad a_{LL^\top}(x,\xi,y) = \langle \xi \rangle^{m(x)+m(y)}$$

and

$$a_{L^{-1}}(x,\xi,y) = \langle \xi \rangle^{-m(x)} + b(x,\xi,y)$$

with a corresponding pseudo-differential operator $B \in \Psi_{1,\delta}^{-\epsilon}$ for sufficiently small $\epsilon > 0$.

The following lemma plays a crucial role in the wavelet compression of $a_L(x,D)$.

4

Lemma 2.4. *For any $0 < \delta < 1$, the Schwartz kernel $K_A(x,y)$ of $A = LL^\top \in \Psi_{1,\delta}^{2m(x)}$ satisfies the Caldéron-Zygmund type estimate*

$$|D_x^\alpha D_y^\beta K_A(x,y)| \leq C_{\alpha,\beta,\delta}|x-y|^{-(1+m(x)+m(y)+(1-\delta)(|\alpha|+|\beta|))}, \tag{2.6}$$

where $x \neq y$, $|x-y| \lesssim 1$. For $A^{-1} = (LL^\top)^{(-1)} \in \Psi_{1,\delta}^{-2m(x)}$, a corresponding estimate holds

$$|D_x^\alpha D_y^\beta K_{A^{-1}}(x,y)| \leq C_{\alpha,\beta,\delta}|x-y|^{-(1-m(x)-m(y)+(1-\delta)(|\alpha|+|\beta|))}. \tag{2.7}$$

For $|x-y| \to \infty$ the kernels are rapidly decaying in the sense that

$$|D_x^\alpha D_y^\beta K(x,y)| \lesssim |x-y|^{-N}$$

for every $N > 0$ and for $K = K_A$ and $K = K_{A^{-1}}$.

The technique for validating of the above assertion is standard, e.g. Chap. VI of [32]. For the reader's convenience we provide a proof. Let $\widehat{\chi} \in C_0^\infty(\mathbb{R})$ be a cut-off function, $0 \leq \widehat{\chi}(x) \leq 1$, $\widehat{\chi}(x) = 1$ if $x \in [0,1]$, with support $supp\widehat{\chi} \subset [0,2]$, such that $\widehat{\Psi}(\xi) = \widehat{\chi}(\xi) - \widehat{\chi}(2\xi)$ defi nes a Littlewood Paley decomposition cf. [32] , i.e.

$$\widehat{\chi}(\xi) + \sum_{j=1}^\infty \widehat{\Psi}(2^{-j}\xi) = 1 \,, \forall \xi \in \mathbb{R}.$$

The extended symbol $a(x,\xi,y) = (1+\xi^2)^{(m(x)+m(y))/2} = \langle\xi\rangle^{m(x)+m(y)}$ of $A = LL^\top$ can be decomposed into the sum

$$a(x,\xi,y) = a_0(x,\xi,y) + \sum_{j=1}^\infty a_j(x,\xi,y) := a(x,\xi,y)\widehat{\chi}(\xi) + \sum_{j=1}^\infty a(x,\xi,y)\widehat{\Psi}(2^{-j}\xi). \tag{2.8}$$

From this we obtain the corresponding decomposition of the operator

$$A = A_0 + \sum_{j=1}^\infty A_j. \tag{2.9}$$

Here the operators A_j are defi ned by their Schwartz kernels

$$k_j(x,z,y) = \frac{1}{2\pi}\int_{\xi\in\mathbb{R}} e^{i\xi\cdot z}a_j(x,\xi,y)d\xi \,, \ z = x - y\,. \tag{2.10}$$

Lemma 2.5. *Let $(x,z,y) \mapsto k_j(x,z,y)$ be the kernel functions defined by (2.10). Then for all $M \geq 0$ and $z \neq 0$ there holds*

$$|D_x^\alpha D_y^\beta D_z^\gamma k_j(x,z,y)| \leq C_{M,\alpha,\beta}|z|^{-M}2^{j(-M-1-|\gamma|+\delta(|\alpha|+|\beta|)+m(x)+m(y))} \tag{2.11}$$

Proof. Due to defi nition (2.10) partial integration yields

$$(-z)^\mu D_z^\gamma D_x^\alpha D_y^\beta k_j(x,z,y) = \frac{1}{2\pi}\int e^{i\xi\cdot z}D_\xi^\mu[(i\xi)^\gamma D_x^\alpha D_y^\beta a_j(x,\xi,y)]d\xi \tag{2.12}$$

5

We observe that the integrand is supported on $\{\xi : 2^{j-1} \le |\xi| \le 2^{j+1}\}$ and the volume of this support is bounded by $c2^j$. Therefore, we can estimate the above integral (2.12) directly using

$$|D_x^\alpha D_y^\beta D_\xi^\mu a_j(x, \xi, y)| \lesssim \langle \xi \rangle^{m(x)+m(y)-|\mu|+\delta(|\alpha|+|\beta|)},$$

by

$$
\begin{aligned}
|z^\mu D_x^\alpha D_y^\beta D_z^\gamma k_j(x, z, y)| &\le \frac{1}{2\pi} \int |D_\xi^\mu[\xi^\gamma D_x^\alpha D_y^\beta a_j(x, \xi, y)]| d\xi \\
&\le C_{\mu,\alpha,\beta} 2^{j(1+m(x)+m(y)+|\gamma|-|\mu|+\delta(|\alpha|+|\beta|))}.
\end{aligned}
$$

Choosing $M = |\mu|$ gives the desired result. $\qquad \square$

We can now prove Lemma 2.4.

Proof. Since

$$K(x, y) = k_0(x, x-y, y) + \sum_{j=1}^\infty k_j(x, x-y, y)$$

it is sufficient to estimate

$$D_x^\alpha D_y^\beta k_0(x, x-y, y) + D_x^\alpha D_y^\beta \sum_{j=1}^\infty k_j(x, x-y, y). \tag{2.13}$$

From Lemma 2.5 we conclude the estimate

$$|D_x^\alpha D_y^\beta k_j(x, x-y, y)| \le C_{M,\alpha,\beta}|x-y|^{-M} 2^{j(M+1-(1-\delta)(|\alpha|+|\beta|)+m(x)+m(y))}.$$

First we consider the case $|z| = |x - y| \le 1$ and decompose the sum (2.13)

$$k_0 + \sum_{j=1}^\infty k_j = \sum_{2^j \le |z|^{-1}} k_j + \sum_{2^j \ge |z|^{-1}} k_j.$$

Setting $M = 0$ in Lemma 2.5 the first sum can be majorized by a multiple of

$$\sum_{2^j \le z^{-1}} 2^{j(1+m(x)+m(y)-(1-\delta)(|\alpha|+|\beta|))} \lesssim |z|^{-(1+m(x)+m(y)-(1-\delta)(|\alpha|+|\beta|))}$$

provided that $1 + m(x) + m(y) - (1 - \delta)(|\alpha| + |\beta|) > 0$.

To estimate the second sum $\sum_{2^j \ge |z|^{-1}} k_j$ we choose

$$M > 1 + m(x) + m(y) - (1 - \delta)(|\alpha| + |\beta|)$$

and obtain an upper bound

$$C|z|^M \sum_{2^j \ge |z|^{-1}} 2^{-j(1-M+m(x)+m(y)-(1-\delta)(|\alpha|+|\beta|))} \lesssim |z|^{-((1+m(x)+m(y)-(1-\delta)(|\alpha|+|\beta|))}.$$

To conclude the proof we consider $|z| \ge 1$. Then Lemma 2.5 provides, for all $N := M - m(x) + m(y) - (1 - \delta)(|\alpha| + |\beta|) > 0$, that (2.13) can be bounded by $C_{N,\alpha,\beta}|z|^{-N}$. $\qquad \square$

Using analogous arguments as above, we obtain

Theorem 2.6. *Let $A \in \Psi^{2m(x)}_{\rho,\delta}$ for some $0 < \delta < \rho \leq 1$. Then the Schwartz-kernel $K_A(x,y)$ of A satisfies for all $\alpha, \beta \in \mathbb{N}_0$ and all $x, y \in \mathbb{R}$ with $x \neq y$*

$$|D_x^\alpha D_y^\beta K_A(x,y)| \leq C_{\alpha,\beta,\delta}|x - y|^{-(1+m(x)+m(y)+(\rho-\delta)(|\alpha|+|\beta|))}. \tag{2.14}$$

In the applications we have in mind, the spatial domain is the unit interval $I = (0,1)$ rather than all of $\mathbb{R}$ and the variable order $m(x)$ is only defined on I. There holds $0 < m(x) < 1$, for all $x \in [0,1]$. For technical purposes we extend $m(x) = s + \tilde{m}(x)$ such that there exists $\epsilon > 0$ with $\epsilon < m(x) < 1 - \epsilon$ and $\tilde{m} \in \mathcal{S}(\mathbb{R})$.

3 Variable Order Sobolev Spaces

We define the Sobolev space $H^{m(x)}(\mathbb{R})$ of variable order $m(x) \in [0,1)$ on $\mathbb{R}$ through the $H^{m(x)}(\mathbb{R})$ Sobolev norm given by

$$\|u\|^2_{H^{m(x)}(\mathbb{R})} := \|\Lambda^{m(x)}u\|^2_{L_2(\mathbb{R})} + \|u\|^2_{L_2(\mathbb{R})}. \tag{3.1}$$

We also introduce the inner product

$$\langle u, v \rangle_{m(x)} := \langle \Lambda^{m(x)}u, \Lambda^{m(x)}v \rangle + \langle u, v \rangle,$$

so that $\langle u, u \rangle_{m(x)} = \|u\|^2_{H^{m(x)}(\mathbb{R})}$. The space $H^{-m(x)}(\mathbb{R})$ is defined as the dual space of $H^{m(x)}(\mathbb{R})$ with respect to the $\langle ., . \rangle$ - duality. An equivalent norm in $H^{-m(x)}(\mathbb{R})$ is given by

$$\langle u, v \rangle_{-m(x)} := \langle \Lambda^{m(x)}u, \Lambda^{m(x)}v \rangle + \langle \Lambda^{-2\overline{m}}u, v \rangle, \qquad \|u\|^2_{H^{-m(x)}(\mathbb{R})} = \langle u, u \rangle_{-m(x)}.$$

We remark that due to the symbolic calculus in [16] referred to in Lemma 2.3, an equivalent definition used by [16, 18] is

$$\|u\|_{H^{m(x)}(\mathbb{R})} := \|(\Lambda^{m(x)})^\top u\|_{L_2(\mathbb{R})}. \tag{3.2}$$

On a bounded interval $I = (a,b) \subset \mathbb{R}$ we define for a variable order $0 \leq m(x) < 1$ the space

$$\tilde{H}^{m(x)}(I) := \left\{ u|_I \,\middle|\, u \in H^{m(x)}(\mathbb{R}), \quad u|_{\mathbb{R}\setminus\overline{I}} = 0 \right\}.$$

It coincides with the closure of $C_0^\infty(I)$ (the space of smooth functions with support compactly contained in I) with respect to the norm

$$\|u\|_{\tilde{H}^{m(x)}(I)} := \|\tilde{u}\|_{H^{m(x)}(\mathbb{R})}, \tag{3.3}$$

where $\tilde{u}$ is the zero extension of u to all of $\mathbb{R}$. It follows from Lemma 2.4 applied to $L = \Lambda^{m(x)}$ and from the positive definiteness of the Dirichlet form of the variable order Riesz potential $\Lambda^{m(x)}$ on $C_0^\infty(\mathbb{R})$ that for $0 < \underline{m} \leq m(x) \leq \overline{m} < 1$ an intrinsic norm on $\tilde{H}^{m(x)}(I)$ is given by

$$\|u\|^2_{\tilde{H}^{m(x)}(I)} := \|u\|^2_{L^2(I)} + |u|^2_{\tilde{H}^{m(x)}(I)}, \quad |u|^2_{\tilde{H}^{m(x)}(I)} := \int_\mathbb{R} \int_\mathbb{R} \frac{|\tilde{u}(x) - \tilde{u}(y)|^2}{|x - y|^{1+m(x)+m(y)}} dy dx. \tag{3.4}$$

Evidently, $\tilde{H}^{m(x)}(I) \subset H^{m(x)}(\mathbb{R})$. We also define

$$H^{m(x)}(I) = \{u \in \mathcal{D}(I) : \text{there exists an extension } lu \in H^{m(x)}(\mathbb{R}), lu|_I = u\}$$

supplied with the quotient norm $\|u\|_{H^{m(x)}(I)} := \inf\{\|lu\|_{m(x)} : lu|_I = u\}$. Then $H^{m(x)}(I)$ is isomorphic to the quotient space $H^{m(x)}(\mathbb{R})/\tilde{H}^{m(x)}(\mathbb{R}\backslash I)$.

Spaces of order $m(x) \le 0$ are defined by duality

$$(\tilde{H}^{m(x)}(I))' = H^{-m(x)}(I) \quad \text{and} \quad (H^{m(x)}(I))' = \tilde{H}^{-m(x)}(I). \tag{3.5}$$

We remark that if $m(a) > \frac{1}{2}$ then functions $u \in \tilde{H}^{m(x)}(I)$ have zero Dirichlet trace, i.e. $u(a) = 0$.

4 Spline Wavelets with Complementary Boundary Conditions

Since we discretize the parabolic equation (1.8) in $(0,T) \times I$ in the spatial variable with spline wavelet bases for $V = \tilde{H}^{m(x)}(I)$, we briefly recapitulate basic definitions and results on wavelets from e.g. [6] and the references there.

The primal wavelet bases $\psi_{l,k}$ span finite dimensional spaces

$$\mathcal{W}^l := \text{span} \{\psi_{l,k} : k \in \nabla_l\}, \quad \mathcal{V}^L := \bigoplus_{l=-1}^{L-1} \mathcal{W}^l \quad l = -1, 0, 1 \ldots,$$

and the dual spaces are defined analogously in terms of the dual wavelets $\tilde{\psi}_{l,k}$ by

$$\widetilde{\mathcal{W}}^l := \text{span} \{\tilde{\psi}_{l,k} : k \in \nabla_l\}, \quad \tilde{\mathcal{V}}^L := \bigoplus_{l=-1}^{L-1} \widetilde{\mathcal{W}}^l \quad l = -1, 0, 1 \ldots,$$

In the sequel we will require the following properties of these functions

1. Biorthogonality: the basis functions $\psi_{l,k}, \tilde{\psi}_{l,k}$ satisfy

$$\langle \psi_{l,k}, \tilde{\psi}_{l',k'} \rangle = \delta_{l,l'}\delta_{k,k'}. \tag{4.1}$$

2. Local support: the diameter of the support is proportional to a meshsize 2^{-l},

$$\text{diam supp } \psi_{l,k} \lesssim 2^{-l} \ , \ \text{diam supp } \tilde{\psi}_{l,k} \sim 2^{-l}. \tag{4.2}$$

3. Conformity: the basis functions should be sufficiently regular, i.e.

$$\mathcal{W}^l \subset \tilde{H}^1(I) \ , \ \widetilde{\mathcal{W}}^l \subset H^\delta(I) \text{ for some } \delta > 0 \ , \ l \ge -1. \tag{4.3}$$

Furthermore $\bigoplus_{l=-1}^\infty \mathcal{W}^l, \bigoplus_{l=-1}^\infty \widetilde{\mathcal{W}}^l$ are supposed to be dense in $L_2(\Omega)$

4. Vanishing moments: The primal basis functions $\psi_{l,k}$ are assumed to satisfy vanishing moment conditions up to order $p^* + 1 \geq p$

$$\langle \psi_{l,k}, x^\alpha \rangle = 0 , \ \alpha = 0, \ldots, d = p^* + 1, l \geq 0, \tag{4.4}$$

and for all dual wavelets, except the ones at each end point, one has

$$\langle \tilde{\psi}_{l,k}, x^\alpha \rangle = 0 , \ \alpha = 0, \ldots, d = p + 1, l \geq 0. \tag{4.5}$$

At the end points the dual wavelets satisfy only

$$\langle \tilde{\psi}_{l,k}, x^\alpha \rangle = 0 , \ \alpha = 1, \ldots, d = p + 1, l \geq 0. \tag{4.6}$$

On $I = (0,1)$ we set $\nabla_l = \{k = 1, \ldots, M_l\}$, where usually $M_l = 2^l$. We remark, that the third condition implies zero Dirichlet condition, namely $\psi_{l,k}(0) = \psi_{l,k}(1) = 0$. To guarantee this condition one has to sacrifice the vanishing moment property of one wavelet, e.g. $\psi_{l,k}$, $l = 0, \ldots$, at each end point 0 or 1. A systematic and general construction for arbitrary order biorthogonal spline wavelets is presented in [9]. Sufficiently far apart from the end points of $(0,1)$, biorthogonal wavelet (e.g. [6] and the references there) bases are used in this approach. Using biorthogonal wavelets in the case $p = 1$, piecewise linear spline wavelets vanishing outside $I = (0,1)$ are obtained by simple scaling. The interior wavelets have two vanishing moments and are obtained from the mother wavelet $\psi(x)$ which takes the values $(0, -\frac{1}{6}, -\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}, -\frac{1}{6}, 0, 0, 0)$ at the points $(0, \frac{1}{8}, \frac{1}{4}, \frac{3}{8}, \frac{1}{2}, \frac{5}{8}, \frac{3}{4}, \frac{7}{8}, 1)$ by scaling and translations: $\psi_{l,k}(x) := 2^{l/2}\psi(2^{l-3}x - k + 2)$ for $2 \leq k \leq 2^l - 3$ and $l \geq 3$. At the left boundary $k = 1$, we use the piecewise linear function ψ_{left} defined by the value $(0, \frac{5}{8}, \frac{-3}{4}, \frac{-1}{4}, \frac{1}{4}, \frac{1}{8}, 0, 0, 0)$ and $\psi_{right}(x) = \psi_{left}(1 - x)$. For more details we refer to [11].

The following simplified basis functions are shown to be quite efficient for the present applications [23]. They have the advantage of an extremely small support. Nevertheless, the dual wavelets do not permit compact support, but they are exponentially decaying. I.e.,

$$\left| \tilde{\Psi}_e(x) \right| \leq C \exp(-\kappa |x|), \quad \kappa > 0, x \in \mathbb{R}. \tag{4.7}$$

We emphasize that we do need the local support only of the primal wavelets for matrix compression. Therefore, this property is important for an efficient implementation. But we have assumed local support of the dual wavelets only for simplifying the present proof of the norm equivalence. However it is very likely that the present results remain valid for exponentially decaying dual wavelets.

The biorthogonal wavelets in the case $p = 1$ are continuous, piecewise linear spline wavelets vanishing outside $I = (0,1)$ (for general intervals $I = (a,b)$, they are obtained by simple scalings). The interior wavelets have two vanishing moments and are obtained from the mother wavelet $\psi(x)$ taking values $(0, -\frac{1}{2}, 1, -\frac{1}{2}, 0)$ at $(0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1)$ by scaling and translation: $\psi_{l,k}(x) := 2^{l/2}\psi(2^{l-1}x - (2k-1)2^{-2})$ for $1 \leq k \leq 2^l - 2$ and $l \geq 2$.

The boundary wavelets are constructed from the continuous, piecewise linear functions ψ_*, with values $(0, 1, -\frac{1}{2}, 0)$ at $(0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4})$, and ψ^*, taking values $(0, -\frac{1}{2}, 1, 0)$ at $(\frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1)$: $\psi_0^l = \psi_*(2^{l-1}x)$ and $\psi_{l,2^l-1} = 2^{l/2}\psi^*(2^{l-1}x - 2^{l-1} + 1)$. The following results are known for wavelets satisfying the above requirements (e.g., [6]).

For all $v_h = \sum_{l=0}^{L-1} \sum_{k=1}^{M^l} v_{l,k}\psi_{l,k} \in V_h = \mathcal{V}^L$, $h \sim 2^{-L}$, there holds the norm equivalence

$$\|v_h\|_{\tilde{H}^s(I)}^2 \cong \sum_{l=0}^{L} \sum_{k=1}^{M^l} |v_{l,k}|^2 2^{2ls}, \tag{4.8}$$

9

for all $0 \le s < \frac{3}{2}$. Validity of (4.8) in the variable order spaces $\tilde{H}^{m(x)}(I)$ will be shown in Theorem 5.9 below. Using the abbreviation $\lambda = (l,k) \in \mathcal{I} := \{\lambda = (l,k) = l = -1, 0, 1, \dots, k = 1, \dots M_l\}$ any function $v \in \tilde{H}^s(I)$, $0 \le s \le p+1$, and, due to $\tilde{H}^{\underline{m}}(I) \subset \tilde{H}^{m(x)}(I) \subset \tilde{H}^{\overline{m}}(I)$, any function in $\tilde{H}^{m(x)}(I)$ can be represented in the wavelet series

$$ v = \sum_{l=0}^{\infty} \sum_{k=1}^{M^l} v_{l,k}\psi_{l,k} = \sum_{\lambda \in \mathcal{I}} v_\lambda \psi_\lambda. $$

This makes it possible to define an approximation operator $Q_h : \tilde{H}^{m(x)}(I) \to V_h$, defined by

$$ Q_h v = \sum_{l=0}^{L} \sum_{k=1}^{M^l} v_{l,k}\psi_{l,k}. \tag{4.9} $$

For $0 \le s < \frac{3}{2} \le t \le p+1$, we have the approximation property (e.g. [6])

$$ \|v - Q_h v\|_{\tilde{H}^s(I)} \le C h^{t-s} \|v\|_{H^t(I)}. \tag{4.10} $$

5 Matrix Compression and Norm Equivalences

Consider $\Lambda^{m(x)} \in \Psi_{1,\delta}^{m(x)}$ with symbol $\langle\xi\rangle^{m(x)} \in S_{1,\delta}^{m(x)}$. Then $M := (\Lambda^{m(x)})^\top \Lambda^{m(x)}$ is, for $u \in C_0^\infty(\mathbb{R})$, given by the oscillatory integral

$$ ((\Lambda^{m(x)})^\top \Lambda^{m(x)}) u(x) := \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{i\xi(x-y)} \langle\xi\rangle^{m(y)+m(x)} u(y) \, dy \, d\xi. $$

For preconditioning as well as to establish norm equivalences in the spaces $\tilde{H}^{m(x)}(I)$ of variable order, we will consider the stiffness matrices of the operators $\Lambda^{m(x)}$ in the wavelet basis ψ_λ, i.e.

$$ \mathbf{M} = (\langle \Lambda^{m(x)}\psi_{\lambda'}, \Lambda^{m(x)}\psi_\lambda \rangle)_{\lambda,\lambda' \in \mathcal{I}} \tag{5.1} $$

and

$$ \mathbf{M}^{(-1)} = (\langle \Lambda^{-m(x)}\psi_{\lambda'}, \Lambda^{-m(x)}\psi_\lambda \rangle)_{\lambda,\lambda' \in \mathcal{I}} \tag{5.2} $$

We recall $\overline{m} := \sup_{x \in \mathbb{R}} m(x)$ and $\underline{m} := \inf_{x \in \mathbb{R}} m(x)$. For $\lambda = (l,k)$, we define the extended support Ω_λ of a wavelet basis function ψ_λ by

$$ \Omega_\lambda := \Omega_{l,k} = \bigcup_{l' \ge l} \{\operatorname{supp}\psi_{\lambda'} : \operatorname{supp}\psi_\lambda \cap \operatorname{supp}\psi_{\lambda'} \ne \emptyset\}. \tag{5.3} $$

For $\lambda = (l,k)$ we also define

$$ \underline{m}_\lambda := \inf\{m(x) : x \in \Omega_\lambda\} \quad \text{and} \quad \overline{m}_\lambda := \inf\{m(x) : x \in \Omega_\lambda\}. \tag{5.4} $$

Lemma 5.1. *Let $\psi_{l,k}, \psi_{l',k'}$ be two wavelet basis functions with compact support and with $p^*+1 \ge 2$ vanishing moments. If $dist(\operatorname{supp}\psi_{l,k}, \operatorname{supp}\psi_{l',k'}) > 0$, then we have for each $0 < \delta < 1$ the estimate*

$$ |\langle \Lambda^{m(x)}\psi_{l,k}, \Lambda^{m(x)}\psi_{l',k'}\rangle| \le C_\delta 2^{-(l+l')(1/2+d)} \operatorname{dist}(\operatorname{supp}\psi_{l,k}, \operatorname{supp}\psi_{l',k'})^{-(1+2\overline{m}+2(1-\delta)d)} \tag{5.5} $$

for come constant $C_\delta > 0$ independent of l, l' as well as

$$ |\langle \Lambda^{-m(x)}\widetilde{\psi}_{l,k}, \Lambda^{-m(x)}\widetilde{\psi}_{l',k'}\rangle| \le C_\delta 2^{-(l+l')(1/2+d)} dist(\operatorname{supp}\widetilde{\psi}_{l,k}, \operatorname{supp}\widetilde{\psi}_{l',k'})^{-(1-2\underline{m}+2(1-\delta)d)}. \tag{5.6} $$

Proof. We define the operator $A = (\Lambda^{m(x)})^\top \Lambda^{m(x)}$ with extended symbol

$$a(x, \xi, y) = \langle \xi \rangle^{m(x)+m(y)}$$

and denote by $K_A(\cdot, \cdot)$ the corresponding Schwartz kernel.

Following standard arguments in wavelet matrix compression (e.g., [6, 29]) we obtain

$$
\begin{aligned}
|\langle \Lambda^{m(x)} \psi_{l,k}, \Lambda^{m(x)} \psi_{l',k'} \rangle| &\lesssim \left| \int \int K_A(x,y) \overline{\psi_{l,k}(x)} \psi_{l',k'}(y) dx dy \right| \\
&\lesssim (\operatorname{diam} \operatorname{supp} \psi_{l,k})^d (\operatorname{diam} \operatorname{supp} \psi_{l',k'})^d (\int |\psi_{l,k}(x)dx|)(\int |\psi_{l',k'}(y)dy|) \\
&\quad \times \sup_{x \in \operatorname{supp} \psi_{l,k}} \sup_{y \in \operatorname{supp} \psi_{l',k'}} |D_x^d D_y^d K_A(x,y)| \\
&\lesssim 2^{-(l+l')(1/2+d)} \operatorname{dist}(\operatorname{supp} \psi_{l,k}, \operatorname{supp} \psi_{l',k'})^{-(1+2(1-\delta)d+2\overline{m})}.
\end{aligned}
$$

In the last step, we have used the fact that $A \in \Psi_{1,\delta}^{2m(x)} \subset \Psi_{1,\delta}^{2\overline{m}}$ together with the Caldérón Zygmund property of $K_A(x,y)$ shown in Lemma 2.4. Since $B = (\Lambda^{-m(x)})^\top \Lambda^{-m(x)} \in \Psi_{1,\delta}^{-2\underline{m}}$ a similar argument proves (5.6). $\qquad \square$

Since the boundary wavelets of the dual wavelet system do not have vanishing moments, we treat the corresponding matrix coefficients separately.

Lemma 5.2. *If among* $\widetilde{\psi}_{l,k}$, $\widetilde{\psi}_{l',k'}$ *only the latter wavelet has no vanishing moments, and if*

$$\operatorname{dist}(\operatorname{supp} \widetilde{\psi}_{l,k}, \operatorname{supp} \widetilde{\psi}_{l',k'}) > 0$$

then we have for each $0 < \delta < 1$ *the estimate*

$$|\langle \Lambda^{-m(x)} \widetilde{\psi}_{l,k}, \Lambda^{-m(x)} \widetilde{\psi}_{l',k'} \rangle| \leq C_\delta 2^{-l(1/2+d)} \operatorname{dist}(\operatorname{supp} \psi_{l,k}, \operatorname{supp} \psi_{l',k'})^{-(1-2\underline{m}+(1-\delta)d)}. \quad (5.7)$$

We need an estimate treating the case that the supports of two wavelet functions overlap. First we investigate the diagonal matrix entries.

Lemma 5.3. *Let* $0 \leq m(x)$, $x \in \mathbb{R}$ *and* $\underline{m}_\lambda$, $\overline{m}_\lambda$, $\lambda = (l, k)$ *as defined in (5.4), then*

$$2^{2l\underline{m}_\lambda} \lesssim |\langle \Lambda^{m(x)} \psi_{l,k}, \Lambda^{m(x)} \psi_{l,k} \rangle| \lesssim 2^{2l\overline{m}_\lambda}. \quad (5.8)$$

For Λ^{-1} *holds correspondingly*

$$2^{-2l\overline{m}_\lambda} \lesssim |\langle \Lambda^{-m(x)} \widetilde{\psi}_{l,k}, \Lambda^{-m(x)} \widetilde{\psi}_{l,k} \rangle| \lesssim 2^{-2l\underline{m}_\lambda}. \quad (5.9)$$

Proof. Observing $\langle \xi \rangle^{m(x)} \langle \xi \rangle^{-\overline{m}} \in S_{1,\delta}^0$ for any $\delta > 0$, and since $\Lambda^{m(x)} : \tilde{H}^{m(x)}(I) \to L^2(I)$ is one-to-one [18], we estimate

$$
\begin{aligned}
\langle \Lambda^{m(x)} \psi_{l,k}, \Lambda^{m(x)} \psi_{l,k} \rangle &= \|\Lambda^{m(x)} \psi_{l,k}\|_{L^2(I)}^2 \lesssim \|\Lambda^{\overline{m}_\lambda} \psi_{l,k}\|_{L^2(I)}^2 \\
&\lesssim \int_\xi |\int_x e^{-ix\xi} \langle \xi \rangle^{\overline{m}_\lambda} \psi_\lambda(x) dx|^2 d\xi \\
&\lesssim 2^{2l\overline{m}_\lambda}.
\end{aligned}
$$

By similar arguments we obtain using $\inf\{\theta(x) : x \in \Omega_\lambda\} \geq \underline{m}_\lambda$ the estimates

$$\langle \Lambda^{-m(x)}\tilde{\psi}_{l,k}, \Lambda^{-m(x)}\tilde{\psi}_{l,k}\rangle = \frac{1}{2\pi}\int_\xi |\int_x e^{-ix\xi}\langle\xi\rangle^{-m(x)}\tilde{\psi}_{l,k}(x)dx|^2 d\xi \lesssim \|\Lambda^{-\underline{m}_\lambda}\tilde{\psi}_{l,k}\|_0^2 \lesssim 2^{-2l\underline{m}_\lambda}.$$

We get also lower bounds. For some $C > 0$ independent of λ holds

$$\langle \Lambda^{m(x)}\psi_{l,k}, \Lambda^{m(x)}\psi_{l,k}\rangle = \frac{1}{2\pi}\int_\xi |\int_x e^{-ix\xi}\langle\xi\rangle^{m(x)}\psi_{l,k}(x)dx|^2 d\xi \geq \|\Lambda^{-\underline{m}_\lambda}\psi_{l,k}\|_0^2 \geq C2^{2l\underline{m}_\lambda}.$$

Similar arguments prove the remaining inequality. $\qquad\square$

Lemma 5.4. *In case* $\operatorname{supp}\psi_\lambda \cap \operatorname{supp}\psi_{\lambda'} \neq \emptyset$ *and* $l \leq l'$, *the following estimate holds for* $0 \leq s + \overline{m}_\lambda < \gamma$,

$$|\langle\Lambda^{m(x)}\psi_{l,k}, \Lambda^{m(x)}\psi_{l',k'}\rangle| \lesssim 2^{-|l-l'|s}2^{l\overline{m}_\lambda + l'\overline{m}_{\lambda'}} \tag{5.10}$$

Proof. We suppose that $l \geq l'$. Then, for $0 \leq s < \gamma - \overline{m}$,

$$\begin{aligned}
|\langle\Lambda^{m(x)}\psi_{l,k}, \Lambda^{m(x)}\psi_{l',k'}\rangle| &\leq \|\Lambda^{m(x)-s}\psi_{l,k}\|_0\|\Lambda^{m(x)+s}\psi_{l',k'}\|_0 \\
&\lesssim \|\Lambda^{\overline{m}_{\lambda'}-s}\psi_{l',k'}\|_0\|\Lambda^{\overline{m}_\lambda+s}\psi_{l,k}\|_0 \\
&\lesssim 2^{l\overline{m}_\lambda + l'\overline{m}_{\lambda'}}2^{-s|l-l'|}.
\end{aligned}$$

$\qquad\square$

We need to treat the dual wavelets in $H^{-m(x)}(I)$ analogously.

Lemma 5.5. *In case* $\operatorname{supp}\tilde{\psi}_\lambda \cap \operatorname{supp}\tilde{\psi}_{\lambda'} \neq \emptyset$ *and* $l \leq l'$, *the following estimate holds for* $0 \leq s + \overline{m}_\lambda < \gamma$,

$$|\langle\Lambda^{-m}\tilde{\psi}_{l,k}, \Lambda^{-m}\tilde{\psi}_{l',k'}\rangle| \lesssim 2^{-|l-l'|s}2^{-l\underline{m}_\lambda - l'\underline{m}_{\lambda'}} \tag{5.11}$$

We remark that Lemmas 5.3-5.5 include the boundary wavelets as well.

Lemma 5.6. *For* $\lambda = (l,k) \in \mathcal{I}$ *one has*

$$\overline{m}_\lambda - \underline{m}_\lambda \lesssim 2^{-l} \tag{5.12}$$

Proof. Due to $m \in C^1(\mathbb{R})$ (actually, only $m \in C^{0,1}(\mathbb{R})$ is required) there holds

$$0 \leq \overline{m}_\lambda - \underline{m}_\lambda \leq \operatorname{diam}(\Omega_\lambda)\sup_{x\in\Omega_\lambda}|m'(x)| \lesssim 2^{-l}\sup_{x\in\Omega_\lambda}|m'(x)|.$$

$\qquad\square$

Definition 5.7. *A matrix* $\mathbf{A} = (a_{\lambda,\lambda'})_{\lambda,\lambda'\in\mathcal{I}}$ *will be called* compressible, *if there exist constants* $a, b > 0$ *such that for all* $\lambda = (l,k)$, $\lambda' = (l',k')$ *holds*

$$|a_{\lambda,\lambda'}| \lesssim 2^{-(a+1/2)|l-l'|}(1 + d(\lambda,\lambda'))^{-1-b} \tag{5.13}$$

with $d(\lambda,\lambda') := 2^{\min\{\ell,\ell'\}}\operatorname{dist}(\operatorname{supp}\psi_\lambda, \operatorname{supp}\psi_{\lambda'})$ *if at least one wavelet has vanishing moments of order* $d^* \geq 2$ *and* $d(\lambda,\lambda') := +1$ *otherwise.*

Let us introduce the diagonal matrices

$$\mathbf{D}^{-m(x)} := (2^{-\overline{m}_\lambda} \delta_{\lambda,\lambda'})_{\lambda,\lambda' \in \mathcal{I}} \ , \ \ \mathbf{D}^{m(x)} := (2^{\underline{m}_\lambda} \delta_{\lambda,\lambda'})_{\lambda,\lambda' \in \mathcal{I}}. \tag{5.14}$$

Proposition 5.8. *The matrices*

$$\mathbf{A} = (a_{\lambda,\lambda'})_{\lambda,\lambda' \in \mathcal{I}} := \mathbf{D}^{-m(x)} \mathbf{M} \mathbf{D}^{-m(x)} \text{ and } \mathbf{A}^{(-1)} := \mathbf{D}^{m(x)} \mathbf{M}^{(-1)} \mathbf{D}^{m(x)}$$

with $\mathbf{M}$, $\mathbf{M}^{-1}$ *as in (5.1), (5.2) are compressible: there exists $s > 0$ such that*

$$|a_{\lambda,\lambda'}| \lesssim 2^{-|l-l'|(s+1/2)} (1 + d(\lambda,\lambda'))^{-1-2(d-\overline{m})(1-\delta)} \tag{5.15}$$

and

$$|a_{\lambda,\lambda'}^{(-1)}| \lesssim 2^{-|l-l'|(s+1/2)} (1 + d(\lambda,\lambda'))^{-1-2(d-\underline{m})(1-\delta)}. \tag{5.16}$$

Proof. If $d(\lambda, \lambda') \lesssim 1$ the assertion follows from Lemma 5.4 and Lemma 5.5.

If $d(\lambda, \lambda') \geq 1$ we infer from Lemma 5.1, that for $l \leq l'$ holds

$$\begin{aligned}
|a_{\lambda,\lambda'}| &\lesssim & 2^{l(1/2+\overline{m}+2(1-\delta)d)-d)} 2^{-l'(1/2+d-\overline{m})} (1 + d(\lambda,\lambda'))^{-(1+2\overline{m}+(1-\delta)2d)} \\
&\lesssim & 2^{(l-l')(1/2+s)} (1 + d(\lambda,\lambda'))^{-(1+2\overline{m}+(1-\delta)2d))},
\end{aligned}$$

provided that $0 < s = \min\{\overline{m} + 2(1-\delta)d - d, d - \overline{m}\}$, which is possible since $\delta > 0$ can be arbitrarily small and $\overline{m} < 1$ if X_t is quadratic, pure jump.

An estimate for $\mathbf{A}^{(-1)}$ can be derived similarly. The case that the dual wavelets are boundary wavelets is treated similarly if only one of the wavelets has vanishing moments.

The case that both wavelets do not have vanishing moments is treated in the same way as the case of overlapping support. $\qquad \square$

The previous result is the basis for the numerical analysis of variable order pseudodifferential equations. It implies multilevel norm equivalences in the variable order Sobolev spaces $\tilde{H}^{m(x)}(I)$ and diagonal preconditioning for wavelet discretizations of the variable order pseudo-differential operators $\Lambda^{m(x)}$.

Theorem 5.9. *Let $u = \sum_{\lambda \in \mathcal{I}} u_\lambda \psi_\lambda = \mathbf{u}^\top \Psi \in \tilde{H}^{m(x)}(I)$. Then for $\mathbf{D}^{m(x)}$ as in (5.14) holds the multilevel norm equivalence in variable order spaces:*

$$\|u\|_{\tilde{H}^{m(x)}(I)}^2 \sim \mathbf{u}^\top (\mathbf{D}^{m(x)})^\top \mathbf{D}^{m(x)} \mathbf{u} \sim \mathbf{u}^\top \mathbf{D}^{2m(x)} \mathbf{u}. \tag{5.17}$$

For $v = \sum_{\lambda \in \mathcal{I}} v_\lambda \widetilde{\psi}_\lambda = \mathbf{v}^\top \widetilde{\Psi} \in H^{-m(x)}(I)$ holds correspondingly

$$\|v\|_{H^{-m(x)}(I)}^2 \sim \mathbf{v}^\top \mathbf{D}^{-2m(x)} \mathbf{v}. \tag{5.18}$$

The spectral condition number of $\mathbf{A} = \mathbf{D}^{-m(x)} \mathbf{M} \mathbf{D}^{-m(x)}$ is uniformly bounded in h.

Proof. As a first step, we need to show that $\operatorname{cond} \mathbf{D}^{-m(x)} \mathbf{D}^{m(x)} \leq C$. Recalling the definition we observe that $\mathbf{D}^{-m(x)} \mathbf{D}^{m(x)} = \operatorname{diag}(2^{-l\underline{m}_\lambda} 2^{l\overline{m}_\lambda})$. Due to Lemma 5.6 we get $\lim_{l \to \infty} 2^{l|\underline{m}_\lambda - \overline{m}_\lambda|} =$

$\lim_{l\to\infty} 2^{lC(2^{-l})} = 1$. Therefore we estimate $1 \le 2^{l|\underline{m}_\lambda - \overline{m}_\lambda|} \le C$. It is well known that the spectral norm of a compressible matrix is bounded, see e.g. [6] Theorem 4.6.1. Therefore, it follows from this result together with Proposition 5.8 that the spectral norm of the matrices $\mathbf{A}$ and $\mathbf{A}^{-1}$ are bounded, which means $\operatorname{cond}\mathbf{A} \sim 1$. Thus the matrices $\mathbf{M}$, $\mathbf{D}^{m(x)}(\mathbf{D}^{m(x)})^\top$ and $\mathbf{D}^{2m(x)}$ are spectrally equivalent. Since $\|u\|^2_{\tilde{H}^{m(x)}(I)} = \mathbf{u}^\top\mathbf{M}\mathbf{u}$, we have established the first assertion (5.17). The second equivalence (5.18) follows analogously.

$\square$

An immediate consequence of the norm-equivalence (5.17) is the uniform boundedness of the quasi-interpolants Q_h in the variable order norms $\tilde{H}^{m(x)}(I)$.

Corollary 5.10. *The quasi-interpolants Q_h defined in (4.9) are uniformly bounded in $\tilde{H}^{m(x)}(I)$, i.e. there is $C > 0$ such that*

$$\forall u \in \tilde{H}^{m(x)}(I), \quad \forall h > 0: \qquad \|Q_h u\|_{\tilde{H}^{m(x)}(I)} \le C\|u\|_{\tilde{H}^{m(x)}(I)}.$$

Using the variable order norm equivalence (5.17) we can extend the approximation property (4.10) to variable order spaces.

Corollary 5.11. *Let $u \in H^s(I) \cap \tilde{H}^1(I)$ for some $1 \le s \le p+1$. Then for the quasi-interpolant $u_h = Q_h u = \sum_{l=0}^{L}\sum_{j=1}^{M^l} u_j^l \psi_j^l$ in (4.9) there holds for $0 < \overline{m} < 1 \le s \le p+1$ the Jackson estimate*

$$\|u - u_h\|^2_{\tilde{H}^{m(x)}(I)} \lesssim \int_I 2^{2L(m(x)-s)}(|D^s u(x)|^2 + |u(x)|^2)dx \lesssim 2^{2L(\overline{m}-s)}\|u\|^2_{H^s(I)} \qquad (5.19)$$

Proof. For multi-indices $\lambda = (l,k), \mu = (L,k') \in \mathcal{I}$, we introduce the notation $\lambda \succeq \mu$ if $l := |\lambda| \ge |\mu| =: L$ and $\operatorname{supp}\psi_\lambda \cap \operatorname{supp}\psi_\mu \ne \emptyset$. For $s \ge \frac{3}{2}$ we choose $s' < s$ with $1 \le s' < \frac{3}{2}$, otherwise we set $s' = s$, we observe that $\underline{m}_\lambda - s' \le (\overline{m}_\mu - s') < 0$ holds for all $\lambda \succeq \mu$. Therefore we conclude from the norm equivalence (5.17)

$$
\begin{aligned}
\|u - u_h\|^2_{\tilde{H}^{m(x)}(I)} &\sim \sum_{l>L}\sum_{k=1}^{M_l} 2^{2|\lambda|\underline{m}_\lambda}|u_{l,k}|^2 = \sum_{l>L}\sum_{k=1}^{M_l} 2^{2|\lambda|(\underline{m}_\lambda - s')}2^{2s'|\lambda|}|u_\lambda|^2 \\
&\lesssim \sum_{\mu\in\nabla_L} 2^{2L(\overline{m}_\mu - s')}\sum_{\lambda\succeq\mu} 2^{2s'|\lambda|}|u_\lambda|^2,
\end{aligned}
$$

where $\nabla_L = \{\mu = (L,k') : k' = 1, \ldots, M_L\}$.

Let $\mu = (L,k')$, $L = |\mu|$ and $\square_\mu := [2^{-|\mu|}k', 2^{-|\mu|}(k'+1)]$. Then, due to the norm equivalence (5.17) and the approximation property (4.10) we have

$$\sum_{\mu\in\nabla_L}\sum_{\lambda\succeq\mu} 2^{2s'|\lambda|}|u_\lambda|^2 \lesssim \sum_{\mu\in\nabla_L} 2^{2L(s'-s)}\int_{\square_\mu}(|D^s u(x)|^2 + |u(x)|^2)dx.$$

Recalling that $2^{L\underline{m}_\mu} \sim 2^{Lm(x)} \sim 2^{L\overline{m}_\mu}$ holds for $x \in \square_\mu$, we obtain the final result

$$\|u - u_h\|^2_{\tilde{H}^{m(x)}(I)} \lesssim \int_I 2^{2L(m(x)-s)}(|D^s u(x)|^2 + |u(x)|^2)dx.$$

14

Recalling that $2^{L\underline{m}_\lambda} \sim 2^{Lm(x)} \sim 2^{L\overline{m}_\lambda}$ holds for $x \in \square_\lambda$, we obtain the final result

$$\|u - u_h\|_{\tilde{H}^{m(x)}(I)}^2 \lesssim \int_I 2^{2L(m(x)-s)}(|D^s u(x)|^2 + |u(x)|^2)dx.$$

$\square$

Lemmas 5.1 to 5.5 addressed only the generators $\Lambda^{m(x)}$. We focussed on this case since it is crucial for establishing the norm equivalences (5.17). For wavelet solution of the Kolmogoroff equation (1.8) in log-linear complexity we need the compressibility of the moment matrices in wavelet bases also for generators which are more general than $\Lambda^{m(x)}$.

Theorem 5.12. *Let $A \in \Psi_{\rho,\delta}^{2m(x)}(\mathbb{R})$ for some $0 < \delta < \rho \le 1$, and let $\{\psi_\lambda\}_\lambda$ be a system of compactly supported spline wavelets in $I = (0,1)$ of polynomial degree $p \ge 1$ as introduced in Section 4 with complementary boundary conditions and $d \ge p+1$ vanishing moments if $\operatorname{supp}(\psi_\lambda) \cap \partial I = \emptyset$.*

Then the preconditioned moment matrix

$$\mathbf{A} := \mathbf{D}^{-m(x)}((\langle A\psi_{\lambda'}, \psi_\lambda \rangle)_{\lambda,\lambda' \in \mathcal{I}}\mathbf{D}^{-m(x)} \tag{5.20}$$

has condition number which is uniformly bounded in L and $\mathbf{A}$ is compressible: there are $C > 0$, $s > 0$ such that for all $\lambda, \lambda' \in \mathcal{I}$ holds

$$|\mathbf{A}_{\lambda,\lambda'}| \le C2^{-|l-l'|(s+1/2)}(1 + d(\lambda,\lambda'))^{-1-2(d-\overline{m})(\rho-\delta)} \tag{5.21}$$

The proof of Theorem 5.12 is completely analogous to that of Proposition 5.8, using Lemmas 5.1 to 5.5 with the kernel estimates (2.14) in place of (2.6).

6 Option Pricing under Variable Intensity Feller Processes

With the norm equivalences (5.17), (5.18) we develop linear complexity numerical solution algorithms for the pricing equation (1.8). Since elements of the algorithm are identical to [26, 22, 23], we focus only on the essential differences in the analysis due to the variable order $m(x)$ of A_X.

6.1 Weak Form of the Pricing Equation. Well-posedness.

We consider (1.8) with nonlocal operator $A \in \mathcal{L}(V, V^*)$ of variable order $2m(x) \in (0,2)$ in $V = \tilde{H}^{m(x)}(I)$, $V^* = H^{-m(x)}(I)$ with Dirichlet form

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

This Dirichlet form is continuous, i.e. there exists a constant $\alpha > 0$ such that

$$|a(u,v)| \le \alpha \|u\|_V \|v\|_V, \qquad \forall u, v \in V. \tag{6.1}$$

It is well known that problem (1.8) has a unique (weak) solution $u(t)$, and there holds the a priori estimate (see e.g. [19])

$$\|u\|_{\mathcal{C}(\overline{J}, L^2(I))} + \|u\|_{L^2(J,V)} + \|\dot{u}\|_{L^2(J,V^*)} \le C\|u_0\|_{L^2(I)}$$

15

provided that for some constants $\gamma > 0$, $C \geq 0$ the Dirichlet form $a(\cdot, \cdot)$ satisfies the Gårding inequality, i.e.,

$$\forall\, u \in V \qquad \Re a(u, u) \geq \gamma \|u\|_V^2 - C \|u\|_{L^2(I)}^2. \tag{6.2}$$

Remark 6.1. Without loss of generality, we may assume in (6.2) that $C = 0$ since the substitution $\bar{u} = e^{-\beta t} u$ results in (1.8) with the shifted operator $A + \beta I$ which is definite for sufficiently large $\beta > 0$ by (6.2) ($\beta > 2C$ will suffice).

Note that for asymmetric distributions of positive and negative jumps in X_t, the generator A needs not to be self-adjoint, i.e. $a(\cdot, \cdot)$ is generally non-symmetric.

6.2 Gårding Inequality. Analyticity of T_t

To establish (6.2), we start with the variable order Riesz potential $A = \Lambda^{2m(x)} \in \Psi_{1,\delta}^{2m(x)}$ given by

$$Au(x) := \frac{1}{2\pi} \int \int e^{i\xi \cdot (x-y)} \langle \xi \rangle^{2m(x)} u(y) dy d\xi, \quad u \in C_0^\infty(I), \tag{6.3}$$

and introduce the operator $L = \Lambda^{m(x)} \in \Psi_{1,\delta}^{m(x)}$.

Lemma 6.2. *For every $\delta > 0$ sufficiently small there exist $\varepsilon > 0$ and pseudo-differential operators $B, R \in \Psi_{1,\delta}^{-\varepsilon}$ with symbols $b(x, \xi), r(x, \xi) \in S_{1,\delta}^{-\varepsilon}$ such that*

$$A = L \circ L^\top + L \circ B + R. \tag{6.4}$$

Proof. Since the the extended symbol $a(x, \xi, y)$ of $LL^\top$ is in $S_{1,\delta}^{2\overline{m}}$, and since $I \subset \mathbb{R}$ is compact we can assume that $A(x, D, y)$ is compactly supported. We also may apply standard pseudo differential operator calculus and expand

$$a(x, \xi, y) = a_0(x, \xi) + a_1(x, \xi) + r_1(x, \xi, y) = \langle \xi \rangle^{2m(x)} + D_y D_\xi (\langle \xi \rangle)^{m(x)+m(y)}|_{y=x} + r_1(x, \xi, y).$$

Therefore, the subprincipal symbol of $a(x, \xi)$ is

$$a_1(x, \xi) = \langle \xi \rangle^{m(x)} \langle \xi \rangle^{m(x)-1} \log \langle \xi \rangle.$$

We represent the corresponding operator $a_1(x, D)$, using symbolic calculus, as

$$a_1(x, D) = L \circ \langle D \rangle^{m(x)-1} \log \langle D \rangle + R_2 = L \circ B + R_1,$$

where $R_2 \in \Psi_{1,d}^{-\varepsilon}$ for sufficiently small ε. The symbol of the operator B is $b(x, \xi) = \langle \xi \rangle^{m(x)-1} \log \langle \xi \rangle \in S_{I,\delta}^{-\varepsilon}$, provided that $0 < \varepsilon < 1 - \delta - \overline{m}$. Setting $R = R_1 + R_2$ gives the claimed result. $\square$

Lemma 6.3. *Let $B \in \Psi_{1,\delta}^{-\varepsilon}$ and $b \in S_{1,\delta}^{-\varepsilon}$ as in Lemma 6.2. Then, for each $\epsilon > 0$ there exists $b_\epsilon \in S_{1,\delta}^{-2}$ such that*

$$|(b - b_\epsilon)(x, \xi)| \leq \epsilon. \tag{6.5}$$

16

Proof. Since $b \in S_{1,\delta}^{-\epsilon}$, for any $\epsilon > 0$ there exists R_ϵ such that

$$|b(x,\xi)| \le \epsilon\,, \ |\xi| \ge R_\epsilon$$

holds. Let χ be the cut-off function defined in the proof of Lemma 2.4, we set $\hat{\chi}_{R_\epsilon}(\xi) := \hat{\chi}((2R)_\epsilon^{-1}(\xi - \frac{1}{2}))$ and $b_\epsilon(x,\xi) := b(x,\xi)\chi_{R_\epsilon}(\xi)$. Then

$$b_\epsilon(x,\xi) = 0 \quad x \in \mathbb{R}, \quad |\xi| > 2R_\epsilon$$

i.e. $b_\epsilon \in S_{1,\delta}^{-k}$ for all $k \in \mathbb{N}$ and b_ϵ satisfies the desired estimates (6.5). $\qquad\square$

Lemma 6.4. *Let the operators B be defined as in Lemma 6.2, and B_ϵ be defined by the symbol b_ϵ in Lemma 6.3. Then*

$$|\langle L \circ (B - B_\epsilon)u, v\rangle| \lesssim \epsilon \|u\|_{\tilde{H}^{m(x)}(I)} \|v\|_{\tilde{H}^{m(x)}(I)} \tag{6.6}$$

and

$$|\langle L \circ B_\epsilon u, u\rangle| \lesssim \epsilon \|u\|_{L_2(I)}^2 \tag{6.7}$$

hold for all $u, v \in \widetilde{H}^{m(x)}(I)$.

Proof. Since u, v have compact support, we can without assume loss of generality that $x \mapsto b(x,\xi) - b_\epsilon(x,\xi)$ has compact support with respect to x, which simplifies the proof. We proceed in a standard way (see e.g [32] Chapter VI, §2), by proving that under this assumption Lemma 6.3 ensures that for all $\epsilon > 0$ there exists B_ϵ satisfying

$$\|(B - B_\epsilon)u\|_{L_2(\mathbb{R})} \le \epsilon \|u\|_{L_2(I)}.$$

To this end, let us consider the function $\hat{d}(\lambda,\xi) = \frac{1}{2\pi} \int e^{-ix\lambda}(b(x,\xi) - b_\epsilon(x,\xi))dx$. Then, due to the local support of $x \mapsto b(x,\xi) - b_\epsilon(x,\xi)$, for each $N \in \mathbb{N}$, there holds an estimate

$$\sup_\xi \hat{d}(\lambda,\xi) \le C_N \epsilon (1 + |\lambda|)^{-N}.$$

We estimate

$$
\begin{aligned}
(B - B_\epsilon)u(x) &= \frac{1}{2\pi} \int\int e^{i\xi(x-y)}(b(x,\xi) - b_\epsilon(x,\xi))u(y)dyd\xi \\
&= \frac{1}{2\pi} \int\int (\int \hat{d}(\lambda,\xi)e^{i\lambda x}d\lambda)e^{i\xi(x-y)}u(y)dyd\xi \\
&= \int (e^{ix\lambda}\int(e^{ix\xi}\hat{d}(\lambda,\xi)\hat{u}(\xi)d\xi))d\lambda =: \int (T_\lambda u)(x)d\lambda.
\end{aligned}
$$

By Plancherels Theorem we get

$$\|T_\lambda u\|_{L_2(\mathbb{R})} \lesssim \sup_\xi |\hat{d}(\lambda,\xi)|\|\hat{u}\|_{L_2(\mathbb{R})} \lesssim \epsilon(1 + |\lambda|)^{-2}\|u\|_{L_2(\mathbb{R})}.$$

Integrating with respect to λ gives

$$\|(B - B_\epsilon)u\|_{L_2(\mathbb{R})} \lesssim \int \|T_\lambda u\|_{L_2(\mathbb{R})}d\lambda \lesssim \int ((1 + |\lambda|)^{-2})\epsilon\|u\|_{L_2(\mathbb{R})}d\lambda \lesssim \epsilon\|u\|_{L_2(\mathbb{R})}.$$

17

Therefore, we estimate

$$
\begin{aligned}
|\langle L \circ (B - B_\epsilon)u, v\rangle| &\leq \|(B - B_\epsilon)u\|_{L_2(\mathbb{R})}\|\Lambda^{m(x)}v\|_{L_2(\mathbb{R})} \\
&\leq \epsilon\|u\|_{L_2(I)}\|v\|_{\tilde{H}^{m(x)}(I)} \\
&\leq \epsilon\|v\|_{\tilde{H}^{m(x)}(I)}\|u\|_{\tilde{H}^{m(x)}(I)},
\end{aligned}
$$

for all $u, v \in \widetilde{H}^m(I)$, which concludes the first estimate (6.6).

Since $\overline{m} < 1$ and $b_\epsilon \in S_{1,\delta}^{-2}$, standard pseudo differential operator calculus gives $L \circ B_\epsilon \in \Psi_{1,\delta}^{-\varepsilon}$ and we obtain

$$
|\langle L \circ B_\epsilon u, u\rangle| \lesssim \|u\|_{L_2(I)}^2.
$$

$\square$

Now we establish the Gårding inequality in an important special case. We recall the definition (3.3) of the $\tilde{H}^{m(x)}(I)$-norm.

Lemma 6.5. *The variable order Riesz potential operator A given by (6.3) satisfies the Gårding inequality: there exist constants $\gamma > 0$ and $C \geq 0$ such that*

$$
\forall u \in \tilde{H}^{m(x)}(I): \quad \Re a(u, u) = \frac{1}{2}\left(\langle Au, u\rangle + \langle u, Au\rangle\right) \geq \gamma\|u\|_{\tilde{H}^{m(x)}(I)}^2 - C\|u\|_{L_2(I)}^2. \tag{6.8}
$$

Furthermore, the operator A is continuous,

$$
|\langle Au, v\rangle| \lesssim \|u\|_{\tilde{H}^{m(x)}(I)}\|v\|_{\tilde{H}^{m(x)}(I)} \quad \forall u, v \in \widetilde{H}^{m(x)}(I). \tag{6.9}
$$

Proof. We decompose

$$
A = L \circ L^\top + L \circ (B - B_\epsilon) + L \circ B_\epsilon + R
$$

where $R \in \Psi_{1,\delta}^{-\varepsilon}$. A similar expression holds for the adjoint $A^\top$. Using Lemma 6.4 there exist constants $\gamma > 0, C' \geq 0$ such that for $\epsilon > 0$ sufficiently small

$$
\begin{aligned}
\forall u \in \tilde{H}^{m(x)}(I): \quad \langle Au, u\rangle + \langle u, Au\rangle &\geq 2\langle \Lambda^{m(x)}u, \Lambda^{m(x)}u\rangle - \epsilon\|u\|_{\tilde{H}^{m(x)}(I)}^2 - C\|u\|_{L_2(I)}^2 \\
&\geq \gamma\|u\|_{\tilde{H}^{m(x)}(I)}^2 - C'\|u\|_{L_2(I)}^2,
\end{aligned}
$$

which proves the Gårding inequality (6.8).

The second assertion follows by the same ingredients observing the L_2-continuity of R and B_ϵ. $\square$

The previous theorem has been shown in [12] for a larger class of operators under the more restrictive condition $\overline{m} - \underline{m} < \frac{1}{2}$.

The Gårding inequality (6.8) holds for the operator $A = \Lambda^{2m(x)}$ defined in (6.3). Lemma 6.5 is, however, the key ingredient in establishing a Gårding inequality for variable order operators $A(x, D) \in \Psi_{\rho,\delta}^{2m(x)}$ which are elliptic in a suitable sense.

18

Theorem 6.6. *Let* $A(x, D) \in \Psi_{\rho,\delta}^{2m(x)}$ *be a pseudodifferential operator of variable order* $2m(x)$, $0 < m(x) < 1$, *given by (2.3) with symbol* $a(x, \xi) \in S_{\rho,\delta}^{2m(x)}$ *for some* $0 < \delta < \rho \leq 1$ *for which there exists* $\gamma > 0$ *with*

$$\forall x, \xi \in \mathbb{R}^d : \qquad \Re e\, a(x, \xi) + 1 \geq \gamma \langle \xi \rangle^{2m(x)}. \tag{6.10}$$

Then $A(x, D) \in \Psi_{\rho,\delta}^{2m(x)}$ *satisfies a Gårding inequality in the variable order space* $\tilde{H}^{m(x)}(I)$*: there are constants* $\gamma > 0$ *and* $C \geq 0$ *such that*

$$\forall u \in \tilde{H}^{m(x)}(I) : \quad \Re a(u, u) \geq \gamma \|u\|_{\tilde{H}^{m(x)}(I)}^2 - C\|u\|_{L_2(I)}^2, \tag{6.11}$$

and

$$\exists \lambda > 0 \quad \text{such that } A(x, D) + \lambda I : \tilde{H}^{m(x)}(I) \to H^{-m(x)}(I) \quad \text{is invertible.} \tag{6.12}$$

Proof. Let $I = (a, b) \subset \mathbb{R}$ be a bounded interval and $m(x) : I \to (0, 1)$ a variable order. As in the proof of Lemma 6.5, we can decompose A into

$$A = N \circ N^\top + N(B - B_\epsilon) + N \circ B_\epsilon + R,$$

where N is an injective elliptic operator on I with symbol $n(x, \xi)$ satisfying $n(x, \xi) \geq c\langle \xi \rangle^{m(x)}$ and B, B_ϵ and R have similar properties as in Lemma 6.2 - 6.4. Due to these assumptions there is $c > 0$ such that for all $u \in \tilde{H}^{m(x)}(I)$ we have $\|Nu\|_{L^2(I)} \geq c\|u\|_{\tilde{H}^{m(x)}(I)}$. This implies in the same fashion as in the previous proof

$$\Re a(u, u) = \frac{1}{2} \left(\langle Au, u \rangle + \langle u, Au \rangle \right) = c\gamma \left((Nu, Nu) - \|u\|_{L^2(I)}^2 \right) \geq \gamma'(\|u\|_{\tilde{H}^{m(x)}(I)}^2 - C'\|u\|_{L^2(I)}^2).$$

Where we conclude (6.8) with Lemmas 6.2 to 6.5.

The invertibility (6.12) of $A(x, D) + \lambda I$ follows from (6.8): it implies that for $\lambda > 0$ sufficiently large, $A(x, D) + \lambda I$ is $\tilde{H}^{m(x)}(I)$-coercive. $\quad\square$

The Gårding inequality (6.11) implies time-analyticity of the semigroup T_t corresponding to $X_t \in FP_{\rho,\delta}^{2m(x)}$ and $A \in \Psi_{\rho,\delta}^{2m(x)}$. As in [30], Theorem 1, we obtain from (6.8)

Theorem 6.7. *Let* $X \in FP_{\rho,\delta}^{2m(x)}$ *be a quadratic, pure jump Feller process of variable intensity* $2m(x)$ *with characteristic function* $a(x, \xi)$ *as in (1.6) belonging to* $S_{\rho,\delta}^{2m(x)}$*, which we assume to be coercive in the sense that (6.10) holds. Then* $A(x, D)$ *is the infinitesimal generator of the Feller semigroup* T_t *of* X_t *defined in (1.2). This semigroup is analytic in the sense, that there exist* $C, d > 0$*, such that for all* $t > 0$, $l \in \mathbb{N}_0$ *and* $0 \leq \theta \leq 1$ *holds*

$$\left\| T^{(l)}(t) \right\|_{\mathcal{L}(V_\theta, V)}^2 \leq Cd^{2l+1-2\theta} \Gamma(2l + 2 - 2\theta) t^{-2(l+1)+2\theta}. \tag{6.13}$$

Here, $V_\theta = (H^{-m(x)}(I), \tilde{H}^{m(x)}(I))_{\theta,2}$ *denotes the real interpolation space with index* $0 \leq \theta \leq 1$ *with the convention that* $V = V_1 = \tilde{H}^{m(x)}(I)$.

Proof. If the negative definite characteristic function $a(x, \xi)$ in (1.6) belongs to $S_{\rho,\delta}^{2m(x)}$, by Theorem 1.3 of [28], the operator $A(x, D)$ extends to a Feller generator given by (1.3). Assumption (6.10) implies by Theorem 6.6 and by Theorem 1 in [30] the analytic regularity estimate (6.13) for the Feller semigroup T_t. $\quad\square$

6.3 Discretization of the Kolmogoroff Equation

We discretize (1.8) with wavelets on $I = (a, b)$, and, to exploit the time-analyticity (6.13) of $u(t)$, by a hp-discontinuous Galerkin (dG) scheme on $J = (0, T)$, following [30, 22, 26]. In our analysis, we will need to consider functions in $V = \widetilde{H}^{m(x)}(I)$ with additional regularity. Therefore, for $s \geq 0$, we define the spaces

$$\mathcal{H}^s(I) = \begin{cases} \widetilde{H}^s(I) & \text{for } s \leq \overline{m}, \\ V \cap H^s(I) & \text{for } s > \overline{m}. \end{cases}$$

By $\|\cdot\|_s$, we denote the corresponding norm in $\mathcal{H}^s(I)$.

6.3.1 Spatial Semi-Discretization by Wavelets

We first discretize (1.8) with respect to the space variable. Hence, let $\mathcal{T}^0$ be a fixed coarse partition of I. Furthermore, define the mesh $\mathcal{T}^l$, for $l > 0$, recursively by bisection of each interval in $\mathcal{T}^{l-1}$. We assume that our computational mesh $\mathcal{T}_h$ (with mesh size h) is obtained in this way as $\mathcal{T}^L$, for some $L > 0$, with $h = C2^{-L}$.

The finite element space $V_h \subset V$ used for the discretization is the space of all continuous piecewise polynomials of degree $p \geq 1$ on the triangulation $\mathcal{T}_h$ which vanish on the boundary ∂I. 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 (1.8) reads: Given $u_0 \in L^2(I)$, find $u_h \in H^1(J, V_h)$ such that

$$\frac{d}{dt}(u_h, v_h) + a(u_h, v_h) = 0, \qquad \forall\, v_h \in V_h, \quad \text{and} \quad u_h(0) = P_h u_0. \tag{6.14}$$

Here, P_h is the L^2 projection onto V_h. Using the stability of the quasi interpolant Q_h in (4.9), i.e. Corollary 5.10, the following result on the spatial semi-discretization can be proved as in [23], Section 5, using the approximation property Corollary 5.11.

Theorem 6.8. *Let* (6.1) *and* (6.8) *be satisfied. Furthermore, we assume that the operator A is of the form $A = A_0 + \mathcal{B}$, where $A_0 = C_0 \Lambda^{2m(x)} \in \Psi_{1,\delta}^{2m(x)}$, $C_0 > 0$ and $\delta > 0$ arbitrarily small, is a variable order Riesz potential, and $\mathcal{B} : V \to V^*$ is a compact perturbation in $V = \tilde{H}^{m(x)}(I)$.*

Then, for any $t > 0$, there holds

$$\|u(t) - u_h(t)\|_V \leq C \min\{1, h^{p+1} t^{-\frac{p+1}{m}}\}. \tag{6.15}$$

Here, $C > 0$ is a constant independent of h and t, and u, u_h are solutions of (1.8) *and* (6.14), *respectively.*

In order to compute a fully discrete approximation (in space and time) to the parabolic problem (1.6), systems of linear equations have to be solved in each implicit time step (of the corresponding time discretization). To obtain an efficient solution method, we use wavelet matrix compression of the dense matrices corresponding to the operator A into sparse ones and we use optimal preconditioning. The methodology is as in [22, 23], but in the variable order case it is based on the norm equivalence Theorem 5.9 and the approximation property Corollary 5.11. After compression and preconditioning,

the systems may be solved in linear or log-linear complexity. Since the proofs are analogous to what was done in [22, 23], we only give brief outlines of the argument here.

The restriction of the Dirichlet form a to $V_h \times V_h$ corresponds to a matrix $\mathbf{A}$ with entries $A_{\lambda,\lambda'} = A_{(l,k),(l',k')} = a(\psi_{l',k'}, \psi_{l,k})$ which, due to the Caldéron-Zygmund estimates (2.14) and the vanishing moment conditions (4.4) - (4.6), decay with increasing distance of their supports. Hence, we can define a compressed matrix $\widetilde{\mathbf{A}}$ and a corresponding sesquilinear form $\widetilde{a}$ by replacing some of the small entries in $\mathbf{A}$ with zero:

$$\widetilde{A}_{(l,k),(l',k')} = \begin{cases} A_{(l,k),(l',k')} & \text{if } \operatorname{dist}(\Omega_{l,k}, \Omega_{l',k'}) \leq \delta_{l,l'} \text{ or } \Omega_{l,k} \cap \partial I \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{6.16}$$

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

$$\delta_{l,l'} := c \max\{2^{-L+\widehat{\alpha}(2L-l-l')}, 2^{-l}, 2^{-l'}\}, \tag{6.17}$$

with some parameters $c > 0$ and $0 < \widehat{\alpha} \leq 1$, and $\Omega_{l,k} = \operatorname{supp} \psi_{l,k}$.

By continuity (6.9) and coercivity (6.8) with $C' = 0$ (which can always be achieved by Remark 6.1) of the Dirichlet form a, we have the norm equivalence

$$\|u\|_a = |a(u,u)|^{\frac{1}{2}} \cong \|u\|_V \cong \|u\|_{\tilde{H}^{m(x)}(I)}.$$

As in [25, Proposition 3.2], we have

Proposition 6.9. *Assume c in (6.17) is sufficiently large. Then, there exists $0 < \widetilde{\beta} \leq \widetilde{\alpha} < \infty$ independent of h such that*

$$|\widetilde{a}(u_h, v_h)| \leq \widetilde{\alpha} \|u_h\|_a \|v_h\|_a, \tag{6.18}$$

and

$$\Re\widetilde{a}(u_h, u_h) \geq \widetilde{\beta} \|u_h\|_a^2, \tag{6.19}$$

for all $u_h, v_h \in V_h$.

Moreover, there holds [25]

Proposition 6.10. *Assume c in (6.17) is sufficiently large. Then, there exists $0 < \tau < 1$ independent of h such that for all $L > 0$ the following consistency condition is satisfied:*

$$|a(u_h, v_h) - \widetilde{a}(u_h, v_h)| \leq \tau \|u_h\|_a \|v_h\|_a, \quad \forall u_h, v_h \in V_h. \tag{6.20}$$

In addition, if

$$\widehat{\alpha} > \frac{2p+2}{2p+2+2\inf\{x : m(x)(\rho-\delta)\}} = \frac{p+1}{p+1+\underline{m}(\rho-\delta)}, \tag{6.21}$$

where $\widehat{\alpha}$ is the constant from (6.17), then, there holds

$$|a(Q_h u, v_h) - \widetilde{a}(Q_h u, v_h)| \leq Ch^{s-\overline{m}} \|u\|_s \|v_h\|_V, \quad \forall u \in \mathcal{H}^s(\Omega), \forall v_h \in V_h, \tag{6.22}$$

and

$$|a(Q_h u, Q_h v) - \widetilde{a}(Q_h u, Q_h v)| \leq Ch^{s+s'-2\overline{m}} \|u\|_s \|v\|_{s'}, \quad \forall u \in \mathcal{H}^s(\Omega), \forall v \in \mathcal{H}^{s'}(\Omega), \tag{6.23}$$

for all $\overline{m} \leq s, s' \leq p+1$. Here, Q_h is a projection or quasi-interpolation operator from (4.9).

The matrix compression (6.16) reduces the number of nonzero elements from N^2 in $\mathbf{A}$ to N times a logarithmic term in $\widetilde{\mathbf{A}}$; see e.g. [29, 6]. More precisely, there holds:

Proposition 6.11. *For $\widehat{\alpha} < 1$ in (6.21), the number of nonzero elements in $\widetilde{\mathbf{A}}$ is $\mathcal{O}(N \log N)$. If $\widehat{\alpha} = 1$, then the number of nonzero elements in $\widetilde{\mathbf{A}}$ is $\mathcal{O}(N(\log N)^2)$.*

The matrix compression from the previous section induces, instead of (6.14), a perturbed spatial semi-discretization of (1.8): find $\widetilde{u}_h \in V_h$ such that

$$\frac{d}{dt}(\widetilde{u}_h, v_h) + \widetilde{a}(\widetilde{u}_h, v_h) = 0, \qquad \forall\, v_h \in V_h. \quad \text{with } \widetilde{u}_h(0) = P_h u_0. \tag{6.24}$$

6.3.2 Discontinuous Galerkin Time Discretization

In order to obtain a fully discrete approximation (in space and time) to the parabolic problem (1.8), we discretize (6.24) in time using a discontinuous Galerkin timestepping following [30].

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

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

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 $\underline{r} = \{r_m\}_{m=1}^M$. Discontinuous Galerkin time stepping yields an approximate solution in the space

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

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

With these definitions, the fully discrete dG scheme for the solution of problem (1.8), respectively (6.24), reads as follows: find $\widetilde{U}_h^{dG} \in \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h)$ such that for all $W \in \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h)$

$$\widetilde{B}_{dG}(\widetilde{U}_h^{dG}, W) = F_{dG}(W), \tag{6.25}$$

where

$$\widetilde{B}_{dG}(U, W) = \sum_{m=1}^M \int_{I_m} \left((\dot{U}, W) + \widetilde{a}(U, W) \right) dt + \sum_{m=1}^{M-1} ([\![U]\!]_m, W_m^+) + (U_0^+, W_0^+), \tag{6.26}$$

and

$$F_{dG}(W) = (P_h u_0, W_0^+). \tag{6.27}$$

We recall from [30] that the solution $\widetilde{U}_h^{dG} \in \mathcal{S}^{\underline{r}}(\mathcal{M}, V_h)$ of the dG method (6.25) is uniquely defined and that the dG method (6.25) 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 $\widetilde{U}_h^{dG} \in \mathcal{P}_{r_n}(I_n, V_h)$, $1 \leq n \leq M$, by solving the variational problem

$$\int_{I_n} \left((\partial_t \widetilde{U}_h^{dG}, W) + \widetilde{a}(\widetilde{U}_h^{dG}, W) \right) dt + (\widetilde{U}_{n-1}^{dG+}, W_{n-1}^+) = (\widetilde{U}_{n-1}^{dG-}, W_{n-1}^+), \tag{6.28}$$

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

6.3.3 Convergence of the Fully Discrete Scheme

The solution operator of the parabolic problem (1.8) generates a holomorphic semi-group (see e.g. [30]). Therefore, the solution $u(t)$ of (1.8) is analytic with respect to t for all $t > 0$. However, due to the non-smoothness of the initial data, the solution may be singular at $t = 0$. The aim of this section is to 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 6.12. *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*

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

Definition 6.13. *A polynomial degree 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 $\lfloor \mu m \rfloor = \max\{q \in \mathbb{N}_0 : q \leq \mu m\}$.*

We have the following a priori error estimate on the fully discrete (perturbed) dG method.

Theorem 6.14. *Let $\rho \in (0, 2]$, and u be the solution of the parabolic problem (1.8) on $J \times \Omega = (0, T) \times (a, b)$, with initial data $u_0 \in \mathcal{V}_\theta$, for some $0 < \theta \leq 1$. Let the assumptions of Theorem 6.8 be satisfied. Moreover, let (6.18)–(6.23) hold.*

Then, there exist $\mu_0, m_0 > 0$ such that for all geometric partitions $\mathcal{M}_{M,\gamma}$ (cf. Definition 6.12) with $0 < \gamma < 1$ and $M \geq m_0 |\log h|$, and all polynomial degree vectors $\vec{r}$ on $\mathcal{M}_{M,\gamma}$ (cf. Definition 6.13) with slope $\mu > \mu_0$, the fully discrete solution $\widetilde{U}_h^{dG}$ to (1.8) obtained by (6.25) on the finite element space $\mathcal{S}^{\underline{r}}(\mathcal{M}_{M,\gamma}, V_h)$ satisfies the a priori error estimate

$$\|u(T) - \widetilde{U}_h^{dG}(T)\|_V \leq CT^{-\frac{p+1}{\widetilde{m}}} h^{p+1} \tag{6.29}$$

Here, h and p are the mesh size and the polynomial degree of the spatial discretization, respectively, and $C > 0$ is a constant independent of h.

6.4 Iterative Solution of the Linear Systems

The dG time stepping scheme (6.25) requires the solution of a nonsymmetric linear system in each time step. These systems can be solved approximately by incomplete GMRES iteration without loss

in convergence rates in the error estimate (6.29). Based on the norm equivalences (5.17), it can be shown as in [23, 26] that the overall complexity of this fully discrete dG time stepping is linear (up to some logarithmic terms). We indicate the argument and state the main result.

The dG time stepping scheme (6.28) requires a linear system of size $(r_m + 1)N^L$ to be solved in each time step $m = 1, 2, \ldots, M$. 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 $1 \leq m \leq M$, let $\{\phi_j = \sqrt{j + \frac{1}{2}} L_j\}_{j=0}^{r_m}$, 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 $\widetilde{U}_{h,m}^{dG}(x, t) = \widetilde{U}_h^{dG}|_{I_m}(x, t)$ and $W_m = W|_{I_m}$ in (6.28) as

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

the variational formulation (6.28) reads: find $(\widetilde{U}_{h,m,j}^{dG})_{j=0}^{r_m} \in (V_h)^{\underline{r}}$ such that there holds for all $(W_{m,i})_{i=0}^{r_m} \in (V_h)^{\underline{r}}$,

$$\sum_{i,j=0}^{r_m} C_{ij}(\widetilde{U}_{h,m,j}^{dG}, W_{m,i})_{L^2(\Omega)} + \frac{k_m}{2} \sum_{i=0}^{r_m} \widetilde{a}(\widetilde{U}_{h,m,j}^{dG}, W_{m,i}) = \sum_{i=0}^{r_m} f_{m,i}(W_{m,i}), \qquad (6.30)$$

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

$$C_{ij} = \sigma_{ij}\sqrt{\left(i + \frac{1}{2}\right)\left(j + \frac{1}{2}\right)}, \quad \sigma_{ij} = \begin{cases} (-1)^{i+j} & \text{if } j > i \\ 1 & \text{else} \end{cases}, \qquad (6.31)$$

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

From now on, for the sake of readability, we will drop the subscript m. Then, 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, (6.28) takes the matrix form

$$\mathbf{R}\underline{u} = \underline{f} \quad \text{with} \quad \mathbf{R} = \mathbf{C} \otimes \mathbf{M} + \frac{k}{2}\mathbf{I} \otimes \widetilde{\mathbf{A}}, \qquad (6.32)$$

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

In [26] it has been found that the system (6.32) of size $(r+1)N^L$ can be reduced to solving $r+1$ linear systems of size N^L. To this end, let $\mathbf{C} = \mathbf{Q}\mathbf{T}\mathbf{Q}^\top$ 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 $\lambda_1, \lambda_2, \ldots, \lambda_{r+1}$ of $\mathbf{C}$. Multiplying (6.32) by $\mathbf{Q}^\top \otimes \mathbf{I}$ from the left results in the linear system

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

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

$$\left(\lambda_{j+1}\mathbf{M} + \frac{k}{2}\widetilde{\mathbf{A}}\right)\underline{w}_j = \underline{s}_j \tag{6.33}$$

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

By (6.33), a dG-time step of order r amounts to solving $r + 1$ linear systems of the form

$$\mathbf{B} = \lambda\mathbf{M} + \frac{k}{2}\widetilde{\mathbf{A}}. \tag{6.34}$$

Here, λ is an eigenvalue of $\mathbf{C}$ in (6.31). Estimates on the eigenvalues of $\mathbf{C}$ have been established in [26].

For the preconditioning of (6.34), we define the matrix $\mathbf{S}$ and the scaled matrix $\widehat{\mathbf{B}} \in \mathbb{R}^{N^L} \times \mathbb{R}^{N^L}$ by

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

where $\mathbf{D}^{m(x)}$ is defined in (5.14).

The preconditioned linear equations corresponding to (6.33) are solved approximately with incomplete GMRES(m_0) iteration (restarted every $m_0 \geq 1$ iterations). There holds:

Theorem 6.15. *Let the assumptions of Theorem 6.14 hold. Then, choosing the number and order of time steps such that $M = r = \mathcal{O}(|\log h|)$ and in each time step $n_G = \mathcal{O}(|\log h|)^5$ GMRES iterations,*

$$\|u(T) - \widehat{U}_h^{dG}(T)\|_V \leq CT^{-\frac{p+1}{m}}h^{p+1}, \tag{6.35}$$

where $\widehat{U}_h^{dG}$ denotes the (perturbed) dG approximation of the exact solution u to (1.8) obtained by the incomplete GMRES(m_0) method.

Applying the matrix compression techniques from Section 5, the judicious combination of geometric mesh refinement and linear increase of polynomial degrees in the dG time-stepping scheme (Theorem 6.14), and an appropriate number of GMRES iterations (Theorem 6.15), results in an linear (up to some logarithmic terms) overall complexity of the fully discrete scheme (6.25) for the solution of the parabolic problem (1.8).

Theorem 6.16. *Under the assumptions of Theorems 6.14 and 6.15, the fully discrete scheme (6.25) with $n_G = \mathcal{O}(|\log h|)^5$ GMRES(m_0) iterations per time step yields $\widehat{U}_h^{dG}(T)$ in at most $\mathcal{O}(N(\log N)^8)$ operations, where $N = N^L = \dim V_h = \mathcal{O}(h^{-1})$ is the number of spatial degrees of freedom.*

References

[1] R.F. Bass, Uniqueness in law for pure jump Markov processes. Probab. Theor. Rel. Fields **79** (1988), 271-287.

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

[3] S. Beuchler, R. Schneider and C. Schwab, Multiresolution weighted norm equivalences and applications, Numer. Math. **98** (2004), 67-97.

[4] F. Black and M. Scholes, The Pricing of Options and Corporate Liabilities, Journal of Political Economy **81** (1973), 637–654

[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] A. Cohen, Numerical Analysis of Wavelet Methods, Studies in Mathematics and its Application **32**, North Holland-Elsevier (2003).

[7] R. Cont and P. Tankov, Financial modelling with jump processes, Chapman and Hall/CRC Press, 2004

[8] Ph. Courrège, Sur la forme intégro-differentielle des opérateurs de C_K^∞ dans C satisfaisant du principe du maximum, Sém. Théorie du Potentiel **38**, 1965/66.

[9] W. Dahmen and R. Schneider, Wavelets with complementary boundary conditions, Function spaces on the cube, *Result. Math.* **34** (1998), 255–293.

[10] F. Delbaen and W. Schachermayer, The Variance-Optimal Martingale Measure for Continuous Processes, Bernoulli **2** (1996) 81–105

[11] H. Harbrecht and R. Stevenson: Wavelets with patchwise cancellation properties, Preprint No. 1311, Department of Mathematics, University of Utrecht (The Netherlands), 2004.

[12] W. Hoh, Pseudodifferential Operators with negative definite symbols of variable order, Rev. Math. Iberoamericana **16** (2000) 219-241.

[13] L. Hörmander: The Analysis of Linear Partial Differential Operators I: Distributions and Fourier Analysis, 2nd Ed., Springer Verlag, 1995

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

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

[16] K. Kikuchi and A. Negoro, On Markov Process generated by pseudodifferential operator of variable order, Osaka J. Math. **34**(1997) 319-335.

[17] H. Kumano-go: Pseudo-Differential Operators, The MIT Press, Cambridge, Massachusetts 1981.

[18] H.G. Leopold, On function spaces of variable order of differentiation, Forum Math. **3**(1991) 1-21.

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

[20] A.-M. Matache, P.A. Nitsche and C. Schwab, Wavelet Galerkin Pricing of American contracts on Lévy driven assets, (in review).

[21] A.-M. Matache and T. von Petersdorff and C. Schwab, *Fast Deterministic Pricing of Options on Lévy Driven Assets*, Mathematical Modelling and Numerical Analysis **38** (2004) 37–71

[22] A.-M. Matache and C. Schwab and T. P. Wihler, *Fast Numerical Solution of Parabolic Integro-Differential Equations with Applications in Finance* , Tech. Report No. 1954, Institute for Mathematics and its Applications (IMA), University of Minnesota, Jan. 2004, to appear in SIAM J. Sci. Computing (2005).

[23] A.-M. Matache and C. Schwab and T. P. Wihler, *Linear Complexity solution of parabolic integrodifferential equations* Tech. Report No. 1990, Institute for Mathematics and its Applications (IMA), University of Minnesota, Sep. 2004 (in review)
http://www.ima.umn.edu/preprints/sep2004/sep2004.html

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

[25] T. von Petersdorff and C. Schwab, *Wavelet Discretizations of Parabolic Integro-Differential Equations*, SIAM J. Numer. Anal. **41** (2003), 159–180

[26] T. von Petersdorff and C. Schwab, *Numerical Solution of Parabolic Equations in High Dimensions*, Mathematical Modelling and Numerical Analysis, **38** (2004) 93-128.

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

[28] R. Schilling On the existence of Feller processes with a given generator, Preprint (2004)
http://www.mathematik.uni-marburg.de/ schilling/research

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

[30] D. Schötzau and C. Schwab, hp-discontinuous Galerkin time-stepping for parabolic problems, C.R.Acad. Sci. Paris **333** (2001) 1121–1126.

[31] M.A. Shubin, Pseudodifferential Operators and Spectral Theory, Springer Verlag, Berlin, Heidelberg, New York, 1985.

[32] E.M. Stein, Harmonic Analysis, Princeton Univ. Press, Princeton, N.J. 2002.