A Flexible Galerkin Scheme for Option Pricing in Lévy Models*

Maximilian Gaß[†] and Kathrin Glau[‡]

Abstract. One popular approach to option pricing in Lévy models is through solving the related partial integro differential equation (PIDE). For the numerical solution of such equations powerful Galerkin methods have been put forward e.g. by Hilber, Reichmann, Schwab, Winter (2013). As in practice large classes of models are maintained simultaneously, flexibility in the driving Lévy model is crucial for the implementation of these powerful tools. In this article we provide a tool that enables the implementation of finite element Galerkin methods *flexibly in the model*. To this end we exploit the Fourier representation of the infinitesimal generator, i.e. the related symbol, which is explicitly available for the most relevant Lévy models. Empirical studies for the Merton, NIG and CGMY model confirm the numerical feasibility of the tool.

Key words. Lévy processes, partial integro differential equations, pseudo-differential operators, symbol, option pricing, Galerkin approach, finite element method

AMS subject classifications. 91G80, 60G51, 35S10, 65M60

1. Introduction.

In computational finance, methods to solve partial differential equations come into play, when both run-time and accuracy matter. In contrast to Monte Carlo simulation for example, run-time is very appealing and a deterministic and conservative error analysis is established and well understood. In addition, compared to Fourier methods, the possibility to capture path-dependent features like early exercise and barriers is naturally built in. Within these appealing features lies the capacity to attract interest from academia and satisfy the needs of the financial industry alike.

In academia, a series of publications by Cont and Voltchkova in 2005 [10], Hilber, Reich, Schwab and Winter in 2009 [17], Jackson, Jaimungal and Surkov in 2012 [21] Salmi, Toivanen and Sydow in 2014 [24], Itkin in 2015 [19], Glau in 2016 [16], and the monograph of Hilber, Reichmann, Schwab and Winter in 2013 [18] have opened the theory to include even more sophisticated models of Lévy type, resulting in Partial *Integro* Differential Equations (PIDEs). The theoretical results have been validated by sophisticated numerical studies. In this context, Schwab and his working group in particular have taken the lead and unveiled the potential of PIDE theory in high generality and for practical purposes in the financial industry. Combining state of the art compression techniques with a wavelet finite element setup has resulted in a numerical framework for option pricing in advanced and multivariate jump models, which thereby moved academic boundaries.

Two standard methods are available for solving PIDEs, that is the finite difference approach and the finite element method (FEM). More recently, also radial basis methods have been pushed forward to solve pricing PIDEs. For all of these concepts implementations for a variety of models and option types have already been developed: Finite difference schemes

*We thank Ernst Eberlein for valuable feedback to this manuscript.

[†]Department for Mathematics at Technical University of Munich (maximilian.gass@mytum.de).

[‡]Department for Mathematics at Technical University of Munich and School of Mathematical Sciences, Queen Mary University of London (kathrin.glau@tum.de and k.glau@qmul.ac.uk).

solving PIDEs for pricing European and barrier options with an implementation for Merton and Variance Gamma are provided by Cont and Voltchkova in 2015 [10], [9]. The method has been further developed in different directions, we mention one example, by Itkin and Carr in 2012 [20], who exploit a special representation of the equation tailored to jump diffusions with jump intensity of tempered stable type. Wavelet-Galerkin methods for PIDEs related to a class generalizing tempered stable Lévy processes are derived by Matache, Nitsche and Schwab in 2005 [23] for American options and e.g. by Marazzani, Reichmann and Schwab in 2012 [22] for a high-dimensional extension. A Fourier time stepping scheme combining PIDE with fast Fourier transform methods has been proposed in Jackson, Jaimungal and Surkov in 2012 [21]. Radial basis approaches for the Merton and Kou model, American and European options are provided by Chan and Hubbert in 2014 [7] and further developed for CGMY models by Brummelhuis and Chan in 2014 [4].

In the financial industry an awareness of the full potential of these tools is yet to be developed. Advocating the advancement of numerical methods one must acknowledge what practice cherishes most. Due to model uncertainty and behavioral characteristics of different portfolios, financial institutions need to deal with a number of different pricing models in parallel. Or, in the words of Föllmer in [13]: "In any case, the signal towards the practitioners of risk management is clear: Do not commit yourself to a single model, remain flexible, vary the models in accordance with the problem at hand, always keeping in mind the worst case scenario."[1] Desirable features that the numerical environment must offer include

(1) a degree of accuracy reaching levels relevant to practical applications that can be measured and controlled by a theoretical error analysis,
(2) fast run times,
(3) low and feasible implementational and maintenance cost,
(4) a flexibility of the toolbox towards different options and models.

An implementation that is flexible in the driving model as well as in the option type first of all requires a problem formulation covering the collectivity of envisaged models and options. In view of feature (1), a unified approach to the error analysis of the resulting schemes is of equal importance. Galerkin methods, accruing from the Hilbert space formulation of the Kolmogorov equation, seem to be predestined to deliver the adequate level of abstraction for this task. It is precisely this abstract level that makes Galerkin methods flexible in the option type and the dimension of the underlying driving process. Consequently, even though Galerkin methods seem to be more involved at first glance in comparison to finite difference schemes, they still promise to lead to a lucid code that is easy to maintain and to extend, and that allows clear an extensive convergence and error analysis. This is of great importance for implementation and controlling methodological risk in finance. Moreover, Galerkin methods allow for efficient compression techniques such as wavelet-compressions, see [18], and reduced order modeling, see e.g. [8], [5]. We therefore consider the finite element, or more general Galerkin methods, worth exploring further for financial applications.

Unfortunately, although flexibility towards models goes well with the abstract formulation, the finite element method faces numerical challenges when implementing Lévy model based pricing tools. More precisely, the Lévy operator that determines the stiffness matrix is of

[1] Translated from German.

integro differential type. Firstly, the resulting matrix is densely populated and in general not symmetric. Secondly, and even more severe, the matrix entries typically are not explicitly available. Instead, they require the evaluation of double integral terms possibly involving a numerically inaccessible Lévy measure. In these cases, a thorough analysis of the respective integrals may lead to approximation schemes deriving the stiffness matrix entries with the required precision. Pursuing this way, however, most likely results in a model specific scheme, contradicting requirement (4).

In this paper we develop a new methodology for option pricing in Lévy models using finite elements which is flexible in the choice of model. We address this goal by expressing the operator in the Fourier space. This means accessing the model specific information via the symbol, and we call the resulting tool the *symbol method*. In contrast to the operator, the symbol is explicitly available for a variety of models and is thus numerically superior. Further advantages will be highlighted in subsequent sections. It is worth mentioning a conceptual relation of this new approach to the Fourier time stepping scheme of [21]. Both methods result in PIDE discretizations that rely on the symbol of the driving Lévy process. While we propose to express the bilinear form in the Galerkin representation via the symbol, the methods of [21] are based on applying the Fourier transform to the pricing PIDEs and is not related to Galerkin approximations.

Section 2 introduces the theoretical framework for our PIDEs of interest and their weak formulation. The next section describes the solution scheme, that is the Galerkin approximation in space. We investigate the scheme with regard to the numerical challenges arising during its implementation. Section 4 introduces the symbol method itself. All components of the FEM solver are expressed in Fourier space. The subsequent numerical evaluation of the stiffness matrix entries is supported by an elementary approximation result. Several examples of symbols for well-known Lévy models confirm the wide applicability of the method and its numerical advantages. The actual implementation of the symbol method poses new challenges. We propose two different ways to tackle these challenges and to obtain a convergent and flexible scheme. As first approach, we propose to mollify the classic hat functions in Section 5. We analyse the error in detail and under standard conditions, obtain the same rate of convergence as for the case without mollification. Section 6 introduces an alternative approach by choosing splines as basis functions. The numerical studies in Section 7 confirm theoretically prescribed rates of convergence and validate the claim of numerical feasibility.

2. Kolmogorov equations for option pricing in Lévy models.

We first introduce the underlying stochastic processes, the Kolmogorov equation, its weak formulation as well as the solution spaces of our choice.

2.1. Lévy processes.

Let a stochastic basis $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{0\le t\le T}, P)$ be given and let L be an $\mathbb{R}^d$-valued *Lévy process* with characteristics $(b, \sigma, F; h)$, i.e. for fixed $t \ge 0$ its characteristic function is given by

$$E\,\mathrm{e}^{i\langle \xi, L_t\rangle} = \mathrm{e}^{-tA(-\xi)} \quad \text{for every } \xi \in \mathbb{R}^d, \tag{1}$$

where the *symbol of the process* is defined as

$$A(\xi) := \frac{1}{2}\langle \xi, \sigma\xi\rangle + i\langle \xi, b\rangle - \int_{\mathbb{R}^d}\left(\mathrm{e}^{-i\langle \xi, y\rangle} - 1 + i\langle \xi, h(y)\rangle\right) F(\mathrm{d}y). \tag{2}$$

Here, σ is a symmetric, positive semi-definite $d \times d$-matrix, $b \in \mathbb{R}^d$, and F is a Lévy measure, i.e. a positive Borel measure on $\mathbb{R}^d$ with $F(\{0\}) = 0$ and $\int_{\mathbb{R}^d} (|x|^2 \wedge 1) F(\mathrm{d}x) < \infty$. Moreover, h is a truncation function i.e. $h : \mathbb{R}^d \to \mathbb{R}^d$ such that $h(x) = x$ in a neighborhood of 0 and $\int_{\{|x|>1\}} h_j(x) F(\mathrm{d}x) < \infty$, where h_j denotes the j-th component of the truncation function h for all $j = 1, \ldots, d$. The *Kolmogorov operator of a Lévy process* L with characteristics $(b, \sigma, F; h)$ is given by

$$
\begin{aligned}
\mathcal{A}\varphi(x) := &- \frac{1}{2} \sum_{j,k=1}^{d} \sigma^{j,k} \frac{\partial^2 \varphi}{\partial x_j \partial x_k}(x) - \sum_{j=1}^{d} b^j \frac{\partial \varphi}{\partial x_j}(x) \\
&- \int_{\mathbb{R}^d} \Big(\varphi(x+y) - \varphi(x) - \sum_{j=1}^{d} \frac{\partial \varphi}{\partial x_j}(x)\, h_j(y) \Big) F(\mathrm{d}y)
\end{aligned} \tag{3}
$$

for every $\varphi \in C_0^\infty(\mathbb{R}^d)$.

2.2. Kolmogorov equation in variational form. Key for the variational formulation of the Kolmogorov equation

$$
\partial_t u + \mathcal{A}u = f \tag{4}
$$

$$
u(0) = g \tag{5}
$$

is the definition of the bilinear form

$$
a(\varphi, \psi) := \int_{\mathbb{R}^d} (\mathcal{A}\varphi)(x) \psi(x)\, \mathrm{d}x \qquad \text{for all } \varphi, \psi \in C_0^\infty(\mathbb{R}^d). \tag{6}
$$

It is one of the major advantages of variational formulations of evolution equations that solution spaces of low regularity, as compared to the space C^2 for example, are incorporated in an elegant way. Departing from the space $C_0^\infty(\mathbb{R}^d)$ of smooth functions with compact support, we can select from a large variety of function spaces V that are characterized by the following assumption.

(A1) V and H are Hilbert spaces such that $C_0^\infty(\mathbb{R}^d)$ is dense in V and there exists a continuous embedding from V into H.

Existence and uniqueness of a variational solution critically hinges on the following two properties of the bilinear form:

(A2) *Continuity*: There exists a constant $C > 0$ such that

$$
\big| a(\varphi, \psi) \big| \le C \|\varphi\|_V \|\psi\|_V \qquad \text{for all } \varphi, \psi \in C_0^\infty(\mathbb{R}^d).
$$

(A3) *Gårding inequality*: There exists constants $G > 0$ and $G' \ge 0$ such that

$$
a(\varphi, \varphi) \ge G \|\varphi\|_V^2 - G' \|\varphi\|_H^2 \qquad \text{for all } \varphi \in C_0^\infty(\mathbb{R}^d).
$$

We observe that due to (A1) and (A2), the bilinear form a possesses a unique continuous bilinear extension $a : V \times V$ that is continuous, i.e. for a constant $C > 0$ we have $\big| a(\varphi, \psi) \big| \le C \|\varphi\|_V \|\psi\|_V$ for all $\varphi, \psi \in V$. Also (A3) holds for all $\varphi \in V$.

As V is separable, this is also true for H and one can find a continuous embedding from H to the dual space V^* of V, i.e. (V, H, V^*) is a Gelfand triplet. We then denote by $L^2\big(0,T;H\big)$ the space of all functions $u : [0,T] \to H$ such that for every $h \in H$ the map $s \mapsto \langle u(s), h\rangle$ is Borel measurable and $\int_0^T \|u(t)\|_H^2 \,\mathrm{d}t < \infty$. Moreover, we denote by $\partial_t u$ the derivative of u with respect to time in the distributional sense. For a detailed definition, which relies on the Bochner integral, we refer to Section 24.2 in [29]. The Sobolev space

$$W^1(0,T;V,H) := \Big\{ u \in L^2\big(0,T;V\big) \,\Big|\, \partial_t u \in L^2\big(0,T;V^*\big)\Big\} \tag{7}$$

will play the role of the solution space in the variational formulation of the Kolmogorov equation (4), (5).

Definition 1. *Let $f \in L^2\big(0,T;V^*\big)$ and $g \in H$. Then $u \in W^1(0,T;V,H)$ is a* variational solution *of Kolmogorov equation (4), if for almost every $t \in (0,T)$,*

$$\langle \partial_t u(t), v\rangle_H + a(u(t), v) = \langle f(t)|v\rangle_{V^*\times V} \quad \text{for all } v \in V \tag{8}$$

and $u(t)$ converges to g for $t \downarrow 0$ in the norm of H.

Remark 2. *Assumptions (A1)–(A3) guarantee the existence and uniqueness of a variational solution $u \in W^1(0,T;V,H)$ of (8), see for instance Theorem 23.A in [30].*

2.3. Solution spaces. Expression (6) is based on the L^2-scalar product and is appropriate for variational equations in Sobolev spaces. Then, typically $H = L^2$. For Kolmogorov equations for option prices the initial condition g in (5) plays the role of the (logarithmically transformed) payoff function of the option. For a call option with strike K it is of the form $x \mapsto (S_0\,\mathrm{e}^x - K)^+$, for a digital up and out option it is given by $x \mapsto \mathbb{1}_{\mathrm{e}^x<b}$ for some $b \in \mathbb{R}$. We thus have to observe that the initial condition g is not square integrable for most of the typical cases of interest. Therefore, we base our analysis more generally on exponentially weighted L^2 spaces: For $\eta \in \mathbb{R}^d$ let

$$L^2_\eta(\mathbb{R}^d) := \big\{ u \in L^1_{loc}(\mathbb{R}^d) \,|\, u\,\mathrm{e}^{\langle \eta, \cdot\rangle} \in L^2(\mathbb{R}^d)\big\}, \quad \|u\|_{L^2_\eta} := \bigg(\int_{\mathbb{R}^d} \big|u(x)\big|^2\,\mathrm{e}^{2\langle \eta, x\rangle}\,\mathrm{d}x\bigg)^{1/2}$$

and

$$a(\varphi,\psi) := \langle \mathcal{A}\varphi, \psi\rangle_{L^2_\eta} = \int_{\mathbb{R}^d} (\mathcal{A}\varphi)(x)\psi(x)\,\mathrm{e}^{2\langle\eta,x\rangle}\,\mathrm{d}x \quad \text{for all } \varphi,\psi \in C_0^\infty(\mathbb{R}^d). \tag{9}$$

We notice that all assertions of the precedent section, concerning assumptions (A1)–(A3) and variational equations hold for bilinear form a defined by (9) instead of a from (6) as well.

As solution spaces V we consider weighted Sobolev-Slobodeckii spaces. These have proven to apply to a large set of option types and models. We refer to [12] and [16], where particularly Feynman-Kac type formulas have been derived linking European and path-dependent options to weak solutions of Kolmogorov equations in Sobolev-Slobodeckii spaces.

To introduce the spaces, we denote by $C_0^\infty(\mathbb{R}^d)$ the set of smooth real-valued functions with compact support in $\mathbb{R}^d$ and let

$$\widehat{\varphi}(\xi) = \mathcal{F}(\varphi)(\xi) := \int_{\mathbb{R}^d} \mathrm{e}^{i\langle\xi,x\rangle}\,\varphi(x)\,\mathrm{d}x \tag{10}$$

be the Fourier transform of $\varphi \in C_0^\infty(\mathbb{R}^d)$ and $\mathcal{F}^{-1}$ be its inverse. We define the *exponentially weighted Sobolev-Slobodeckii space* $H_\eta^\alpha(\mathbb{R}^d)$ with index $\alpha \geq 0$ and weight $\eta \in \mathbb{R}^d$ as the completion of $C_0^\infty(\mathbb{R}^d)$ with respect to the norm $\|\cdot\|_{H_\eta^\alpha}$ given by

$$\|\varphi\|_{H_\eta^\alpha}^2 := \int_{\mathbb{R}^d} \big(1+|\xi|\big)^{2\alpha} \big|\mathcal{F}(\varphi)(\xi - i\eta)\big|^2 \,\mathrm{d}\xi. \tag{11}$$

By construction $H_\eta^\alpha(\mathbb{R}^d)$ is a separable Hilbert space and we denote its dual space by $\big(H_\eta^\alpha(\mathbb{R}^d)\big)^*$.

3. Implementational Challenges.

Based on this theoretical introduction we are now in the position to focus on its implementation and related numerical questions.

3.1. Abstract Galerkin approximation in space.

For a countable Riesz basis $\{\varphi_1, \varphi_2, \ldots\}$ of V we define

$$V_N := \operatorname{span}\{\varphi_1, \ldots, \varphi_N\} \qquad \text{for all } N \in \mathbb{N}.$$

Since V is dense in H, we may further choose g_N in V_N such that $g_N \to u(0)$ in H. For each fixed $N \in \mathbb{N}$ the semidiscrete problem is defined by restricting (8) to the finite dimensional space: *Find a function* $u_N \in W^1(0,T;V_N;H\cap V_N)$ *that satisfies for all* $\chi \in C_0^\infty(0,T)$ *and* $\varphi \in V_N$,

$$\begin{aligned} -\int_0^T \langle u_N(t), \varphi\rangle_{L^2}\, \dot\chi(t)\,\mathrm{d}t + \int_0^T a\big(u_N(t),\varphi\big)\,\chi(t)\,\mathrm{d}t &= \int_0^T \langle f(t)|\varphi\rangle_{V^*\times V}\,\chi(t)\,\mathrm{d}t \\ u_N(0) &= g_N. \end{aligned} \tag{12}$$

As a result of the elegant Hilbert space formulation, the semidiscrete problem (12) is uniquely solvable and the convergence of the sequence u_N to u is guaranteed, see Theorem 23.A and Remark 23.25 in [30].

The major advantage of equation (12) in regard to implementation is that it suffices to insert the basis functions as test functions. Thus, denoting $g_N = \sum_{j=1}^N \alpha_j \varphi_j$ and $u_N(t) := \sum_{j=1}^N U_j(t)\varphi_j$ we arrive at

$$\begin{aligned} \sum_{l=1}^N \dot U_l(t)\langle \varphi_l, \varphi_k\rangle_{L^2} + \sum_{l=1}^N U_l(t) a\big(\varphi_l,\varphi_k\big) &= \langle f(t)|\varphi_k\rangle_{V^*\times V} \\ U_j(0) &= \alpha_j \quad \text{for all } j = 1,\ldots,N. \end{aligned}$$

Written in matrix form the problem is to find $U : [0,T] \to \mathbb{R}^N$ such that

$$\mathbf{M}\dot U(t) + \mathbf{A}U(t) = \mathbf{F}(t) \tag{13}$$

$$U(0) = \alpha, \tag{14}$$

where the *right hand side (vector)* $\mathbf{F}$ is given by $\mathbf{F} = (F_1, \ldots, F_N)^\top$ with $F_j(t) = \langle f(t)|\varphi_j\rangle_{V^*\times V}$ for $j = 1,\ldots,N$, $\alpha = (\alpha_1,\ldots,\alpha_N)^\top$, and the *mass matrix* $\mathbf{M}$ and *stiffness matrix* $\mathbf{A}$ are given by

$$M_{kl} = \langle \varphi_l, \varphi_k\rangle_{L^2}, \qquad A_{kl} = a\big(\varphi_l,\varphi_k\big) \qquad \text{for all } k,l = 1,\ldots,N. \tag{15}$$

3.2. Fully discrete scheme. As fully discrete scheme, we approximate (13) with a θ scheme in time, namely

$$\mathbf{M}\frac{U^{m+1}-U^m}{\Delta t}+\mathbf{A}U^{m+\theta}(t)=\mathbf{F}^{m+\theta}(t) \tag{16}$$

$$U(0)=\alpha, \tag{17}$$

where $U^{m+\theta}=\theta U^{m+1}+(1-\theta)U^m$, $F^{m+\theta}$ accordingly, and $\theta\in[0,1]$.

3.3. Flexible implementation for different driving Lévy processes. We inspect equations (13) and (14) in regard to flexibility towards different options as well as models. All ingredients depend on the choice of the basis. While M is independent of the specific problem at hand, F and α represent the input data and therefore may vary for different option types. The stiffness matrix A carries the information of the driving process. So in order to obtain flexibility towards model types, we need a *generic way to compute the entries of the stiffness matrix.* For smooth basis functions with compact support and solution spaces without weighting, i.e. $\eta=0$, according to (3) and (6), the stiffness matrix entries are given by

$$\begin{aligned}a\big(\varphi_l,\varphi_k\big)=&\sum_{i,j=1}^d\frac{\sigma^{i,j}}{2}\int_{\mathbb{R}^d}\frac{\partial}{\partial x_j}\varphi_l(x)\frac{\partial}{\partial x_i}\varphi_k(x)\,\mathrm{d}x-\sum_{i=1}^d b^i\int_{\mathbb{R}^d}\frac{\partial}{\partial x_i}\varphi_l(x)\varphi_k(x)\,\mathrm{d}x\\ &-\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\Big(\varphi_l(x+y)-\varphi_l(x)-\sum_{i=1}^d\frac{\partial}{\partial x_i}\varphi_l(x)\,h_i(y)\Big)F(\mathrm{d}y)\varphi_k(x)\,\mathrm{d}x.\end{aligned} \tag{18}$$

Typical basis functions are not smooth. Therefore it is not a priori clear if the integral representation (18) extends to the usual basis functions. Observe that an extension of this representation requires some care: For a large and important class of pure jump Lévy processes, the solution spaces are Sobolev-Slobodeckii spaces of fractional order, i.e. H^α with some $0<\alpha<1$. For functions in H^α with $\alpha<1$, however, the first order weak derivative in (18) is not defined and therewith this integral representation of the bilinear form is not well-defined. Understanding that the basis functions are usually in H^1, we derive the validity of the representation under appropriate assumptions that also include the more challenging case of solution spaces with fractional order derivatives.

Lemma 3. *Let $d=1$. Let a be defined by (9). Assume (A1)–(A3) for a, V and H and denote by $a:V\times V$ its unique bilinear continuous extension. If $H^1_\eta(\mathbb{R})\subset V$, we have for every $\varphi,\psi\in H^1_\eta(\mathbb{R})$,*

$$\begin{aligned}a(\varphi,\psi)=&\frac{\sigma}{2}\int_{\mathbb{R}}\varphi'(x)\psi'(x)\,\mathrm{e}^{2\eta x}\,\mathrm{d}x-b(\eta,\sigma,F)\int_{\mathbb{R}}\varphi'(x)\psi(x)\,\mathrm{e}^{2\eta x}\,\mathrm{d}x\\ &-\int_{\mathbb{R}}\int_{|y|<1}\int_0^y\int_0^z\varphi'(x+v)\,\mathrm{d}v\,\mathrm{d}zF(\mathrm{d}y)\big(\psi'(x)+2\eta\psi(x)\big)\,\mathrm{e}^{2\langle\eta,x\rangle}\,\mathrm{d}x\\ &-\int_{\mathbb{R}}\int_{|y|>1}\big(\varphi(x+y)-\varphi(x)\big)F(\mathrm{d}y)\psi(x)\,\mathrm{e}^{2\langle\eta,x\rangle}\,\mathrm{d}x\end{aligned} \tag{19}$$

with

$$b(\eta,\sigma,F) = b - 2\sigma\eta + \int_{|y|<1} \big(y - h(y)\big) F(\mathrm{d}y) - \int_{|y|>1} h(y) F(\mathrm{d}y).$$

The proof of the Lemma is provided in Section A.1.

Inspecting the expression for the bilinear form, we encounter several numerical challenges due to the integral part—stemming from the jumps of the process:

1. The appealing tridiagonal structure of the stiffness matrix for classic hat functions related to the Black-Scholes equation does not extend to the general Lévy setting. Instead, the stiffness matrix is densely populated. Pleasantly, it is still a Toeplitz matrix.
2. For some choices of Lévy measures and bases the stiffness matrix entries may be derived in closed form. This is for instance the case for the Merton model and piecewise linear basis functions when $\eta = 0$. Following Section 10.6.2 in [18], the stiffness matrix entries may be derived in semi-closed form expressions for a further group of jump intensities including tempered stable, CGMY and KoBoL processes and the choice of piecewise linear basis functions. In general, however, closed form expressions for the stiffness matrix entries, when arbitrary models and basis functions are considered, are not available.

An implementation that is flexible in the driving Lévy process therefore has to rely on numerical approximations of the entries of the stiffness matrix. These approximations inevitably affect the accuracy of the solution to the scheme (13)–(14). The following question arises: *How accurate does the integration routine have to be chosen in order to meet a desired accuracy of the solution V?*

In order to gain a first practical insight in the magnitude of the error resulting from an inaccuracy in the stiffness matrix entries, consider Section 3.4.2 in [14]. The numerical investigations presented therein reveal that an impressively high precision of the computation of the entries of the stiffness matrix is required.

4. Fourier approach to the Kolmogorov equation. In regard to the high accuracy the approximation of the stiffness matrix entries needs to achieve, we would like to avoid numerical evaluations of the stiffness matrix entries on the basis of representation (**??**). Seeking for alternative representations, let us point out that the symbol A of the Lévy process is always available. Even more, it is an explicit function of the parameters of the process and thus can be seen as the modelling quantity of the process as the Examples 9–12 show below. We therefore take a Fourier perspective on the variational formulation of the Kolmogorov equation. This is especially promising since the Kolmogorov operator $\mathcal{A}$ of a Lévy process is a pseudo differential operator with symbol A,

$$\mathcal{A}\varphi = \mathcal{F}^{-1}(A\mathcal{F}(\varphi)) \qquad \text{for all } \varphi \in C_0^\infty(\mathbb{R}^d), \tag{20}$$

as elementary manipulations show. Now Parseval's identity yields

$$a(\varphi,\psi) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \mathcal{F}(\mathcal{A}\varphi)(\xi)\overline{\mathcal{F}(\psi)(\xi)}\,\mathrm{d}\xi \tag{21}$$

for all $\varphi, \psi \in C_0^\infty(\mathbb{R}^d)$, respectively,

$$a(\varphi,\psi) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} A(\xi)\mathcal{F}(\varphi)(\xi)\overline{\mathcal{F}(\psi)(\xi)}\,\mathrm{d}\xi. \tag{22}$$

This well-known identity has already proved to be highly beneficial for the analysis of the variational solutions of the Komogorov equations, compare e.g. [18], [15] and [16]. Let us point out the transition from the operator to the symbol from (21) to (22) in the bilinear form and recall its role for the derivation of the stiffness matrix in (15). The resulting alternative representation is key for the flexibility of our numerical approach. Exploiting the symbol will facilitate the numerical implementation considerably.

Lemma 4 (Continuous extension of bilinear forms). *Let A be the symbol of a Lévy process given by the characteristic triplet (b,σ,F). Denote by $\mathcal{A}: C_0^\infty(\mathbb{R}^d,\mathbb{C}) \to C^\infty(\mathbb{R}^d,\mathbb{C})$ the pseudodifferential operator associated with symbol A. Furthermore, denote by $a: C_0^\infty \times C_0^\infty \to \mathbb{C}$ the bilinear form associated with the operator $\mathcal{A}$. Let $\eta \in \mathbb{R}^d$. If*

i) the exponential moment condition

$$\int_{|x|>1} e^{-\langle \eta', x\rangle} F(\mathrm{d}x) < \infty \tag{23}$$

holds for all $\eta' \in \operatorname{sgn}(\eta^1)[0,|\eta^1|]\times\cdots\times\operatorname{sgn}(\eta^d)[0,|\eta^d|]$ and

ii) there exists a constant $C_1>0$ with

$$|A(z)| \le C_1(1+\|z\|)^\alpha \tag{24}$$

for all $z \in U_{-\eta} := U_{-\eta^1}\times\cdots\times U_{-\eta^d}$ with $U_{-\eta^j} = \mathbb{R} - i\operatorname{sgn}(\eta^j)[0,|\eta^j|)$,

then $a(\cdot,\cdot)$ possesses a unique linear extension $a: H_\eta^{\alpha/2}\times H_\eta^{\alpha/2} \to \mathbb{R}$ that can be written as

$$a(\varphi,\psi) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} A(\xi - i\eta)\widehat{\varphi}(\xi - i\eta)\overline{\widehat{\psi}(\xi - i\eta)}\,\mathrm{d}\xi \tag{25}$$

for all $\varphi,\psi \in H_\eta^{\alpha/2}(\mathbb{R}^d)$.

Proof. The proof can be found in [11] using Theorem 4.1 therein and Parseval's identity. ∎

In order to gain first insight in the convergence analysis, we fix a level N in the Galerkin scheme and derive conditions for the convergence of the sequence of weak solutions that we obtain by approximating the stiffness matrix entries. In the implementation in Section 7 below we will also approximately compute the right hand side F of the equation. We therefore more generally consider sequences of stiffness matrices, right hand sides and initial conditions.

As usual, we denote for a given bilinear form $a: V\times V \to \mathbb{R}$ the associated operator $\mathcal{A}: V \to V^*$ defined by $\mathcal{A}(u)(v) := a(u,v)$ for all $u,v\in V$.

Lemma 5. *Let V, H and $a: V\times V\to\mathbb{R}$ satisfy (A1)–(A3). Let $X := \operatorname{span}\{\varphi_1,\dots,\varphi_N\} \subset V$ and for each $n\in\mathbb{N}$ let*

(An1) $f_n, f \in L^2(0,T;H)$ with $f_n \to f$ in $L^2\big(0,T;X^*\big)$,

(An2) $g_n, g \in H$ with $g_n \to g$ in H,

(An3) $a_n : V \times V \to \mathbb{R}$ *be a bilinear form such that for all* $l, k \le N$,

$$\big|(a_n - a)(\varphi_l, \varphi_k)\big| \to 0. \tag{26}$$

Then the sequence of unique weak solutions $u_n \in W^1(0,T;X,H)$ *of*

$$\dot{u}_n + \mathcal{A}_n u_n = f_n, \quad u_n(0) = g_n \tag{27}$$

converges strongly[2] *in* $L^2(0,T;X) \cap C(0,T;H)$ *to the unique weak solution* $u \in W^1(0,T;X,H)$ *of*

$$\dot{u} + \mathcal{A}u = f, \quad u(0) = g. \tag{28}$$

The proof is provided in Section A.2.

Next we introduce our approach to approximate the stiffness matrix entries.

4.1. The symbol method. The key component of a Galerkin FEM solver is the model dependent stiffness matrix $\mathbf{A} \in \mathbb{R}^{N\times N}$. Using expression (18) of Section 3.3 above, the entries of that matrix can be derived. The way the Lévy measure F enters that expression, however, renders the numerical derivation of the matrix rather cumbersome. Additionally, the empirical accuracy study of Section 3.4.2 in [14] emphasizes that utmost care must be taken when the stiffness matrix entries are numerically derived. Consequently, in this section we approach the calculation of the FEM solver components differently. The Fourier approach indicated by Lemma 4 will allow us to access the model information required for the stiffness matrix and all other FEM solver components via the symbol that is associated with the operator. In stark contrast to the operator, the symbol of a Lévy model is numerically accessible in a unified way for a large set of underlying models and we will present several examples highlighting this feature.

Let us state the core lemma of this section. Here we concentrate on basis functions obeying a simple nodal translation property, which is in particular satisfied for classical piecewise polynomial basis functions.

Lemma 6 (Symbol method for bilinear forms). *Let the assumptions of Lemma 4 be satisfied with* $\eta = 0$. *Assume further for* $N \in \mathbb{N}$ *a set of functions* $\varphi_0, \varphi_1, \ldots, \varphi_N \in H_0^{\alpha/2}(\mathbb{R})$ *and nodes* $x_1, \ldots, x_N \in \mathbb{R}$, *such that for all* $j = 1, \ldots, N$

$$\varphi_j(x) = \varphi_0(x - x_j) \qquad \forall x \in \mathbb{R}.$$

Then we have

$$a(\varphi_l, \varphi_k) = \frac{1}{2\pi} \int_{\mathbb{R}} A(\xi) e^{i\xi(x_l - x_k)} \, |\widehat{\varphi_0}(\xi)|^2 \, \mathrm{d}\xi. \tag{29}$$

for all $k, l = 1, \ldots, N$. *If additionally*

$$\Re(A(\xi)) = \Re(A(-\xi)) \quad \textit{and} \quad \Im(A(\xi)) = -\Im(A(-\xi)), \tag{30}$$

[2] Strong convergence in the Hilbert space $L^2(0,T;X)$ means $\|u_n - u\|_{L^2(0,T;X)} \to 0$.

then

$$a(\varphi_l, \varphi_k) = \frac{1}{\pi} \int_0^\infty \Re\left(A(\xi) e^{i\xi(x_l - x_k)}\right) |\widehat{\varphi_0}(\xi)|^2 \,\mathrm{d}\xi \tag{31}$$

for all $k, l = 1, \ldots, N$.

Proof. Elementary properties of the Fourier transform yield

$$\widehat{\varphi_j}(\xi) = e^{i\xi x_j} \widehat{\varphi_0}(\xi) \qquad \forall \xi \in \mathbb{R}. \tag{32}$$

Since $\varphi_j \in H_0^{\alpha/2}(\mathbb{R})$ for all $j = 1, \ldots, N$, the identity (29) follows from identity (25) with $\eta = 0$ above. The second claim (31) is then elementary. ∎

When classic hat functions on an equidistant grid with mesh size $h \in \mathbb{R}$ are chosen as basis functions with

$$\varphi_0(x) = (1 - |x|/h) \mathbb{1}_{|x| \le h} \qquad \forall x \in \mathbb{R} \tag{33}$$

we have

$$\widehat{\varphi_0}(\xi) = \frac{2}{\xi^2 h}(1 - \cos(\xi h)) \qquad \forall \xi \in \mathbb{R}. \tag{34}$$

Corollary 7 (Symbol method for stiffness matrices). *Let A be a univariate symbol with associated operator $\mathcal{A}$ satisfying (24) with $\eta = 0$. Denote by $\varphi_j \in L^1(\mathbb{R})$, $j \in 1, \ldots, N$ the basis functions of a Galerkin scheme associated with an equidistantly spaced grid $\Omega = \{x_1, \ldots, x_N\}$ possessing the property*

$$\varphi_j(x) = \varphi_0(x - x_j) \qquad \forall x \in \mathbb{R}, \tag{35}$$

for some $\varphi_0 : \mathbb{R} \to \mathbb{R}$ with $\varphi_0 \in H_0^{\alpha/2}(\mathbb{R})$. Then, the stiffness matrix $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ of the scheme can be computed by

$$A_{kl} = \frac{1}{2\pi} \int_{\mathbb{R}} A(\xi) e^{i\xi(x_l - x_k)} |\widehat{\varphi_0}(\xi)|^2 \,\mathrm{d}\xi \tag{36}$$

for all $k, l = 1, \ldots, N$.

Proof. The proof is an immediate consequence of Lemma 6. ∎

Remark 8 (On the symbol method for bilinear forms). *From a numerical perspective, the representations of the stiffness matrix entries provided in Lemma 6 and Corollary 7 are highly promising:*

1. *Instead of the double integrals appearing in (18), only one dimensional integrals need to be computed.*
2. *The model specific information is expressed via the symbol $\xi \mapsto A(\xi)$, which for a large set of models is available in form of an explicit function of ξ and the model parameters, a feature that we now can exploit numerically. We give a short list of examples below. For further examples we refer to [15] and [16].*

3. *Representation (36) displays the entries of the stiffness matrix as* Fourier integrals*. Moreover, the nodes appear as Fourier variables. As a consequence, Fast Fourier Transform (FFT) methods can be used to accelerate their simultaneous computation.*
4. *The essential assumption of Lemma 6 and Corollary 7 is that the basis functions are obtained by shifting (and possibly scaling) a "mother" basis function. This is the case for a large and interesting class of bases, including the wavelet bases, and in particular extends to the multivariate case. Therefore the methods we propose and analyse in this article in the univariate setting naturally extend to the multivariate case.*

Expression (3) introduced operators $\mathcal{A}$ for Lévy processes L in terms of the characteristic triplet (b, σ, F). The following examples present the respective symbols for some well known Lévy models, where the asset price follows $S_t = S_0\,\mathrm{e}^{L_t}$ for every $t \geq 0$ and r is the deterministic continuously compounding interest rate.

Example 9 (Symbol in the Black-Scholes (BS) model). *In the Black-Scholes model, determined by the Brownian volatility $\sigma^2 > 0$, the symbol is given by*

$$A(\xi) = A^{bs}(\xi) = i\xi b + \frac{1}{2}\sigma^2\xi^2, \tag{37}$$

with drift b set to

$$b = r - \frac{1}{2}\sigma^2 \tag{38}$$

as required by the no-arbitrage condition.

Example 10 (Symbol in the Merton model). *In the Merton model where $\sigma > 0$, $\lambda > 0$, $\alpha \in \mathbb{R}$ and $\beta > 0$, the symbol computes to*

$$A(\xi) = A^{merton}(\xi) = i\xi b + \frac{1}{2}\sigma^2\xi^2 - \lambda\left(e^{-i\alpha\xi - \frac{1}{2}\beta^2\xi^2} - 1\right) \tag{39}$$

with drift set to

$$b = r - \frac{1}{2}\sigma^2 - \lambda\left(e^{\alpha + \frac{\beta^2}{2}} - 1\right), \tag{40}$$

as required by the no-arbitrage condition.

Example 11 (Symbol in the CGMY model). *In the CGMY model of [6] with $\sigma > 0$, $C > 0$, $G \geq 0$, $M \geq 0$ and $Y \in (1, 2)$, the symbol computes to*

$$A(\xi) = A^{cgmy}(\xi) = i\xi b + \frac{1}{2}\sigma^2\xi^2 - C\Gamma(-Y)\left[(M + i\xi)^Y - M^Y + (G - i\xi)^Y - G^Y\right], \tag{41}$$

for all $\xi \in \mathbb{R}$, with drift b set to

$$b = r - \frac{1}{2}\sigma^2 - C\Gamma(-Y)\left[(M - 1)^Y - M^Y + (G + 1)^Y - G^Y\right] \tag{42}$$

for martingale pricing. This class is a special case of the classes referred to as Koponen and KoBoL in the literature, see e.g. [3] and as tempered stable processes.

Example 12 (Symbol in the NIG model). *With $\sigma > 0$, $\alpha > 0$, $\beta \in \mathbb{R}$ and $\delta > 0$ such that $\alpha^2 > \beta^2$, the symbol of the NIG model is given by*

$$A(\xi) = A^{nig}(\xi) = i\xi b + \frac{1}{2}\sigma^2\xi^2 - \delta\left(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta-i\xi)^2}\right) \tag{43}$$

for all $\xi \in \mathbb{R}$ with drift given by

$$b = r - \frac{1}{2}\sigma^2 - \delta\left(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+1)^2}\right) \tag{44}$$

as required by the no-arbitrage condition.

Implementing (36), we encounter new numerical challenges: From the perturbation study in Section 3.4.2 in [14] we conclude that we need to evaluate the integrals at high precision. Consider first the Black-Scholes model and choose the piecewise linear hat functions as basis elements as a toy example. Applying a standard Matlab integration routine will lead to considerable errors. To understand the effect, let us first consider the oscillatory contribution by the hat functions stemming from the Fourier transform in expression (34) to the integrands in (36). We depict $\widehat{\varphi_0}$ in Figure 1.

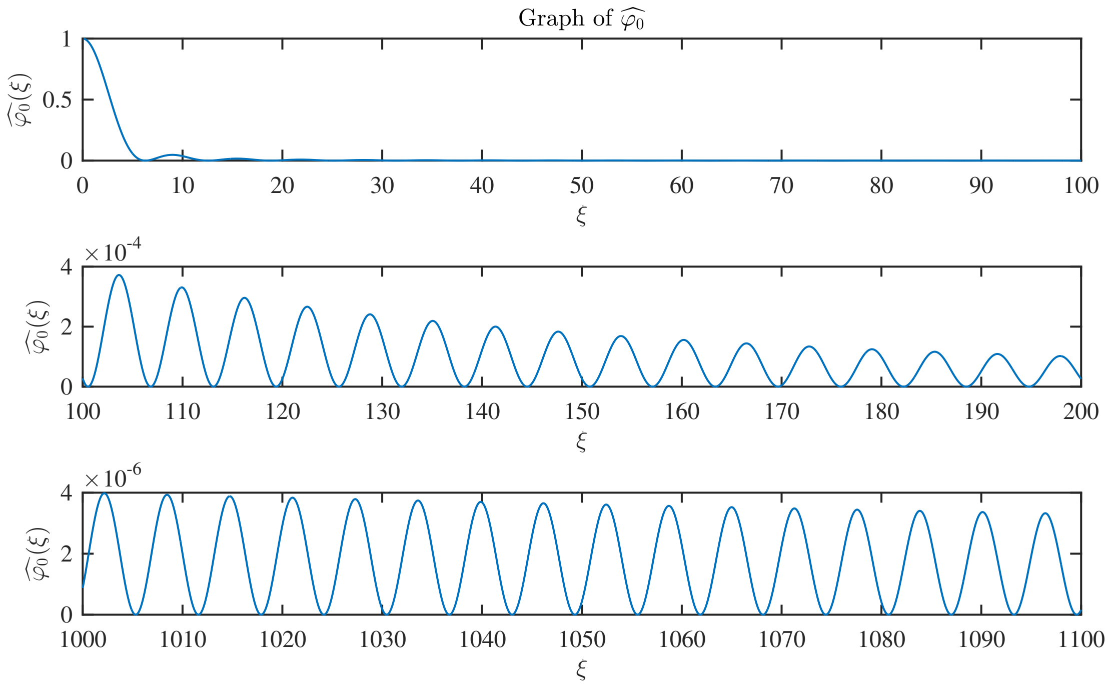

Figure 1. *Consider the hat function φ_0 of expression (33) with $h = 1$. The graph depicts its Fourier transform $\widehat{\varphi_0}$ which is evaluated over three subintervals of $\mathbb{R}^+$. The oscillations and the rather slow decay to zero complicate numerical integration with high accuracy requirements considerably.*

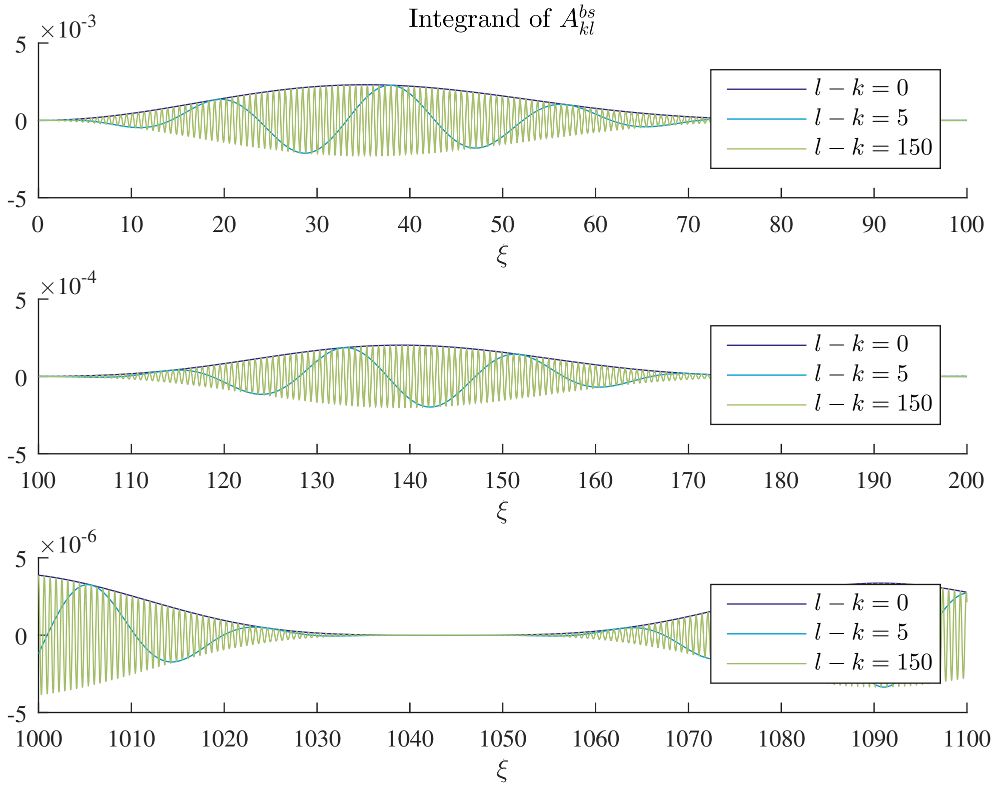

Figure 2. *The integrand for the Black-Scholes stiffness matrix A_{kl} for several values of $l-k$. The grid of the hat functions spans the interval $[-5,5]$ with 150 equidistantly spaced inner nodes and grid fineness $h = 0.0662$. A Black-Scholes solution on this grid would thus be represented by the weighted sum of 150 hat functions. We observe that oscillations of the integrand increase in the value of $|l-k|$ and so does the number of supporting points for naive numerical integration.*

Furthermore, Figure 2 shows several integrands of $\mathbf{A} \in \mathbb{R}^{N\times N}$ in the representation provided by (36) of Corollary 7 with the Black-Scholes symbol of Example 9. Therein, each integrand is evaluated for a different value of $l-k$ over three different subintervals taken from the unbounded integration range. Here, the integrands of A_{kl}, $1 \leq k, l \leq N$, have to be numerically integrated for all $l-k \in \{-(N-1), \dots, -1, 0, 1, \dots, N-1\}$.

The larger $|l-k|$, however, the more severe the numerical challenges for evaluating the integrand, as Figure 2 demonstrates. All integrands illustrated therein decay rather slowly. Additionally, oscillations increase in $|l-k|$. In combination, these two observations seriously threaten a numerically reliable evaluation of the integral. With increasing values of $|l-k|$, the oscillations of the integrand accelerate and the number of necessary supporting points for accurate integration increases. Computation of the stiffness matrix entries along these lines by invoking standard integration routines e.g. based on Matlab's `quadgk` demands considerable run times for accurate results.

These findings show that we need to further investigate the problem to obtain a flexible method to compute the stiffness matrix reliably and with low computational cost. The path that we propose here is to modify the problem in such a way that the resulting integrands decay much faster so that the domain of integration can be chosen considerably smaller and a usual integration routine such as Matlab's function `quadgk` is sufficient. To do so, we first observe that the hat functions, which we used in our toy example, are piecewise linear. While being continuous they are not continuously differentiable everywhere and thus lack smoothness on an elementary level already. This lack of smoothness translates into a slow decay of their Fourier transform or $\widehat{\varphi_0}$, respectively.

Therefore, we propose to replace the piecewise linear basis functions by basis functions that display considerably higher regularity leading to appealing decay properties of the integrands in (36). In the following two sections, we present two different approaches to implement such a problem modification.

5. From classic hat functions to mollified hats.

It is well known that convolution with a smooth function has a smoothing effect on the function that the convolution is applied to. Functions that qualify for this smoothing by convolution are called mollifiers. In order to choose an appropriate mollifier for our purposes—the fast and accurate computation of the integrals in (36), the mollifiers need to display two essential features:

(1) The Fourier transform of the modified basis function needs to be available.
(2) The smoothing effect needs to be steerable through a parameter.

As the Fourier transform of the convolution of two functions is the product of the two Fourier transformed functions, (1) boils down to the availability of the Fourier transform of the mollifier. Since the Fourier transform of standard mollifiers is not available in closed form, we widen the range of the standard mollifiers and allow for non-compact support. More precisely, we call the sequence $m=(m_k)_{k\in\mathbb{N}}$, $m_k\in L^1(\mathbb{R})$ for all $k\in\mathbb{N}$, a *mollifier*, if

1. $m_k\geq 0$, for all $k\in\mathbb{N}$,
2. $\int_{\mathbb{R}} m_k(x)\,\mathrm{d}x=1$, and
3. for all $\varrho>0$ we have the convergence $\int_{[-\varrho,\varrho]^c} m_k(x)\,\mathrm{d}x\to 0$ for $k\to\infty$.

Feature (2) is often required and we follow the usual construction here. By Proposition and Definition 2.14 in [1] we can adjust the influence of mollification by a parameter ε. To this end let $m\in L^1(\mathbb{R})$ with

$$m\geq 0,\qquad\text{and}\qquad\int_{\mathbb{R}} m(x)\,\mathrm{d}x=1. \tag{45}$$

Define

$$m^\varepsilon=\frac{1}{\varepsilon}m\left(\frac{\cdot}{\varepsilon}\right). \tag{46}$$

Then for each $\varrho>0$ we have $\int_{\mathbb{R}} m^\varepsilon(x)\,\mathrm{d}x=1$ and $\int_{[-\varrho,\varrho]^c} m^\varepsilon(x)\,\mathrm{d}x\to 0$ for $\varepsilon\to 0$. Consequently, for each null sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ the sequence $(m^{\varepsilon_k})_{k\in\mathbb{N}}$ is a mollifier in the sense of our definition.

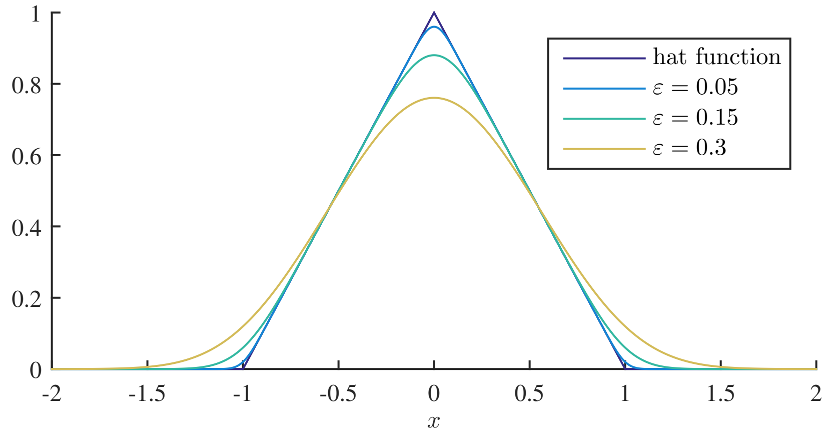

Figure 3. *A comparison between the classic hat function φ_0 on a grid with $h = 1$ and the mollified hat function $\varphi_0^\varepsilon = \varphi_0 * m^\varepsilon_{Gaussian}$ for several values of $\varepsilon \in \{0.05, 0.15, 0.3\}$ using the Gaussian mollifier of Example 13.*

Example 13 (A mollifier based on the Normal distribution). *We present an example for a mollifier. Define*

$$m_{Gaussian}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{47}$$

Then we call $(m^{\varepsilon_k}_{Gaussian})_{k\in\mathbb{N}}$ defined according to (46) a Gaussian mollifier. *The characteristic function of the Gaussian mollifier is known in closed form,*

$$\widehat{m^\varepsilon_{Gaussian}}(\xi) = \exp\left(-\frac{1}{2}\varepsilon^2\xi^2\right), \tag{48}$$

thus exhibiting exponential decay.

It is a well known result, that mollified functions $f * m_k$ converge to f in $L^p(\mathbb{R})$, $1 \le p < \infty$ when k tends to infinity, see for example Satz 2.15 in [1].

Figure 4 displays the decay of the Fourier transform of the mollified hat function in comparison with the behaviour of the hat.

5.1. Convergent Scheme based on mollified hats.

In this section we propose and analyse a convergent fully discrete scheme based on the symbol method of Section 4 and mollified hats. We also analyze the rate of convegence of the scheme. We introduce stronger assumptions that allow us to use the result of [27]. Namely, we assume elliptcity of the bilinear form instead of the weaker assumption that a Gårding inequality.

According to the symbol method introduced in Corollary 7, we solve the θ scheme (16)–(17) with stiffness matrix $\mathbf{A}$ given by equation (36),

$$\mathbf{A}_{kl} = a(\varphi_l, \varphi_k) = \frac{1}{2\pi}\int_{\mathbb{R}} A(\xi) e^{i\xi(x_l - x_k)} \left|\widehat{\varphi_0}(\xi)\right|^2 \mathrm{d}\xi \tag{49}$$

for all $k, l = 1, \dots N$, where φ_l are the hat functions and φ_0 is the hat function at the origin given by (35).

For a light notation let $m_\varepsilon := m^\varepsilon_{\text{Gaussian}}$. Following the approach we introduced in Section 5 to use mollified hats, we replace the stiffness matrix of (33) by

$$\mathbf{A}^\varepsilon_{kl} := \frac{1}{2\pi} \int_{\mathbb{R}} A(\xi) e^{i\xi(x_l - x_k)} \left|\widehat{\varphi_0}(\xi)\right|^2 \left|\widehat{m_\varepsilon}(\xi)\right|^2 \mathrm{d}\xi \tag{50}$$

On the level of the bilinear form this means we replace the bilinear form a by

$$a^\varepsilon(u, v) := \frac{1}{2\pi} \int_{\mathbb{R}} A(\xi) \widehat{u}(\xi) \overline{\widehat{v}(\xi)} \left|\widehat{m_\varepsilon}(\xi)\right|^2 \mathrm{d}\xi. \tag{51}$$

In order to achieve the optimal order of convergence of the thus perturbed θ scheme, we need to choose ε dependent on h, i.e. $\varepsilon = \varepsilon(h)$. Moreover, in an actual implementation, we will need to truncate the range of integration. In order to preserve the two fundamental properties, Gårding inequality and continuity with respect to the solution space V of the original equation, we incorporate here the *asymptotic behaviour of the symbol*. The asymptotic behaviour of the symbol plays a decisive role in the determination of the solution space. To this aim,

let $\widetilde{A} : \mathbb{R} \to \mathbb{R}$ be such that there exists $N > 0$ such that

$$|A(\xi) - \widetilde{A}(\xi)| \leq |A(\xi)|/2 \quad \text{for all } |\xi| > N \tag{52}$$

To illustrate what form $\widetilde{A}$ can take in practice, let us briefly consider a simple example. $\widetilde{A}$ carries the asymptotic behaviour of A and the convergence needs to be fast enough. This is for instance satisfied if we take for A the symbol in Merton's model from Example 10, $A(\xi) = A^{merton}(\xi) = i\xi b + \frac{1}{2}\sigma^2\xi^2 - \lambda(e^{-i\alpha\xi - \frac{1}{2}\beta^2\xi^2} - 1)$ and for $\widetilde{A}$ we use its Brownian part, $\widetilde{A}(\xi) = \frac{1}{2}\sigma^2\xi^2$.

Now let

$$\widetilde{a^\varepsilon}(u, v) := \frac{1}{2\pi} \int_{-N(\epsilon)}^{N(\epsilon)} A(\xi) \widehat{u}(\xi) \overline{\widehat{v}(\xi)} \left|\widehat{m_\varepsilon}(\xi)\right|^2 \mathrm{d}\xi + \frac{1}{2\pi} \int_{[-N(\epsilon), N(\epsilon)]^c} \widetilde{A}(\xi) \widehat{u}(\xi) \overline{\widehat{v}(\xi)} \, \mathrm{d}\xi \tag{53}$$

Now choose $N(\varepsilon) := \frac{\tilde{\delta}}{\varepsilon}$ and $\varepsilon(h) := \delta h$ for some $0 < \delta < 1$, $0 < \tilde{\delta} < \min\{\frac{1}{2\delta^2}, \frac{1}{\sqrt{2}}\}$. Then

$$N^2(\varepsilon)\varepsilon^2 < 1/2, \text{ and } N(\varepsilon(h))(\varepsilon(h))^2 \leq \tilde{\delta}\delta h \quad \text{for all } h. \tag{54}$$

Under standard conditions convergence of the fully discrete version of the (13)–(15) with a θ-scheme in time has been provided in [27]. Assuming the same standard conditions, we show that the resulting fully discrete scheme when replacing in (13)–(15) the bilinear form a by $\widetilde{a^{\epsilon(h)}}$ still leads to a convergent scheme of the same rate.

While the asymptotic behaviour of A is used in the theoretical analysis, numerically the same error behaviour is already achieved when neglecting the second term in (54), compare Section 7.4. This shows the potential of the approach even beyond the cases where the asymptotic behaviour of A is accessible in a simple form that allows to compute the second term in (53).

5.2. Convergence analysis. General assumptions and notation: $I = (a, b) \subset \mathbb{R}$, $H := L^2(I)$, $V^s := H^s(I)$, let V_h^s be a Galerkin space, e.g. the linear space spanned by the hat functions with mesh fineness h. For $\varepsilon > 0$ consider the Gauss kernel $m^{\varepsilon}_{\text{Gaussian}}$ from (46), (13). Now let $\varepsilon : (0, \infty) \to (0, \infty)$ and define $\widetilde{V_h^s} := \{(m^{\varepsilon(h)}_{\text{Gaussian}} * u_h)|_I \,|\, u_h \in V_h^s\}$, where with a slight abuse of notation, we denote by u_h the extension of u_h by zero outside of I in order to define the convolution with $m^{\varepsilon(h)}_{\text{Gaussian}}$. We notice that this extension is not necessarily in $H^s(\mathbb{R})$. We also denote $\widetilde{u_h} := (m^{\varepsilon(h)}_{\text{Gaussian}} * u_h)|_I$.

We denote by $u_h^0 = g_h$ the initial condition of the θ scheme.

Furthermore we set

$$V^t := \begin{cases} \widetilde{H}^s(I) & \text{if } s = \alpha, \\ H^{s+1}(I) \cap \widetilde{H}^s(I) & \text{if } t = \alpha + 1. \end{cases}$$

Finally, set $a(u, v) = \int_{\mathbb{R}^d} A(\xi)\hat{u}(\xi)\overline{\hat{v}(\xi)}\,\mathrm{d}\xi$, $\|u\|_a := \sqrt{a(u,u)}$ and $\|f\|_* := \frac{f(v_h)}{\|v_h\|_a}$.

We consider the following set of conditions that form the basis of the perturbation analysis in [27]:

Conditions 14. Fix index $\alpha \in [0, 1]$.

(A1) (Continuity and coercivity) There exist constants $0 < \beta, \gamma$ such that for all $\xi \in \mathbb{R}$,

$$\beta|\xi|^{2\alpha} \le A(\xi) \le \gamma|\xi|^{2\alpha}.$$

(A2) (Approximation property of the Galerkin space) There exists a family of bounded linear projectors $P_h : V^\alpha \to V_h^\alpha$ and a constant $C_1 > 0$ such that for all $u \in V^{\alpha+1}$

$$\|u - P_h u\|_{V^\alpha} \le C_1 h \|u\|_{V^{\alpha+1}}. \tag{55}$$

(A3) (Inverse property) There is a constant $C_{\mathrm{IP}} > 0$ independent of $h > 0$ such that with $0 \le s \le \alpha$ we have for all $u_h \in V_h^s$

$$\|u_h\|_{V^s} \le C_{\mathrm{IP}}\, h^{-s} \|u_h\|_H. \tag{56}$$

(A4) (Quasi-optimality of the initial condition) There is a constant $C_I > 0$ independent of $h > 0$ such that

$$\|g - g_h\|_H \le C_I \inf_{v_h \in V_h^s} \|g - v_h\|_H. \tag{57}$$

Condition (A1) is equivalent to the continuity and ellipticity of the bilinear form a with respect to V^α. Conditions (A2)–(A4) are basic approximation conditions on the Galerkin spaces. They are not only satisfied for V_h^s being the linear space spanned by the hat functions with mesh fineness h, but also for wavelet approximation spaces, see [27].

We consider an equidistant time grid, $t_m = m * T/(M-1)$, $m = 0, \ldots, M$ and denote $u^m = u(t_m)$, $u^{m+\theta} = \theta u^{m+1} + (1-\theta)u^m$, $u_h^\kappa = \sum_{j=1}^{\dim(V_h^\alpha)} U_j^\kappa \varphi_j$ for $\kappa = m$ or $\kappa = m + \theta$.

Let us first consider the rate of convergence of the θ scheme without perturbation that we directly obtain from Theorem 5.4 of [27], by choosing $\tilde{a} = a$ and $\nu = 0$ $p = \alpha$ and $\alpha = \varrho/2$ in their setting:

Lemma 15 (Convergence rate of the θ scheme). *Assume Conditions 14 and let $u \in W^1(0,T;V^\alpha,H)$ be the weak solution to problem (4)–(5). Then there exists a constant $\overline{C}>0$ such that*

$$\begin{aligned}\|u^M - u_h^M\|_H^2 + \Delta t \sum_{m=0}^{M-1} \|u^{m+\theta} - u_h^{m+\theta}\|_a^2 \le \overline{C}\, h^2 \max_{0\le\tau\le T} \|u(\tau)\|_{V^{\alpha+1}}^2 \\ + \overline{C}\, h^2 \int_0^T \|u(\tau)\|_{V^{\alpha+1}}^2 \,\mathrm{d}\tau \\ + \overline{C} \begin{cases} (\Delta t)^2 \int_0^T \|\dot u(s)\|_*^2 \,\mathrm{d}s, & \forall\theta\in[0,1] \\ (\Delta t)^4 \int_0^T \|\ddot u(s)\|_*^2 \,\mathrm{d}s, & \theta = \frac{1}{2}\,. \end{cases}\end{aligned} \tag{58}$$

Notice that the assertion of the lemma is only meaningful if the regularity of u implies finiteness of the right-hand-side of the equation. In other words, the assertion on the convergence rate implicitly comes with regularity assumptions on the solution u.

5.2.1. Convergence rate for θ scheme, mollified hat. We denote by $(\widetilde{u_h^m})_{m=1,\dots,M}$ the interpolated solution of the θ scheme induced by $\widetilde{a^{\varepsilon(h)}}$.

Proposition 16. *The assertion of Lemma 15 also holds for the solution $(\widetilde{u_h^m})_{m=1,\dots,M}$ of the perturbed θ scheme instead of $(\widetilde{u_h^m})_{m=1,\dots,M}$.*

Proof. In view of Conditions 14, in order to apply Theorem 5.4 of [27], it is enough to verify two conditions for the perturbation of the bilinear form a, namely

(i) There exists a constant $\eta < 1$ independent of h such that

$$\left|a(u,v) - \widetilde{a^{\varepsilon(h)}}(u,v)\right| \le \eta \|u\|_a \|v\|_a \quad \text{for all } u,v\in V^\alpha. \tag{59}$$

(ii) For the family of projectors P_h of Condition (A2) there exists a constant $C>0$ independent of h such that

$$\left|a(P_h u, v_h) - \widetilde{a^{\varepsilon(h)}}(P_h u, v_h)\right| < Ch\|u\|_{V^{\alpha+1}} \|v_h\|_{V^\alpha} \quad \text{for all } u\in V^{\alpha+1}, v_h\in V_h^\alpha. \tag{60}$$

These two conditions are inequalities (3.8) and (3.9) of [27].

Verify (i): Inserting the definition, we see, denoting $N = N(\varepsilon(h))$ that

$$\begin{aligned}\left|a(u,v) - \widetilde{a^{\varepsilon(h)}}(u,v)\right| \le & \frac{1}{2\pi}\left|\int_{-N}^{N} A(\xi)\widehat{u}(\xi)\overline{\widehat{v}(\xi)} \left(1 - \left|\widehat{m_\varepsilon}(\xi)\right|^2\right) \mathrm{d}\xi\right| \\ & + \frac{1}{2\pi}\left|\int_{[-N,N]^c} \big(A - \widetilde{A}\big)(\xi)\widehat{u}(\xi)\overline{\widehat{v}(\xi)}\,\mathrm{d}\xi\right|,\end{aligned}$$

where $\left|\widehat{m_\varepsilon}(\xi)\right|^2 = \mathrm{e}^{-\varepsilon^2\xi^2}$ and $0 \le 1 - \left|\widehat{m_\varepsilon}(\xi)\right|^2 = 1 - \mathrm{e}^{-\varepsilon^2\xi^2} \le \varepsilon^2\xi^2$, and hence

$$\frac{1}{2\pi}\left|\int_{-N}^{N} A(\xi)\widehat{u}(\xi)\overline{\widehat{v}(\xi)} \left(1 - \left|\widehat{m_\varepsilon}(\xi)\right|^2\right) \mathrm{d}\xi\right| \le \varepsilon^2 N^2 \|u\|_a \|v\|_a.$$

Moreover,

$$\frac{1}{2\pi}\left|\int_{[-N,N]^c}(A-\widetilde{A})(\xi)\widehat{u}(\xi)\overline{\widehat{v}(\xi)}\,\mathrm{d}\xi\right| \leq 1/2\int_{[-N,N]^c}|A(\xi)||\widehat{u}(\xi)||\widehat{v}(\xi)|\,\mathrm{d}\xi$$
$$\leq 1/2\|u\|_a\|v\|_a.$$

Summarizing, since $\varepsilon(h)^2N(h)^2 < 1/2$ for h small enough, we have

$$\left|a(u,v)-\widetilde{a^{\varepsilon(h)}}(u,v)\right| \leq \eta\|u\|_a\|v\|_a$$

for some $\eta < 1$.

Verify (ii): We first show the assertion when we replace P_hu by u. We observe that

$$\left|a(u,v_h)-\widetilde{a^{\varepsilon(h)}}(u,v_h)\right| \leq \frac{1}{2\pi}\int_{-N}^{N}\left|A(\xi)\widehat{u}(\xi)\overline{\widehat{v}(\xi)}\right|\left(1-\left|\widehat{m_\varepsilon}(\xi)\right|^2\right)\mathrm{d}\xi$$
$$+\frac{1}{2\pi}\int_{[-N,N]^c}\left|A-\widetilde{A}(\xi)\right|\widehat{u}(\xi)\overline{\widehat{v}(\xi)}\,\mathrm{d}\xi.$$

Using Hölder's inequality, inserting again $1-\left|\widehat{m_\varepsilon}(\xi)\right|^2 = 1-\mathrm{e}^{-\varepsilon^2\xi^2} \leq \varepsilon^2\xi^2$, the continuity condition from (A1) and inequality (54) we get

$$\frac{1}{2\pi}\int_{-N}^{N}\left|A(\xi)\widehat{u}(\xi)\overline{\widehat{v_h}(\xi)}\right|\left(1-\left|\widehat{m_\varepsilon}(\xi)\right|^2\right)\mathrm{d}\xi$$
$$\leq \frac{1}{2\pi}\left(\int_{\mathbb{R}}|A(\xi)|\left|\widehat{v_h}(\xi)\right|^2\mathrm{d}\xi\right)^{1/2}\left(\int_{-N}^{N}|A(\xi)|\left|\widehat{u}(\xi)\right|^2\left(1-\left|\widehat{m_\varepsilon}(\xi)\right|^2\right)^2\mathrm{d}\xi\right)^{1/2}$$
$$\leq \sqrt{1/(2\pi)}\|v_h\|_a\left(\int_{-N}^{N}|A(\xi)|\left|\widehat{u}(\xi)\right|^2\varepsilon^4\xi^4\,\mathrm{d}\xi\right)^{1/2}$$
$$\leq \epsilon^2N\sqrt{1/(2\pi)}\|v_h\|_a\left(\int_{\mathbb{R}}|A(\xi)|\xi^2\left|\widehat{u}(\xi)\right|^2\mathrm{d}\xi\right)^{1/2}$$
$$\leq \epsilon^2N\sqrt{\gamma/(2\pi)}\|v_h\|_a\|u\|_{V^{\alpha+1}}$$
$$\leq h/(2\delta)\sqrt{\gamma/(2\pi)}\|v_h\|_a\|u\|_{V^{\alpha+1}}.$$

Finally,

$$\frac{1}{2\pi}\int_{[-N,N]^c}\left|A(\xi)-\widetilde{A}(\xi)\right|\left|\widehat{u}(\xi)\widehat{v_h}(\xi)\right|\mathrm{d}\xi$$
$$\leq \frac{1}{4\pi}\int_{[-N,N]^c}|A(\xi)|\left|\widehat{u}(\xi)\right|\left|\widehat{v_h}(\xi)\right|\mathrm{d}\xi$$
$$\leq \frac{1}{4\pi}\int_{[-N,N]^c}|A(\xi)|\left|\widehat{u}(\xi)\right|\frac{|\xi|}{N}\left|\widehat{v_h}(\xi)\right|\mathrm{d}\xi$$
$$\leq \frac{1}{4\pi N}\left(\int_{\mathbb{R}}|A(\xi)|\left|\widehat{v_h}(\xi)\right|^2\mathrm{d}\xi\right)^{1/2}\left(\int_{\mathbb{R}}|A(\xi)|\xi^2\left|u(\xi)\right|^2\mathrm{d}\xi\right)^{1/2}$$
$$\leq h\delta/(2\tilde{\delta})\sqrt{\gamma/(2\pi)}\|v_h\|_a\|u\|_{V^{\alpha+1}}.$$

Now we are in a position to derive assertion (ii): By the triangle inequality we have

$$\left|a(P_hu,v_h)-\widetilde{a^{\varepsilon(h)}}(P_hu,v_h)\right|\leq\left|a(u,v_h)-\widetilde{a^{\varepsilon(h)}}(u,v_h)\right|+\left|(a-\widetilde{a^{\varepsilon(h)}}(P_hu-u,v_h)\right|.$$

Invoking inequality (59), (60) for u instead of P_hu and approximation property (A2) of Conditions 14 show the existence of a constant $C>0$ such that

$$\left|a(P_hu,v_h)-\widetilde{a^{\varepsilon(h)}}(P_hu,v_h)\right|<Ch\|u\|_{V^{\alpha+1}}\|v_h\|_{V^\alpha}\quad\text{for all } u\in V^{\alpha+1},v_h\in V_h^\alpha. \tag{61}$$ ■

Before we test the numerical performance of this approach to modify the Galerkin scheme in Section 7 below, we introduce an alternative approach based on splines. We keep the presentation of the second approach shorter since the numerical results are more promising for the mollified hat approach.

6. Splines as basis functions.

Instead of mollification of piecewise linear basis functions, we can alternatively choose basis functions that display higher regularity itself. We therefore investigate a well-established class of finite element basis functions as candidates for our purposes, namely cubic splines. Spline theory applies to a very broad context and we refer the reader to [26] for an introduction and overview. From our perspective, splines are smooth basis functions. Their Fourier transform is accessible and the theory of function spaces they span is well-established. We give the definition of the Irwin-Hall cubic spline that inherits its name from the related probability distribution. We define the univariate *Irwin-Hall spline* $\varphi_0:\mathbb{R}\to\mathbb{R}^+$ by

$$\varphi_0(x)=\frac{1}{4}\begin{cases}(x+2)^3 & ,\ -2\leq x<-1\\ 3|x|^3-6x^2+4 & ,\ -1\leq x<1\\ (2-x)^3 & ,\ 1\leq x\leq 2\\ 0 & ,\ \text{elsewhere}\end{cases} \tag{62}$$

for all $x\in\mathbb{R}$. The spline φ_0 has compact support on $[-2,2]$ and is a cubic spline. We use it to define a spline basis:

Definition 17 (Spline basis functions on an equidistant grid). *Choose $N\in\mathbb{N}$. Assume an equidistant grid $\Omega=\{x_1,\dots,x_N\}$, $x_j\in\mathbb{R}$ for all $j=1,\dots,N$, with mesh fineness $h>0$. Let φ_0 be the Irwin-Hall spline of (62). For $j=1,\dots,N$ define*

$$\varphi_j(x)=\varphi_0((x-x_j)/h)\qquad\forall x\in\mathbb{R}.$$

We call φ_j the spline basis function associated to node *j.*

For a given grid $\Omega=\{x_1,\dots,x_N\}$, $x_j\in\mathbb{R}$, Definition 17 provides the set of spline basis functions that we also use in our numerical implementation, later. In standard literature, the Irwin-Hall basis is usually enriched with additional splines associated with the first and the last node of the grid that provide further flexibility in terms of Dirichlet and Neumann boundary conditions, see for example [26]. We omit the three Irwin-Hall basis functions associated with either of the first and the last grid nodes thus implicitly prescribing Dirichlet, Neumann and second order derivative zero boundary conditions.

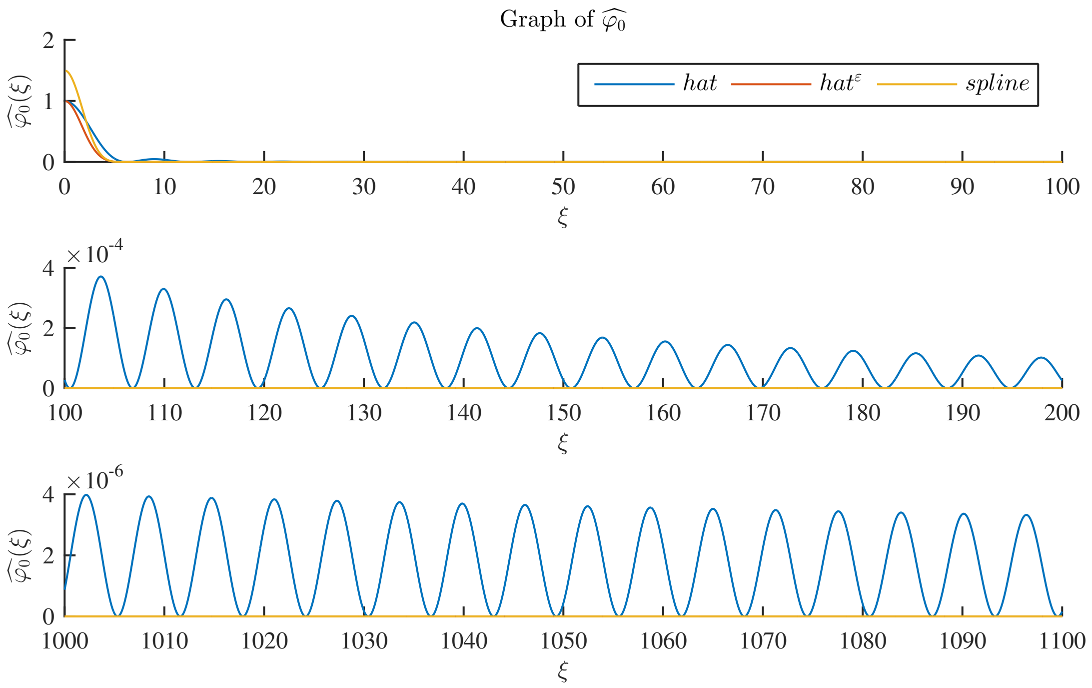

Figure 4. *Graphs of the Fourier transforms of all basis function candidates presented in this section, evaluated over three subintervals of* $\mathbb{R}^+$. *The mesh is chosen with* $h = 1$*, the mollification parameter is set to* $\varepsilon = 0.3h$.

Lemma 18 (Fourier transform of the Irwin-Hall spline). *Let φ_0 be the Irwin-Hall cubic spline of* (62). *Then its Fourier transform $\widehat{\varphi_0}$ is given by*

$$\widehat{\varphi_0}(\xi) = \frac{3}{\xi^4}\left(\cos(2\xi) - 4\cos(\xi) + 3\right) \tag{63}$$

for all $\xi \in \mathbb{R}$.

The proof of the Lemma follows by elementary calculation. This immediately gives the following corollary.

Corollary 19 (Fourier transform of spline basis functions on an equidistant grid). *Choose $N \in \mathbb{N}$. Assume an equidistant grid $\Omega = \{x_1, \ldots, x_N\}$, $x_j \in \mathbb{R}$ for all $j = 1, \ldots, N$, with mesh fineness $h > 0$ and let φ_j be the spline basis function associated with node j as defined in Definition 17. Its Fourier transform is given by*

$$\widehat{\varphi_j}(\xi) = e^{i\xi x_j}\frac{3}{h^3\xi^4}(\cos(2\xi h) - 4\cos(\xi h) + 3)$$

for all $\xi \in \mathbb{R}$.

Figure 4 compares the decay behaviour of the Fourier transforms of all three basis presented function types. As Figure 1 already illustrated, the Fourier transform of the classic hat functions exhibits both slow decay rates and oscillatory behaviour. In stark contrast the Fourier transforms of the mollified hats as well as the Fourier transform of Irwin-Hall splines visually decay to zero instantly. In case of the mollified hat functions this is due to the exponential decay of the Fourier transform of the Gaussian mollifier while for splines Corollary 19 displays a polynomial decay of order 4. In this regard, both alternatives to the classic hat functions are promising candidates for the implementation of the symbol method of Corollary 7. In Section 7 we put that promise to the test. Before that we briefly discuss the error analysis for the symbol method via spline basis functions as presented.

Convergence rate for θ scheme, splines. The spline approximation we consider falls into the framework of approximation with NURBS (non-uniform rational B-splines) of [2]. Since the geometry of our domain is the simplest possible one, namely an interval, large part of the analysis from [2] is not required in our case. Nevertheless, working with splines, we need to replace the standard Sobolev space $\widetilde{H}^1$ by a so-called *"bent" Sobolev space* $\mathcal{H}^1$, where the Sobolev spaces on the individual elements (subintervals in our case), on which the splines are cubic polynomials, are "bent" together by the corresponding regularity conditions at the interfaces, see equation (8) in [2]. Ignoring the boundary conditions we will impose, Lemma 3.3 in [2] provides the approximation property of the spline Galerkin space, (A2) from Conditions 14, and the inverse property, (A3) from Conditions 14, follows from Theorem 4.2 in [2]. Now, since the proofs in [27] do not hinge on the specific properties of the space $\widetilde{H}^1$ (also consult Section 3.6.2 of [14]) Lemma 15 extends to the setting with splines. As one might expect, both the approximation property (A2) from Conditions 14 and the inverse property (A3) are satisfied with a higher order in h, i.e. for h^4. Hence Theorem 5.4 of [27] is valid with an order of h^4. However, all terms on the right-hand side of the estimate in this theorem need to be finite, in particular $\max_{0\leq t\leq T}\|u\|_{\mathcal{H}^4}$, and therewith the respective regularity for the initial value g. As this is not given in our implementation we cannot hope for the order h^4. [3] To summarize we can expect a convergence rate of h^2 as in the case of the approach with mollified hats.

7. Numerical Implementation.

In this section we implement the pricing PIDEs for plain vanilla call and put options and test the two approaches to the symbol method experimentally.

Theorem 20 (Feynman-Kac). *Let $(L_t)_{t\geq 0}$ be a (time-homogeneous) Lévy process. Consider the PIDE*

$$\begin{aligned} \partial_t U^{C,P} + \mathcal{A}U^{C,P} + rU^{C,P} &= 0, \qquad \textit{for almost all } t\in(0,T) \\ U^{C,P}(0) &= g^{C,P}, \end{aligned} \tag{64}$$

where $\mathcal{A}$ is the operator associated with the symbol of $(L_t)_{t\geq 0}$ and $g^{C,P}\in L^2_\eta(\mathbb{R})$. Assume further the assumptions (A1)–(A3) of [11] to hold. Then (64) possesses a unique weak solution

$$U^{C,P}\in W^1(0,T;H^{\alpha/2}_\eta(\mathbb{R}),L^2_\eta(\mathbb{R})) \tag{65}$$

[3] Additional numerical experiments with smooth initial conditions, performed within in a master thesis in the working group, that we do not report in this article in more detail showed the convergence rate of h^4 thus confirming the theoretical discussion from this section.

where $\alpha > 0$ is the Sobolev index of the symbol of $(L_t)_{t\geq 0}$ and $\eta \in \mathbb{R}$ is chosen according to Theorem 6.1 in [11]. If additionally $g_\eta^{C,P} \in L^1(\mathbb{R})$ then the relation

$$U^{C,P}(T-t,x) = \mathbb{E}\left[g^{C,P}(L_{T-t}+x)\right] \tag{66}$$

holds for all $t \in [0,T]$, $x \in \mathbb{R}$.

Proof. For $r = 0$, the result is proved in [11] and follows from Theorem 6.1 therein. For general $r \geq 0$, that proof is easily adapted.

Remark 21. *Setting $g^{C,P} = g^C$ in (64), the payoff profile of a European call option, results in U^C being the price of a European call option. Analogously, setting $g^{C,P} = g^P$, the payoff profile of a European put option, results in U^P being the price of a European call option.*

7.1. Truncation to zero boundary conditions. As we derive prices of plain vanilla European call and put options, the solution to the respective pricing PIDE is defined on the whole real line. As a first step towards a discretization, we want to truncate the domain to bounded interval (a,b) and we choose to implement zero boundary conditions. Under further assumptions, exponential convergence of the truncation error has been shown in [9, Section 4.1]. Here, we follow the standard procedure to subtract an appropriate auxiliary function ψ that matches the asymptotic behavior of $U^{C,P}$. Having chosen ψ, the resulting modified problem for $\phi = U^{C,P} - \psi$ is

$$\begin{aligned} \partial_t\phi(t,x) + (\mathcal{A}\phi)(t,x) + r\phi(t,x) &= f(t,x) \qquad && \forall (t,x) \in (0,T)\times\mathbb{R}, \\ \phi(0,x) &= g_\Psi(x) && \forall x \in \mathbb{R}, \end{aligned} \tag{67}$$

where $g_\Psi(x) = g(x) - \psi(0,x)$ for all $x \in \mathbb{R}$ and the right hand side f is given by

$$f(t,x) := -\left(\partial_t\psi(t,x) + (\mathcal{A}\psi)(t,x) + r\psi(t,x)\right).$$

The solution $U^{C,P}$ to the original problem (64) can easily be restored by $U^{C,P} = \phi + \psi$. Examples for ψ will be presented, later.

The right hand side in vector notation is given by $\mathbf{F}(t^k) = (F_1(t^k),\ldots,F_N(t^k)) \in \mathbb{R}^N$ for each t^k on the time grid with $F_j(\cdot)$, $j = 1,\ldots,N$, given by

$$F_j(t) = -\int_{\mathbb{R}} \left(\partial_t\psi(t,x) + (\mathcal{A}\psi)(t,x) + r\psi(t,x)\right)\varphi_j(x)\,\mathrm{d}x \tag{68}$$

for all $j = 1,\ldots,N$.

We observe that the operator $\mathcal{A}$ applied to the auxiliary function ψ appears in the integral of expression (68). For the same reasons as in the computation of the stiffness matrix entries, we decide to apply the symbol method for the computation of the entries of the right hand side $\mathbf{F} \in \mathbb{R}^N$. We pursue these considerations in the following section.

7.2. Computation of the right hand side F. First, we need to choose an appropriate auxiliary function ψ. As its purpose is to enable us to truncate the domain and insert zero boundary conditions, we need to inspect the limit behaviour of the price value

$$\begin{aligned} &U^C(x,t) \to 0, && x \to -\infty,\ t \in [0,T] \\ &U^C(x,t) \to e^x - Ke^{-rt}, && x \to +\infty,\ t \in [0,T] \end{aligned} \tag{69}$$

for call options and

$$\begin{aligned} &U^P(x,t) \to Ke^{-rt} - e^x, && x \to -\infty,\ t \in [0,T] \\ &U^P(x,t) \to 0, && x \to +\infty,\ t \in [0,T] \end{aligned} \tag{70}$$

for put options. This is the usual way to obtain the auxiliary function. Now, in regard to our specific approach, relying on the Fourier transforms, we identify additional desirable features for the auxiliary function. We denote $\widehat{\psi}(t,z) := \widehat{\psi(t,\cdot)}(z)$. Consider a smooth function $\psi : [0,T] \times \mathbb{R} \to \mathbb{R}$ such that $\psi(t) \in H^{\alpha/2}_\eta(\mathbb{R})$ for all $t \in [0,T]$ for some $\eta \in \mathbb{R}$. Then, for the second summand in (68) we have by applying the symbol method of Lemma 4 that

$$\int_{\mathbb{R}} (\mathcal{A}\psi)(t,x)\varphi_j(x)\,\mathrm{d}x = \frac{1}{2\pi}\int_{\mathbb{R}} A(\xi - i\eta)\widehat{\psi}(t,\xi - i\eta)\overline{\widehat{\varphi_j}(\xi+\eta)}\,\mathrm{d}\xi, \tag{71}$$

where A denotes the symbol of the model. With the above identity, we are able to derive the right hand side $(F_j)_{j=1,\dots,N}$ of the PIDE in vector notation as introduced by (68) in terms of Fourier transforms by

$$F_j = -\frac{1}{2\pi}\int_{\mathbb{R}} \left(\widehat{\partial_t\psi}(t,\xi - i\eta) + (A(\xi - i\eta) + r)\widehat{\psi}(t,\xi - i\eta)\right)\overline{\widehat{\varphi_j}(\xi+\eta)}\,\mathrm{d}\xi. \tag{72}$$

This shows that ψ is numerically suitable for the purpose of localizing the pricing PIDE if ψ is quickly evaluable on the region $[a,b] \times [0,T]$ and the integrals determining F_j can be numerically evaluated fast for all $j = 1,\dots,N$. The first feature allows retransforming the solution of the localized problem into the solution of the original pricing PIDE, while the second grants the fast numerical derivation of the right hand side in equation (67). These considerations lead us to the following list of desirable features for the auxiliary function ψ that is required to obey the respective limit conditions (69), (70):

1. a (semi-)closed expression of the function ψ,
2. a (semi-)closed expression of its Fourier transform $\widehat{\psi}$
3. and fast decay of $|\widehat{\psi}(\xi)|$ and $|\widehat{\partial_t\psi}(\xi)|$ for $|\xi| \to \infty$.

The smoother ψ, the faster $|\widehat{\psi}|$ decays. In the following two subsections we analyze two candidates for ψ that display these desired features.

A first suggestion for ψ consists in using Black-Scholes prices as functions in $x = \log(S_0) \in [a,b]$ and time to maturity $t \in [0,T]$ for localization of the pricing PIDE. We express the price of a European option with payoff profile $g^{C,P}$ in the Black-Scholes model in terms of (generalized) Fourier transforms and define ψ accordingly, as the following Lemma explains.

Lemma 22 (Subtracting Black-Scholes prices). *Choose a Black-Scholes volatility $\sigma^2 > 0$, let $r \geq 0$ be the prevailing risk-free interest rate and set $\eta < -1$ in the case of a call option and $\eta > 0$ for the put. Define ψ to be the associated Black-Scholes price,*

$$\psi(t,x) = \psi^{bs}(t,x) := e^{-\eta x}e^{-rt}\frac{1}{2\pi}\int_{\mathbb{R}} e^{i\xi x}\widehat{g^{C,P}}(-(\xi + i\eta))\varphi^{bs}_{t,\sigma}(\xi + i\eta)\,\mathrm{d}\xi, \tag{73}$$

with $\varphi^{bs}_{t,\sigma}(z) = \mathrm{e}^{tA^{bs}(z)}$*. We denote by* A *the symbol of the associated operator* $\mathcal{A}$*. Then the right hand side* $\boldsymbol{F}\colon [0,T] \to \mathbb{R}^N$ *can be written in the form*

$$F_j(t) = \frac{1}{2\pi}\int_{\mathbb{R}} \Big((A^{bs} - A)(\xi - i\eta)\Big)\widehat{g^{C,P}}(\xi - i\eta)\exp\Big(-t\big(r + A^{bs}(\xi - i\eta)\big)\Big)\overline{\widehat{\varphi_j}(\xi + i\eta)}\,\mathrm{d}\xi \tag{74}$$

for all $j = 1, \dots, N$*.*

Proof. In order to derive the right hand side, we need to represent ψ in Fourier terms. Since for call and put options, $\psi \notin L^1(\mathbb{R})$, we compute the (generalized) Fourier transform of ψ or the Fourier transform of $\psi_\eta = e^{\eta\cdot} g^{C,P}$, respectively. We get

$$\psi_\eta(t,x) = e^{-rt}\frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\xi x}\widehat{g^{C,P}}(\xi - i\eta)\varphi^{\mathrm{bs}}_{t,\sigma}(-(\xi - i\eta))\,\mathrm{d}\xi. \tag{75}$$

The integral in (75) is a Fourier (inversion) integral. We read off

$$\widehat{\psi_\eta}(t,\xi) = \widehat{g^{C,P}}(\xi - i\eta)\exp\Big(-t\Big(r + A^{\mathrm{bs}}(\xi - i\eta)\Big)\Big), \tag{76}$$

where we used the relation between the characteristic function and the symbol of a process. Now,

$$\widehat{\frac{\partial}{\partial t}\psi_\eta}(t,\xi) = -\Big(r + A^{\mathrm{bs}}(\xi - i\eta)\Big)\,\widehat{\psi_\eta}(t,\xi). \tag{77}$$

Finally, since $\psi^{\mathrm{bs}} \in H^{\alpha/2}_\eta(\mathbb{R})$, we have that

$$\int_{\mathbb{R}} (\mathcal{A}\psi^{\mathrm{bs}})(t,x)\varphi_j(x)\,\mathrm{d}x = \frac{1}{2\pi}\int_{\mathbb{R}} A(\xi - i\eta)\widehat{\psi^{\mathrm{bs}}(t,\cdot)}(\xi - i\eta)\overline{\widehat{\varphi_{j_{-\eta}}}(\xi)}\,\mathrm{d}\xi. \tag{78}$$

Collecting our results proves the claim. ∎

The candidate $\psi = \psi^{\mathrm{bs}}$ matches the desired criteria. It is quickly evaluable, since Black-Scholes prices are implemented in many code libraries. Also, the integral in (74) is numerically accessible, since the integrand decays fast. Observe that FFT techniques could be employed to computed $F_j(t)$ for all $j = 1, \dots, N$ simultaneously. A major disadvantage of choosing $\psi = \psi^{\mathrm{bs}}$, however, lies in the fact that $t \in [0,T]$ can not be separated from the integrand in (74). Consequently, $F_j(t^k)$, must be numerically evaluated for each $j = 1, \dots, N$ and $k = 1, \dots, M$, individually. This results in significant numerical cost. We therefore present a second candidate for ψ.

Lemma 23 (Subtracting quasi-hockey stick). *Let* $\sigma_\psi > 0$*. Define* ψ^C *in the call option and* ψ^P *in the put option case by*

$$\begin{aligned} \psi^C(t,x) &= \left(e^x - Ke^{-rt}\right)\Phi(x), && (t,x) \in [0,T]\times[a,b],\\ \psi^P(t,x) &= \left(Ke^{-rt} - e^x\right)(1 - \Phi(x)), && (t,x) \in [0,T]\times[a,b], \end{aligned} \tag{79}$$

where Φ denotes the cumulative distribution function of the normal $\mathcal{N}(0,\sigma_\psi^2)$ distribution. Furthermore, in the call option case choose $\eta < -1$ and $\eta > 0$ in the put option case. Then, the right hand side $\boldsymbol{F}\colon [0,T] \to \mathbb{R}^N$ is given by

$$
F_j(t) = \frac{1}{2\pi}\Bigg(\int_{\mathbb{R}} \big(A(\xi - i\eta) + r\big)\frac{\widehat{f^{\mathcal{N}}}(\xi - i(\eta+1))}{i\xi + (\eta+1)}\overline{\widehat{\varphi_j}(\xi + i\eta)}\,\mathrm{d}\xi - e^{-rt}K\int_{\mathbb{R}} A(\xi - i\eta)\frac{\widehat{f^{\mathcal{N}}}(\xi - i\eta)}{i\xi + \eta}\overline{\widehat{\varphi_j}(\xi + i\eta)}\,\mathrm{d}\xi\Bigg), \tag{80}
$$

for all $j = 1,\dots,N$ with $t \in [0,T]$, where A is the symbol of the associated operator $\mathcal{A}$ and where

$$
\widehat{f^{\mathcal{N}}}(\xi) = \exp\left(-\frac{1}{2}\sigma_\psi^2\xi^2\right),
$$

the Fourier transform of the normal $\mathcal{N}(0,\sigma_\psi^2)$ density.

Proof. We consider the call option case first. To derive the expression for F_j in (80) we need to compute the Fourier transform of (the appropriately weighted) ψ^C. We choose $\eta < -1$ and $t \in [0,T]$ arbitrarily and compute for $K = 1$,

$$
\begin{aligned}
\widehat{\psi_\eta^C(t,\cdot)}(\xi) &= \int_{\mathbb{R}} e^{i\xi x}e^{\eta x}\left(e^x - e^{-rt}\right)\Phi(x)\,\mathrm{d}x \\
&= \int_{\mathbb{R}} e^{i\xi x}e^{(\eta+1)x}\Phi(x)\,\mathrm{d}x - e^{-rt}\int_{\mathbb{R}} e^{i\xi x}e^{\eta x}\Phi(x)\,\mathrm{d}x.
\end{aligned} \tag{81}
$$

Integration by parts and l'Hôpital's rule yield that

$$
\int_{\mathbb{R}} e^{i\xi x}e^{(\eta+1)x}\Phi(x)\,\mathrm{d}x = -\frac{1}{i\xi + (\eta+1)}\int_{\mathbb{R}} e^{i(\xi - i(\eta+1))x}f^{\mathcal{N}}(x)\,\mathrm{d}x, \tag{82}
$$

which can be expressed in terms of the Fourier transform of the normal distribution yielding

$$
\int_{\mathbb{R}} e^{i\xi x}e^{(\eta+1)x}\Phi(x)\,\mathrm{d}x = -\frac{\widehat{f^{\mathcal{N}}}(\xi - i(\eta+1))}{i\xi + (\eta+1)}. \tag{83}
$$

Equivalently, we obtain for the second integral in (81) that

$$
\int_{\mathbb{R}} e^{i\xi x}e^{\eta x}\Phi(x)\,\mathrm{d}x = -\frac{\widehat{f^{\mathcal{N}}}(\xi - i\eta)}{i\xi + \eta}. \tag{84}
$$

Assembling these results we find

$$
\widehat{\psi_\eta^C(t,\cdot)}(\xi) = -\frac{\widehat{f^{\mathcal{N}}}(\xi - i(\eta+1))}{i\xi + (\eta+1)} + e^{-rt}\frac{\widehat{f^{\mathcal{N}}}(\xi - i\eta)}{i\xi + \eta}. \tag{85}
$$

We deduce from (72) that

$$F_j(t) = \frac{1}{2\pi}\Bigg(\int_{\mathbb{R}} \big(A(\xi - i\eta) + r\big)\frac{\widehat{f^{\mathcal{N}}}(\xi - i(\eta+1))}{i\xi + (\eta+1)}\overline{\widehat{\varphi_j}(\xi + i\eta)}\,\mathrm{d}\xi \\ - e^{-rt}\int_{\mathbb{R}} A(\xi - i\eta)\frac{\widehat{f^{\mathcal{N}}}(\xi - i\eta)}{i\xi + \eta}\overline{\widehat{\varphi_j}(\xi + i\eta)}\,\mathrm{d}\xi\Bigg). \tag{86}$$

For the put option case we choose ψ^P as defined in (79). The computations for $\widehat{\psi^P_\eta}$ follow along the same lines as for the call and we get the relation

$$\widehat{\psi^P_\eta(t,\cdot)}(\xi) = \widehat{\psi^C_\eta(t,\cdot)}(\xi) \qquad \forall (t,\xi) \in [0,T]\times\mathbb{R}, \tag{87}$$

for η set to some $\eta > 0$, which proves the claim. ■

Remark 24 (Computational features of ψ^C and ψ^P). *While ψ^C serves as localizing function for the call option case, ψ^P can be used in the put option case. Both candidates are based on the payoff functions of call and put options but avoid the lack of differentiability with respect to x in $x = \log(Ke^{-rt})$ for $t \in [0,T]$. As a consequence, both ψ^C and ψ^P are smooth functions and thus fulfill the requirements collected above when σ_ψ is chosen small enough. Additionally, the two integrals in (80) do not depend on the time variable $t \in [0,T]$ and thus need to be computed only once for each basis function φ_j. This results in a significant acceleration in computational time compared to the suggestion $\psi = \psi^{bs}$ of Lemma 22.*

Algorithm 1 summarizes the abstract structure of a general FEM solver based on the symbol method. By plugging the symbol associated to the model of choice into the computation of line 9 of the algorithm, the solver instantly adapts to that model. In other words, only one line needs to be specified to obtain a model specific solver for option pricing. As Examples 9, 10, 11, 12 and others emphasize, the symbol exists in analytically (semi–)closed form for many models, indeed. Algorithm 1 thus provides a very appealing tool for FEM pricing in practice.

7.3. Implementation of the symbol method. As outlined in sections 5 and 6, we implement two versions of the symbol method. On the one hand, we approximate the entries of the stiffness matrix according to the approach of mollified hats, on the other hand we use Irwin-Hall cubic splines as basic functions. For the mollified hats, we simplify the scheme proposed in Section 5 further. Namely, we omit the second term in the defining equation (53) for the approximate bilinear form and we truncate the first integral at a fixed level. The numerical results already show the convergence rate of h^2 for this simplified version, thanks to the small magnitude of the tail integral.

7.4. Empirical Convergence Results. Now we implement the symbol method for both mollified hats and splines. Finally, we conduct an empirical order of convergence study. We consider the univariate Merton, CGMY and NIG model and investigate the empirical rates of convergence for the different implementations as Table 1 summarizes. For each model and each implemented basis function type enlisted in the table we consider the payoff function

$$g(x) = \max(e^x - 1, 0). \tag{88}$$

Algorithm 1 A symbol method based FEM solver

Choose equidistant space grid x_i, $i = 1, \ldots, N$
Choose basis functions φ_i, $i = 1, \ldots, N$, with $\varphi_i(x) = \varphi_0(x - x_i)$ for some φ_0
Choose equidistant time grid T_j, $j = 0, \ldots, M$
Procedure Compute Mass Matrix $\mathbf{M}$
  Derive the mass matrix $\mathbf{M} \in \mathbb{R}^{N \times N}$ by
    $M_{kl} = \int_{\mathbb{R}} \varphi_l(x)\varphi_k(x)\,\mathrm{d}x \qquad \forall k, l = 1, \ldots, N$
Procedure Compute Stiffness Matrix $\mathbf{A}$
  Derive the stiffness matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ by plugging the symbol A of the chosen model into the following formula and computing
    $A_{kl} = \frac{1}{2\pi} \int_{\mathbb{R}} A(\xi)\, e^{i\xi(x_k - x_l)}\, |\widehat{\varphi_0}(\xi)|^2 \,\mathrm{d}\xi + rM_{kl} \qquad \forall k, l = 1, \ldots, N$
  using numerical integration
Procedure Run Theta Scheme
  Choose a function ψ to subtract from the original pricing problem to obtain a zero boundary problem and retrieve the respective basis function coefficient vectors $\overline{\psi}^k \in \mathbb{R}^N$, $k = 1, \ldots, M$. Consider the suggestions by Lemma 22 or Lemma 23 for plain vanilla European options above.
  Choose an appropriate basis function coefficient vector $U^1 \in \mathbb{R}^N$ matching the initial condition of the transformed problem
  Derive the right hand side vectors $\mathbf{F}^k \in \mathbb{R}^N$, $k = 0, \ldots, M$. Consult Lemma 22 or Lemma 23 matching the choice of ψ.
  Choose $\theta \in [0, 1]$ and run the iterative scheme
  for $k = 0 : (M - 1)$
    $U^{k+1} = (\mathbf{M} + \Delta t\, \theta\, \mathbf{A})^{-1} \big((\mathbf{M} - \Delta t\, (1-\theta)\, \mathbf{A})\, U^k + \theta \mathbf{F}^{k+1} + (1-\theta)\mathbf{F}^k \big)$
  end
Procedure Reconstruct Solution to Original Problem
  Add previously subtracted right hand side ψ to the solution of the transformed problem by computing
    $\widetilde{U}^k = U^k + \overline{\psi}^k, \qquad k = 0, \ldots, M$
  to retrieve the basis function coefficient vectors $\widetilde{U}^k$, $k = 0, \ldots, M$, to the original pricing problem

of a call option with strike $K = 1$. In each study we compute FEM prices for N_k basis functions with

$$N_k = 1 + 2^k, \qquad k = 4, \ldots, 9, \tag{89}$$

resulting in $N_4 = 17$ basis functions in the most coarse and $N_9 = 513$ basis functions in the most granular case. On each grid, the nodes that basis functions are associated with are equidistantly spaced and the supports of the basis functions cover the space interval $\Omega = [-5, 5]$. The time discretization is kept constant with $N_{\text{time}} = 2000$ equidistantly spaced

Model	Symbol	Parameter choices		Implemented basis functions	
				Mollified hats	Splines
Merton	Example 10	$\sigma = 0.15$, $\beta = 0.2$,	$\alpha = -0.04$, $\lambda = 3$	✓	✓
CGMY	Example 11	$C = 0.5$, $M = 27.24$,	$G = 23.78$, $Y = 1.1$	✓	✓
NIG	Example 12	$\alpha = 12.26$, $\delta = 0.52$	$\beta = -5.77$,	✓	✓

Table 1

An overview of the models considered in the empirical order of convergence analysis and their parametrization. For these models, the symbol method is implemented and tested for both mollified hat functions and splines. In all models, the constant risk-less interest rate has been set to $r = 0.03$.

time nodes spanning a grid range of two years up until maturity, thus covering a time to maturity interval of

$$[T_1, T_{N_{\text{time}}}], \qquad \text{with } T_1 = 0 \text{ and } T_{N_{\text{time}}} = 2. \tag{90}$$

For each $k = 4, \dots, 9$, the resulting price surface constructed by N_k basis functions in space and $N_{\text{time}} = 2000$ grid points in time is computed. A comparison of these surfaces is drawn to a price surface of most granular structure based on the same type of basis functions. We call this most granular surface *true* price surface. It rests on $N_{\text{true}} = N_{11} = 1 + 2^{11} = 2049$ basis functions in space and N_{time} grid points in time covering the same grid intervals as above, that is $\Omega = [-5, 5]$ in space and $[0, 2]$ in time, respectively. The underlying FEM implementation is thus based on distances h_{true} between grid nodes that basis functions are associated with of

$$\begin{aligned} h_{\text{true}}^{\text{mollified hat}} &= (5 - (-5))/(2 + 2^{11}) \approx 0.0049, \\ h_{\text{true}}^{\text{splines}} &= (5 - (-5))/(4 + 2^{11}) \approx 0.0049, \\ \Delta t_{\text{true}} &= 2/(2000 - 1) \approx 0.001 \end{aligned} \tag{91}$$

in space and time, respectively. Note that all space grids are designed in such a way that the log-strike $\log(K) = 0$ is one of the space nodes. For each model and method and each $k = 4, \dots, 9$ the (discrete) L^2 error ε_{L^2} is calculated as

$$\varepsilon_{L^2}(k) = \sqrt{\Delta t_{\text{true}} \cdot h_{\text{true}} \cdot \sum_{i=1}^{N_{\text{time}}} \sum_{j=1}^{N_{\text{true}}} \Big(Price_{\text{true}}(i,j) - Price_k(i,j) \Big)^2},$$

wherein $Price_{\text{true}}(i,j)$ is the value of the true pricing surface at space node $j \in \{1, \dots, 1+2^{11}\}$ and time node $i = 1, \dots, 2000$ and $Price_k(i,j)$ is the respective, linearly interpolated value of

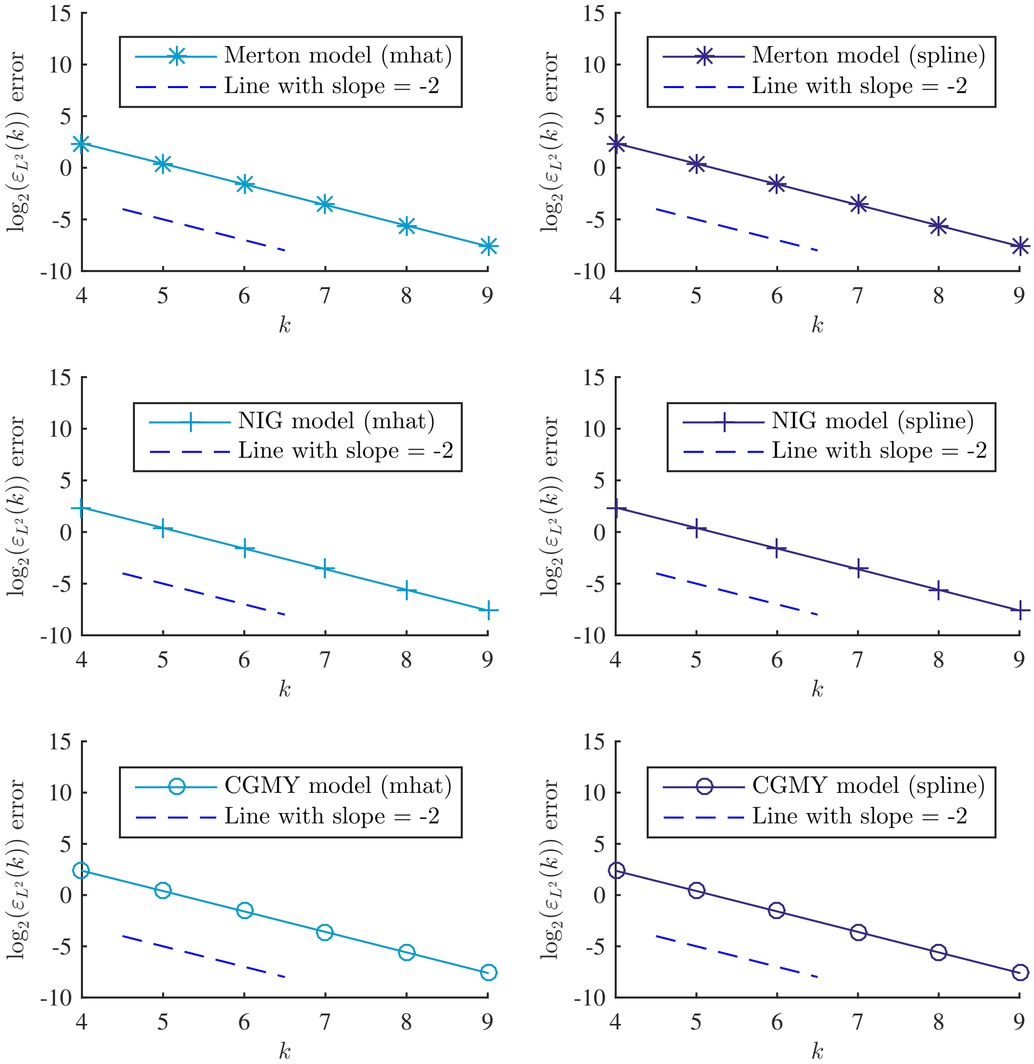

Figure 5. *Results of the empirical order of convergence study for the Merton, the NIG and the CGMY model using mollified hats (left pictures) and splines (right pictures) as basis functions. All models are parametrized as stated in Table 1. Additionally, part of a straight line with (absolute) slope of 2 is depicted in each figure serving as a comparison.*

the coarser pricing surface supported by only N_k basis functions.

Figure 5 summarizes the results of the six studies of empirical order of convergence in the Merton, the NIG and the CGMY model in a symbol based implementation once using mollified hats and once using splines as basis functions. In each implementation and for all considered models, the (discrete) L^2 error decays exponentially with rate 2. The convergence result of Theorem 5.4 by [28] suggest that this is the best possible rate we can hope for, which yields the experimental validation of both approaches.

8. Conclusion and outlook.

We have presented a tool for finite element solvers that allows for an implementation that is highly flexible in the model choice and that maintains numerical feasibility. Invoking the symbol was key. The transition into Fourier space has introduced smoothness as a new requirement to the basis functions. We have presented mollified hats and splines as compatible basis functions in our approach. Several numerical examples have confirmed the convergence rates expected by the theoretical considerations in both cases.

Let us mention several possible extensions of the approach. Firstly, the implementation naturally extends to time-inhomogeneous Lévy models that we neglected here for notational convenience. Secondly, combining the symbol method with wavelet basis functions allows for compression techniques that might further improve the overall numerical performance, as Hilber, Reichmann, Schwab and Winter in [18] outline. Thirdly, the polynomial decay that we observe in our numerical experiments can possibly be improved to exponential rates by invoking an hp-discontinuous Galerkin scheme, see e.g. Schötzau and Schwab in [25]. Fourthly, the method can be extended to multivariate settings. In particular, tensor-based multivariate extensions are conceptually straightforward. Since the domain for financial applications typically is a (hyper)rectangular, tensorized extensions of the basis functions are a natural choice. Both the mollified hats and the splines have natural tensorized generalizations.

Appendix A. Proofs.

A.1. Proof of a more general version of Lemma 3.

Proof. We first consider $\varphi, \psi \in C_0^\infty(\mathbb{R})$.

For $F \equiv 0$ the assertion follows directly from partial integration. Since the Lévy measure may be unbounded around the origin, the representation of the jump part of the bilinear form,

$$a^{jump}(\varphi,\psi) := -\int_{\mathbb{R}}\int_{\mathbb{R}} \Big(\varphi(x+y) - \varphi(x) - \varphi'(x)\,h(y)\Big) F(\mathrm{d}y)\psi(x)\,\mathrm{e}^{2\langle \eta, x\rangle}\,\mathrm{d}x,$$

needs to be carefully derived. In order to exploit the identity

$$\varphi(x+y) - \varphi(x) - y\varphi'(x) = \int_0^y \int_0^z \varphi''(v)\,\mathrm{d}v\,\mathrm{d}z$$

we split the integral with respect to the Lévy measure in three parts, set $c(F) := \int_{|y|<1} \big(y - h(y)\big) F(\mathrm{d}y) - \int_{|y|>1} h(y) F(\mathrm{d}y)$ and obtain

$$\begin{aligned} a^{jump}(\varphi,\psi) := &-\int_{\mathbb{R}}\int_{|y|<1}\int_0^y\int_0^z \varphi''(x+v)\,\mathrm{d}v\,\mathrm{d}z F(\mathrm{d}y)\psi(x)\,\mathrm{e}^{2\langle \eta, x\rangle}\,\mathrm{d}x \\ &- c(F)\int_{\mathbb{R}} \varphi'(x)\psi(x)\,\mathrm{e}^{2\langle \eta, x\rangle}\,\mathrm{d}x \\ &- \int_{\mathbb{R}}\int_{|y|>1} \big(\varphi(x+y) - \varphi(x)\big) F(\mathrm{d}y)\psi(x)\,\mathrm{e}^{2\langle \eta, x\rangle}\,\mathrm{d}x. \end{aligned}$$

Thanks to $\int_0^y \int_0^z \left|\varphi''(v)\right| \mathrm{d}v\,\mathrm{d}z \le cy^2$ with some constant $c > 0$ for all $y \in [-1,1]$ and

$$\begin{aligned}\int_{\mathbb{R}} \int_{|y|<1} \int_0^y \int_0^z \left|\varphi'(x+v)\right| \mathrm{d}v\,\mathrm{d}z F(\mathrm{d}y) \left|\psi'(x) + 2\eta\psi(x)\right| \mathrm{e}^{2\langle \eta, x\rangle}\,\mathrm{d}x \\ \le (1+2\eta)\|\varphi\|_{H^1_\eta}\|\psi\|_{H^1_\eta} \int_{|y|<1} y^2 F(\mathrm{d}y)\end{aligned} \tag{92}$$

we can apply the theorem of Fubini and partial integration to obtain

$$\begin{aligned}-\int_{\mathbb{R}} \int_{|y|<1} \int_0^y \int_0^z \varphi''(x+v)\,\mathrm{d}v\,\mathrm{d}z F(\mathrm{d}y) \psi(x)\,\mathrm{e}^{2\langle \eta, x\rangle}\,\mathrm{d}x \\ = \int_{\mathbb{R}} \int_{|y|<1} \int_0^y \int_0^z \varphi'(x+v)\,\mathrm{d}v F(\mathrm{d}y) \big(\psi'(x) + 2\eta\psi(x)\big)\,\mathrm{e}^{2\langle \eta, x\rangle}\,\mathrm{d}x.\end{aligned}$$

This yields the assertion for $\varphi, \psi \in C_0^\infty(\mathbb{R})$.

Next, we verify that the bilinear form as stated in Lemma 3 is well defined for $\varphi, \psi \in H^1_\eta(\mathbb{R})$ and is continuous with respect to the norm of $H^1_\eta(\mathbb{R})$. For $F \equiv 0$ this is obvious. The assertion follows for the jump part from inequality (92) and

$$\int_{\mathbb{R}} \int_{|y|>1} \left|\varphi(x+y) - \varphi(x)\right| F(\mathrm{d}y) \left|\psi(x)\right| \mathrm{e}^{2\langle \eta, x\rangle}\,\mathrm{d}x \le 2F\big(\mathbb{R}\setminus[-1,1]\big)\|\varphi\|_{L^2_\eta}\|\psi\|_{L^2_\eta}.$$

Thus a from Lemma 3 is a continuous bilinear form on $H^1_\eta(\mathbb{R}) \times H^1_\eta(\mathbb{R})$ that coincides with (9) on the dense subset $C_0^\infty(\mathbb{R}) \times C_0^\infty(\mathbb{R})$. This proves the assertion. ■

A.2. Proof of Lemma 5.

Proof. To prove the assertion, we verify the conditions of Lemma 7.1 in [16], which provides an abstract robustness result for weak solutions. We first observe that the conditions for f_n, f, g_n, g coincide with those of Lemma 7.1 in [16]. Second, we verify conditions (An1)–(An3) of Lemma 7.1 in [16]. Therefore we assign to each $u, v \in X$ the coefficients $\alpha_k(u), \alpha_k(v) \in \mathbb{R}$ for $k \le N$ such that $u = \sum_{k=1}^N \alpha_k(u)\varphi_k$ and $v = \sum_{k=1}^N \alpha_k(v)\varphi_k$. Thanks to the finite dimensionality of X, there exists a constant $\widetilde{C} > 0$ such that for all $u \in X$,

$$\|u\|_V \le \sum_{k=1}^N \left|\alpha_k(u)\right| \|u\|_V \le C'\|u\|_V. \tag{93}$$

Thanks to (26) there exists a sequence $0 < c_n \to 0$ such hat for all $j, k \le N$,

$$\left|(a_n - a)(\varphi_j, \varphi_k)\right| \le c_n \|\varphi_j\|_V \|\varphi_k\|_V. \tag{94}$$

Together with assumption (A2) this yields for all $j, k \le N$,

$$\left|a_n(\varphi_j, \varphi_k)\right| \le C_1 \|\varphi_j\|_V \|\varphi_k\|_V. \tag{95}$$

Inequalities (95) and (93) together yield for all $u, v \in X$,

$$
\begin{aligned}
\big|a_n(u,v)\big| &\le \sum_{j=1}^{N}\sum_{k=1}^{N}\big|\alpha_j(u)\alpha_k(v)\big|\big|a_n(\varphi_j,\varphi_k)\big| \\
&\le C_1\sum_{j=1}^{N}\sum_{k=1}^{N}\big|\alpha_j(u)\alpha_k(u)\big|\|\varphi_j\|_V\|\varphi_k\|_V \\
&\le C_1\widetilde{C}^2\|u\|_V\|v\|_V,
\end{aligned}
$$

which shows that condition (An1) of Lemma 7.1 in [16] is satisfied. Due to inequalities (94) and (93), we have for all $u \in X$,

$$
\begin{aligned}
\big|(a-a_n)(u,u)\big| &\le \sum_{j=1}^{N}\sum_{k=1}^{N}\big|\alpha_j(u)\alpha_k(u)\big|\big|a_n(\varphi_j,\varphi_k)\big| \\
&\le c_n\sum_{j=1}^{N}\sum_{k=1}^{N}\big|\alpha_j(u)\alpha_k(u)\big|\|\varphi_j\|_V\|\varphi_k\|_V \\
&\le c_n\widetilde{C}^2\|u\|_V^2,
\end{aligned}
$$

which shows assumption (An3) of Lemma 7.1 in [16]. Finally, from assumption (A1) and the last inequality for all $u \in X$ we obtain

$$
\begin{aligned}
a_n(u,u) &\ge a(u,u) - \big|(a-a_n)(u,u)\big| \\
&\ge G\|u\|_V^2 - G'\|u\|_H^2 - c_n\widetilde{C}^2\|u\|_V^2,
\end{aligned}
$$

which shows that there exists $N_0 \in \mathbb{N}$ such that condition (An2) of Lemma 7.1 in [16] is satisfied for all $n > N_0$. This shows the assertion of Lemma 5. ■

REFERENCES

[1] H. Alt, *Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung*, Springer Lehrbuch, Springer, 6th ed., 2011.

[2] Y. Bazilevs, L. Beirão da Veiga, J. Cottrell, T. Hughes, and G. Sangalli, *Isogemeotric Analysis: Approximation, stability and error estimates for h-refined meshes*, Mathematical Models and Methods in Applied Sciences, 16 (2006), pp. 1031–1090.

[3] S. I. Boyarchenko and S. Z. Levendorskiĭ, *Barrier options and touch-and-out options under regular Lévy processes of exponential type*, Annals of Applied Probability, 12 (2002), pp. 1261–1298.

[4] R. Brummelhuis and R. Chan, *An RBF scheme for option pricing in exponential Lévy models*, Applied Mathematical Finance, 21 (2014), pp. 238–269.

[5] O. Burkovska, B. Haasdonk, J. Salomon, and B. Wohlmuth, *Reduced basis methods for pricing options with the Black-Scholes and Heston models*, SIAM Journal of Financial Mathematics, 6 (2015), pp. 685–712.

[6] P. Carr, H. Geman, D. B. Madan, and M. Yor, *The fine structure of asset returns: An empirical investigation*, Journal of Business, 75 (2002), pp. 305–332.

[7] R. Chan and S. Hubbert, *Options pricing under the one-dimensional jump-diffusion model using the radial basis function interpolation scheme*, Review of Derivatives Research, 17 (2014), pp. 161–189.
[8] R. Cont, N. Lantos, and O. Pironneau, *A reduced basis for option pricing*, SIAM Journal on Financial Mathematics, 2 (2011), pp. 287–316.
[9] R. Cont and E. Voltchkova, *A finite difference scheme for option pricing in jump diffusion and exponential Lévy models*, SIAM Journal on Numerical Analysis, 43 (2005), pp. 1596–1626.
[10] R. Cont and E. Voltchkova, *Integro-differential equations for option prices in exponential Lévy models*, Finance and Stochastics, 9 (2005), pp. 299–325.
[11] E. Eberlein and K. Glau, *PIDEs for pricing European options in Lévy models – a Fourier approach.* Technische Universität München, 2011.
[12] E. Eberlein and K. Glau, *Variational solutions of the pricing PIDEs for European options in Lévy models*, Applied Mathematical Finance, 21 (2014), pp. 417–450.
[13] H. Föllmer, *Alles richtig und trotzdem falsch?*, Mitteilungen der DMV, 17 (2009), pp. 148–154.
[14] M. Gass, *PIDE methods and concepts for parametric option pricing*, PhD thesis, Technical University Munich, 2016. online version forthcoming.
[15] K. Glau, *Classification of Lévy processes with parabolic Kolmogorov backward equations.* forthcoming in SIAM Journal Theory of Probability and its Applications, 2016.
[16] K. Glau, *Feynman-Kac type formula for Lévy processes with discontionuous killing rate.* accepted for publication in Finance and Stochastics, 2016.
[17] N. Hilber, N. Reich, C. Winter, and C. Schwab, *Numerical methods for Lévy processes*, Finance and Stochastics, 13 (2009), pp. 471–500.
[18] N. Hilber, O. Reichmann, C. Schwab, and C. Winter, *Computational Methods for Quantitative Finance*, Springer, 2013.
[19] A. Itkin, *Efficient solution of backward jump-diffusion PIDEs with splitting and matrix exponentials*, Journal of Computational Finance, 19 (2016), pp. 29–70.
[20] A. Itkin and P. Carr, *Using pseudo-parabolic and fractional equations for option pricing in jump diffusion models*, Computational Economics, 40 (2012), pp. 63–104.
[21] K. Jackson, S. Jaimungal, and V. Surkov, *Fourier space time stepping for option pricing with Lévy Models*, Journal of Computational Finance, 12 (2012), pp. 1–29.
[22] D. Marazzina, O. Reichmann, and C. Schwab, *hp-DGFEM for Kolmogorov-Fokker-Planck equations of multivariate Lévy processes*, Mathematical Models and Methods in Applied Sciences (M3AS), 22 (2012).
[23] A.-M. Matache, P.-A. Nitsche, and C. Schwab, *Wavelet Galerkin pricing of American options on Lévy driven assets*, Quantitative Finance, 5 (2005), pp. 403–424.
[24] S. Salmi, J. Toivanen, and L. von Sydow, *An IMEX-scheme for pricing options under stochastic volatility models with jumps*, SIAM Journal on Scientific Computing, 36 (2014), pp. B817–B834.
[25] D. Schötzau and C. Schwab, *Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method*, SIAM Journal of Numerical Analysis, 38 (2006), pp. 837–875.
[26] L. Schumaker, *Spline Functions: Basic Theory*, Cambridge Mathematical Library, Cambridge University Press, 3rd ed., 2007.
[27] T. von Petersdorff and C. Schwab, *Wavelet discretizations of parabolic integrodifferential equations*, SIAM Journal on Numerical Analysis, 41 (2003), pp. 159–180 (electronic).
[28] T. von Petersdorff and C. Schwab, *Wavelet discretizations of parabolic integrodifferential equations*, SIAM Journal on Numerical Analysis, 41 (2003), pp. 159–180.
[29] J. Wloka, *Partial Differential Equations*, Cambridge University Press, 1987.
[30] E. Zeidler, *Nonlinear Functional Analysis and its Applications*, vol. II/A, Springer, 1990. Linear Monotone Operators.