Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Swiss Federal Institute of Technology Zurich

Semester Thesis

Variance-optimal hedging error in geometric Lévy models

Author:
Yavor Stoev

Supervisor:
Prof. Dr. Christoph Schwab

Degree Programme:
Applied Mathematics MSc

July 5, 2010

Abstract

This semester thesis discusses the theoretical and numerical considerations concerning the solution of the variance-optimal hedging problem in geometric Lévy models. It is based on the PhD thesis of Vesenmayer [13].

Contents

1 Introduction

The aim of this semester thesis is to present a numerical method for computation of the variance-optimal hedging error of European style option in geometric (i.e. exponential) Lévy market models. Such models are in general incomplete and thus we cannot obtain a perfect hedge for every contingent claim. Therefore we choose as criterium the variance-optimal hedging error and we try to find a trading strategy which minimizes this error with respect to an European contingent claim. In [2] a probabilistic expression in terms of the so-called carré-du-champ operator for the variance-optimal trading strategy is derived. Moreover, in [13] it is proved that the minimized hedging error solves a partial integro-differential equation (PIDE) involving the generator of the underlying process. We successively apply standard numerical methods to solve this PIDE - deriving the variational formulation, localization, discretization in space and time. We also strive for efficiency and use matrix compression methods as well as discontinuous Galerkin time-stepping as developed in [8] for dealing with parabolic integro-differential equations. We provide more detailed probabilistical treatment and state main results concerning the error analysis of the numerical method. Implementation issues are also discussed.

We give some remarks on notation. Apart from using f' to denote first derivative and $\dot{f}$ to denote derivative with respect to time, we use $D^{j_1,\ldots,j_n}_{i_1,\ldots,i_n} f$ for denoting the partial derivative of the function $f(x_1,\ldots,x_m)$ with respect to the variables $(x_{i_1},\ldots,x_{i_n})$ of order $(j_1,\ldots,j_n)$. $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{N}$ denote the sets of all real, complex and positive integer numbers respectively, $\mathbb{N}_0$ denotes the set of positive integers including 0 and $\mathbb{R}_{\geq 0}$ the set of nonnegative real numbers. C^p denotes the space of p-times continuously differentiable functions on $\mathbb{R}$ and L^p the standard Lebesgue spaces on $\mathbb{R}$ for $1 \leq p \leq \infty$. With (f,g) we denote the L^2 scalar product, i.e. $(f,g) := \int_{\mathbb{R}} f(x)\overline{g(x)}dx$. $\Re z$ and $\Im z$ denote the real and imaginary parts of the complex number z. We also occasionally denote the exponential function by exp.

2 Preliminaries

2.1 Bochner, Sobolev and Hölder spaces

In the realm of parabolic differential equations, Bochner spaces play a crucial role, because the solutions lie in such spaces. Assume $J \subset \mathbb{R}$ and X a Banach space with norm denoted by $\|\cdot\|_X$.

Definition 2.1. *Given a function $f : J \to X$, where we identify all functions which are almost everywhere equal on J, and $1 \leq p \leq \infty$ the Bochner spaces are:*

$$L^p(J; X) := \{f : J \to X; f \text{ is measurable and } \|f\|_{L^p(J;X)}\},$$
$$L^\infty(J; X) := \{f : J \to X; f \text{ is measurable and } \|f\|_{L^\infty(J;X)}\},$$

where

$$\|f\|_{L^p(J;X)} := \left(\int_J \|f(s)\|_X^p ds\right)^{1/p},$$
$$\|f\|_{L^\infty(J;X)} := \operatorname{ess\,sup}_J \|f\|_X,$$

We can also generalize the notion of Sobolev spaces in this setting. Let $H^0(J, X) := L^2(J, X)$ and for $k \in \mathbb{N}$ we define:

$$H^k(J, X) := \{f \in H^{k-1}(J, X); f(t) = f(s_0) + \int_{s_0}^t g(s)ds$$
$$\text{for some } s_0 \in J \text{ and } g \in H^{k-1}(J, X)\},$$

with the corresponding norm

$$\|f\|_{H^k(J,X)} := \|f\|_{H^{k-1}(J,X)} + \|g\|_{H^{k-1}(J,X)}.$$

We recall also the fractional generalization of the spaces C^r (here $r \in \mathbb{N}_0$) of r-times continuously differentiable functions - namely the Hölder spaces. Define now, for $r \in \mathbb{R}$, $r \geq 0$ and $r = [r] + \{r\}$, where $0 \leq \{r\} < 1$ and $f : \mathbb{R} \to \mathbb{R}$:

$$\|f\|_C := \sup_{x \in \mathbb{R}} |f(x)|,$$
$$\|f\|_{C^{[r]}} := \sum_{k=0}^{[r]} \|D^k f\|_C,$$
$$\|f\|_{C^r} := \|f\|_{C^{[r]}} + \sup_{y \neq 0} \frac{|D^{[r]} f(x + y) - D^{[r]} f(x)|}{|y|^r},$$
$$C^r := \{f \in C^{[r]}; \|f\|_{C^r} < \infty\}.$$

We will need a weaker notion than smoothness, provided by the so called Sobolev spaces of fractional order. In order to introduce them, recall the definition of the Fourier and the inverse Fourier transforms on the space L^1 of Lebesgue integrable complex-valued functions:

$$\mathcal{F}f := \hat{f}(z) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-izx} f(x) dx$$
$$(\mathcal{F}^{-1}f)(x) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{izx} f(z) dz,$$

We have the following properties for $f, g \in L^1$ (see [10] Theorem 7.7, Theorem 7.2)

1. The inversion formula $f(x) = \mathcal{F}^{-1}(\mathcal{F}f)(x)$ holds almost everywhere with respect to the Lebesgue measure.

2. Defining the convolution of f and g by

$$(f * g)(x) := \int_{\mathbb{R}} f(x-y)g(y)dy,$$

we have that $\widehat{f * g} = \sqrt{2\pi}\hat{f}\hat{g}$ and if, additionally, f and g lie in L^2, the space of square integrable complex-valued functions, we have $\widehat{fg} = \sqrt{2\pi}\hat{f} * \hat{g}$.

The Fourier transform can be extended from $L^1 \cap L^2$ to L^2 such that $\mathcal{F} : L^2 \to L^2$ is linear isometry that preserves the L^2 scalar product and is called the Fourier-Plancherel transform (see [10] Theorem 7.9 ff). Important role for further extending the fourier transform will play the following space:

Definition 2.2. *A function $f : \mathbb{R} \to \mathbb{C}$ is called rapidly decreasing if:*

$$\|f\|_{\mathcal{S},N} := \sup_{k \le N} \sup x \in \mathbb{R}(1+x^2)^N |D^k f(x)| < \infty,$$

for all $N \in \mathbb{N}_0$. We denote the space of all rapidly decreasing functions by $\mathcal{S}$. $\mathcal{S}$ is also called Schwartz space, and is locally-convex topological space with respect to the countable collection of semi-norms $(\|.\|_{\mathcal{S},N})_{N\in\mathbb{N}}$ (see [10] Theorem 1.37). Its dual space is the so called space of tempered distributions:

$$\mathcal{S}' := \{f \in \mathcal{L}(\mathcal{S}, \mathbb{C}); f \textit{ is continuous}\},$$

where $\mathcal{L}(\mathcal{S}, \mathbb{C})$ is the space of all complex-valued linear functionals on $\mathcal{S}$.

Now we extend $\mathcal{F}$ on $\mathcal{S}'$ for every $f \in \mathcal{S}'$ as:

$$\mathcal{F}(f)(\Phi) := f(\mathcal{F}(\Phi)) \qquad \forall \Phi \in \mathcal{S},$$

and it turns out that $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$ is continuous with continuous inverse and coincides with the Fourier and Fourier-Plancherel transforms on L^1 and L^2 respectively (see [10] Definition 7.14 ff). Now, by using the isometry on L^2, we can define the Sobolev spaces and norms for $s \in \mathbb{R}$ as follows:

$$H^s := \{f \in \mathcal{S}'; \|f\|_{H^s} < \infty\}$$

where

$$\|f\|_{H^s} := \|\mathcal{F}^{-1}\left((1+|.|^2)^{s/2}\mathcal{F}f(.)\right)\|_{L^2}.$$

H^s is called Sobolev space of order s. It is a Hilbert space with scalar product:

$$(f,g)_{H^s} := \int_{\mathbb{R}} (1+|z|^2)^s \hat{f}(z)\overline{\hat{g}(z)}dz.$$

We also have that $(H^s)^* := \mathcal{L}(H^s,\mathbb{C})$, the dual of H^s, is isomorphic to H^{-s} (see [1] Theorem 7.63). So we can define the so-called *duality pairing* as follows:

$$\langle f,g\rangle_{(H^s)^*\times H^s} := (f,g)_{L^2} \qquad \forall f \in H^{-s},\ g \in H^s.$$

To state some of the error estimation and stability results we would also need a weighted version of the above spaces.

Definition 2.3. *Let $\omega \in \mathbb{R}$ and denote by $\mathcal{S}_\omega$ the space of all $f \in C^\infty$ such that $e^{\omega\cdot} f(\cdot) \in \mathcal{S}$ with corresponding system of semi-norms:*

$$\|f\|_{\mathcal{S}_\omega,N} := \|e^{\omega\cdot} f\|_{\mathcal{S},N}.$$

The dual space shall be denoted by $\mathcal{S}'_{-\omega}$.

By definition $\mathcal{S}_\omega \subset \mathcal{S}$ and therefore $\mathcal{S}' \subset \mathcal{S}'_{-\omega}$. Notice that if $f \in \mathcal{S}_\omega$ then $\hat{f}(\cdot + i\omega)$ exists. It follows that we can extend the Fourier transform to $\mathcal{S}'_{-\omega}$ by defining for every $u \in \mathcal{S}'_{-\omega}$ the functional $\hat{u}(\cdot - i\omega) \in \mathcal{S}'$ as

$$\hat{u}(\cdot - i\omega)(\hat{v}(\cdot + i\omega)) := u(v) \qquad \forall v \in \mathcal{S}_\omega.$$

Now we can define the weighted Sobolev spaces and norms of order $s \in \mathbb{R}$ as follows:

$$H^s_\omega := \{f \in \mathcal{S}'_\omega; \|f\|_{H^s_\omega} < \infty\}$$

where

$$\|f\|_{H^s_\omega} := \left(\int_{\mathbb{R}} (1+|\xi|^2)^s |\hat{f}(\xi + i\omega)|^2 d\xi\right)^{1/2}.$$

Similarly, H^s_ω is a Hilbert space with scalar product:

$$(f,g)_{H^s_\omega} := \int_{\mathbb{R}} (1+|z|^2)^s \hat{f}(z+i\omega)\overline{\hat{g}(z+i\omega)}dz.$$

We will also define for short $\|f\|^2_{H^s_{\omega_1,\omega_2}} := \|f\|^2_{H^s_{\omega_1}} + \|f\|^2_{H^s_{\omega_2}}$.
Again, $(H^s_\omega)^*$, the dual of H^s_ω, is isomorphic to $H^{-s}_{-\omega}$ and the duality pairing is as follows:

$$\langle f,g\rangle_{(H^s_\omega)^* \times H^s_\omega} := \int_{\mathbb{R}} \hat{f}(z-i\omega)\overline{\hat{g}(z+i\omega)}dz \qquad \forall f \in H^{-s}_{-\omega},\ g \in H^s_\omega.$$

2.2 Markov processes, semimartingale characteristics and Lévy processes

Denote first by $C_0(\mathbb{R}^d)$ the space of all functions $f : \mathbb{R}^d \to \mathbb{R}$ such that f is continuous and vanishing at infinity and by $C^\infty_c(\mathbb{R}^d)$ the space of all functions $f : \mathbb{R}^d \to \mathbb{R}$ such that f is smooth and with compact support.
Given a probability space $(\Omega, \mathcal{F}, P)$, s stochastic process on it is a mapping X from $\mathbb{R}_{\geq 0}$ into the random variables (in this case we take $\mathbb{R}^d$-valued random variables) and we denote it by $(\Omega, \mathcal{F}, P, (X_t)_{t\geq 0})$. A universal process is the family $(\Omega, \mathcal{F}, P^x, (X_t)_{t\geq 0})_{x\in\mathbb{R}^d}$, such that for every $x \in \mathbb{R}^d$, $(\Omega, \mathcal{F}, P^x, (X_t)_{t\geq 0})$ is stochastic process , P^x is a Borel measurable mapping and $P^x(X_0 = x) = 1$. Let $(\mathcal{F}_t)_{t\geq 0}$ be the filtration generated by the process X.

Definition 2.4. *The universal process X is called a Markov process if for all Borel sets $A \in \mathcal{B}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$:*

$$E^x(\mathbb{1}_A(X_{s+t})|\mathcal{F}_s) = E^{X_s}(\mathbb{1}_A(X_t)) \qquad P^x - a.e.,$$

where $\mathbb{1}_A : \mathbb{R}^d \to \{0,1\}$ is the indicator function of $A \subseteq \mathbb{R}^d$ and E^x denotes the expectation with respect to the probability measure P^x.

Let now X be a Markov process. Define on the space of bounded Borel functions on $\mathbb{R}^d$, $B_b(\mathbb{R}^d)$ the operator family $(T_t)_{t\geq 0}$ by:

$$
\begin{aligned}
T_t &: \mathcal{B}_b(\mathbb{R}^d) \to \mathcal{B}_b(\mathbb{R}^d) \\
T_t f(x) &:= E^x(f(X_t))
\end{aligned}
$$

It is easy to see, by the Markov property, that this family is a semigroup, i.e. for $s, t \geq 0$ we have $T_{s+t} = T_s \circ T_t$. Further, an operator semigroup $(T_t)_{t\geq 0}$ on $C_0(\mathbb{R}^d)$ is called a Feller semigroup if it satisfies the following conditions:

1. $T_t : C_0(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$ is linear and $\|T_t f\|_C \leq \|f\|_C$
2. $\lim_{t\to 0}\|T_t f - f\|_C = 0$, i.e. the semigroup is strongly continuous
3. $0 \leq f \leq 1$ implies $0 \leq T_t f \leq 1$.

A Markov process is called a Feller process if the restriction of $(T_t)_{t\geq 0}$ on $C_0(\mathbb{R}^d)$ is a Feller semigroup.
The generator of a Feller semigroup is the following operator:

$$
\begin{aligned}
A &: D(A) \to C_0(\mathbb{R}^d) \\
Af &:= \lim_{t\to 0} \frac{T_t f - f}{t},
\end{aligned}
$$

where

$$D(A) := \{f \in C_0(\mathbb{R}^d); \lim_{t\to 0} \frac{T_t f - f}{t} \text{ exists with respect to the topology of } C_0(\mathbb{R}^d)\},$$

is its domain. The operator A is densely defined and closed, and uniquely determines $(T_t)_{t\geq 0}$.
An important and tractable class of Feller processes are the Lévy processes. In our probabilistic framework the underlying stochastic process will be a Lévy process.

Definition 2.5. *A Markov process $(\Omega, \mathcal{F}, P^x, (X_t)_{t\geq 0})_{x\in\mathbb{R}^d}$ is called a Lévy process if for every $x \in \mathbb{R}^d$, the stochastic process $(\Omega, \mathcal{F}, P^x, (X_t)_{t\geq 0})$ satisfies:*

1. *$X_0 = x$ a.e.*
2. *For every $n \in \mathbb{N}$ and $0 \leq t_0 < t_1 < \cdots < t_n$, the random variables $X_{t_0} - X_0, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent and have the same distribution (stationarity) as $X_{t_0} - X_0, X_{t_1 - t_0} - X_0, X_{t_2 - t_1} - X_0, \ldots, X_{t_n - t_{n-1}} - X_0$.*

3. For every $t \geq 0$ and $\epsilon > 0$

$$\lim_{s \to t} P^x(|X_s - X_t| > \epsilon) = 0.$$

The generator A of a Lévy process, can be represented in a specific form, namely pseudo differential form. Specifically, for Lévy process (X_t) the generator restricted to $C_c^\infty(\mathbb{R}^d)$ is of the form:

$$Af(x) = -\mathcal{F}^{-1}(\Psi^X(\cdot)\mathcal{F}f(\cdot))(x),$$

where $\Psi^X : \mathbb{R}^d \to \mathcal{C}$ is called the symbol of the pseudo differential operator (PDO) A and is given by:

$$\Psi^X(\xi) = -\lim_{t \to 0} \frac{\hat{P}^x_{X_t - x}(\xi) - 1}{t},$$

where $P^x_{X_t - x}$ is the distribution function of $X_t - x$ under P^x, and therefore Ψ^X is also the characteristic exponent of X. Finally, the Fourier transform of $P^x_{X_t - x}$ is actually the characteristic function of $X_t - x$ defined by:

$$\hat{P}^x_{X_t - x}(\xi) := (2\pi)^{-1/2} E^x(e^{i(X_t - x)^\top \xi})$$

Lévy processes belong to the general class of semimartingales. Semimartingales are stochastic processes with respect to which stochastic integration can be well-defined. They provide a setting in which we can do stochastic analysis. The notion of semimartingale characteristics is helpful in studying the local properties of a semimartingale, i.e. the properties at each time $t \in \mathbb{R}_+$.
Assume that X is $\mathbb{R}^d$-valued semimartingale. A well-known fact is that X can be decomposed as $X_t = X_0 + M_t + A_t$ where M is $\mathbb{R}^d$-valued local martingale and A is $\mathbb{R}^d$-valued finite variation process. Informally, the characteristics of X are the elements of the triplet (B, C, ν) where B is $\mathbb{R}^d$-valued predictable process corresponding to the finite variation part of X, C is $\mathbb{R}^{d \times d}$ continuous process corresponding to the quadratic variation of the continuous martingale part of X, and ν is a predictable random measure on $\mathbb{R}_+ \times \mathbb{R}^d$ corresponding to the expectation of the number of jumps of X. The characteristic triplet of a semimartingale depends on a so-called truncation function. A truncation function is a bounded function $h : \mathbb{R}^d \to \mathbb{R}^d$, such that $h(x)$ is asymptotically equal to x as $x \to 0$. The truncation function enters the expression for B by compensating for the expectation of small

jumps (i.e. jumps with absolute magnitude less than a fixed constant). For precise semimartingale characteristics definition we refer to [6]. Now fix a truncation function $h : \mathbb{R}^d \to \mathbb{R}^d$ defined by $h(x) = \mathbb{1}_{|x|\leq 1} x$.

Definition 2.6. *Let the semimartingale X has characteristics (B, C, ν). Let b, c are predictable processes and F a transition kernel from $(\Omega \times \mathbb{R}_+, \mathcal{P})$ to $(\mathbb{R}^d, \mathcal{B}^d)$, i.e. $F_t(G) = P^x(\Delta X_t \in G)\lambda(t)$ under P^x for some positive valued function $\lambda(t)$ and $\Delta X_t = X_t - X_{t-}$ is the jump of the process X at time t. Then (b, c, F) are called differential characteristics of X if:*

$$\begin{aligned} B_t &= \int_0^t b_s ds, \\ C_t &= \int_0^t c_s ds, \\ \nu([0,t] \times G) &= \int_0^t F_s(G) ds, \qquad \forall G \in \mathcal{B}^d. \end{aligned}$$

For a Lévy process it turns out that the differential characteristics are deterministic and correspond to the so-called Lévy-Khintchine triplet.

Proposition 2.7. *Suppose $((X_t)_{t\geq 0}, P^x)_{x\in\mathbb{R}}$ is a Markov process such that $X - x$ is also a semimartingale for all $x \in \mathbb{R}$. Then $((X_t)_{t\geq 0}, P^x)_{x\in\mathbb{R}}$ is a Lévy process if and only if $X - x$ admits a version of its differential characteristics (b, c, F) (or equivalently (b, σ^2, F))that is deterministic and independent of time. In this case (b, c, F) coincides with the so-called Lévy-Khintchine triplet, i.e. the characteristic function of $X_t - x$:*

$$E(e^{iu(X_t - x)}) = \exp\left[t\left(iub - \frac{1}{2}ucu + \int_{\mathbb{R}} (e^{iuy} - 1 - iuh(y))F(dy)\right)\right].$$

Proof. See [6], Corollary II.4.19 $\square$

We need to ensure that our price process is a martingale, since we will be working in a no-arbitrage framework. The following theorem gives a useful result for exponentials of one-dimensional Lévy processes.

Proposition 2.8. *Let X be a one-dimensional Lévy process with characteristic triplet (b, c, F) and assume that $\int_{|z|>1} |z| F(dz) < \infty$ and $\int_{|z|>1} \mathrm{e}^z F(dz) < \infty$. Then e^X is a martingale if and only if*

$$\frac{c}{2} + b + \int_{\mathbb{R}} (\mathrm{e}^z - 1 - z) F(dz) = 0.$$

Proof. Using independence and stationarity of increments, for $0 \leq s < t$ and expectation with respect to P^0 we get

$$\begin{aligned} E[\mathrm{e}^{X_t}|\mathcal{F}_s] =&\mathrm{e}^{X_s}E[\mathrm{e}^{X_t - X_s}] = \mathrm{e}^{X_s}E[\mathrm{e}^{X_{t-s}}] \\ &= \mathrm{e}^{X_s}\exp\left[(t-s)\left(b + \frac{c}{2} + \int_{\mathbb{R}}(e^y - 1 - h(y))F(dy)\right)\right], \end{aligned}$$

and by using the truncation function $h(x) := x$ we obtain the result. □

We give also the following consequence of Ito formula for semimartingales which we would need for the derivation of the partial integro-differential equation (PIDE) for the hedging error.

Theorem 2.9. *Given a semimartingale X with characteristics (B, C, ν) and a $C_b^2(\mathbb{R}^d)$ function (i.e. bounded twice continuously differentiable) $f : \mathbb{R}^d \to \mathbb{R}^d$ the process*

$$\begin{aligned} f(X) - f(X_0) - \sum_{j \leq d} D_j f(X_-)B^j - \frac{1}{2}\sum_{j,k \leq d} D_{jk} f(X_-)C^{jk} \\ - \left(f(X_- + x) - f(X_-) - \sum_{j \leq d} D_j f(X_-)h^j(x)\right) \bullet \nu \end{aligned}$$

is a local martingale, where $\bullet$ denotes stochastic integration.

Proof. See [6], Theorem II.2.42 □

Using Theorem 2.9, Proposition 2.7 and the expression for the generator of a Lévy process via the characteristic exponent we get the following result which we also use in the derivation of the PIDE for the hedging error.

Corollary 2.10. *For a one-dimensional Lévy process X with characteristic triplet (b, c, F) the generator A is given by*

$$(Af)(x) = \frac{c}{2}D^2 f(x) + bDf(x) + \int_{\mathbb{R}}(f(x+z) - f(x) - zDf(x))F(dz),$$

where $f \in C_b^2(\mathbb{R})$. Moreover, the process $M_t := f(X_t) - \int_0^t (Af)(X_s)ds$ is a local martingale.

2.3 Semigroup approach to parabolic differential equations

For deriving the solution of the hedging problem, we would need to make a connection between the PIDE that we must solve and a certain class of operators that are semigroup generators. Then we would be able to represent the solution of a PIDE by Duhamel's principle as an image of an exponential of an operator and use functional analysis to derive error bounds.
In this section we follow [7]. Let X and H are two Hilbert spaces such that we have the so-called *Gelfand* or *Evolution triple*, $X \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} X^*$, where $\overset{d}{\hookrightarrow}$ denotes a dense embedding and we identify H with its dual H^* by Riesz representation theorem. Let $a : X \times X \to \mathbb{R}$ be a bilinear form satisfying:

1. Continuity: $|a(v,w)| \leq C_1 \|v\|_X \|w\|_X, \qquad \forall v,\ w \in X$
2. Coercivity: $a(v,v) \geq C_2 \|v\|^2, \qquad \forall v \in X,$

where $C_1,\ C_2 > 0$. Assume also $g \in L^2((0,T);X^*)$ and $u_0 \in H$. We want to solve the following variational equation:
Find $u \in H^1((0,T);X^*) \cap L^2((0,T);X)$ such that for all $v \in X$ and $t \in (0,T)$ we have:

$$\begin{aligned} &\frac{d}{dt}\langle u(t), v\rangle_{X^* \times X} + a(u(t), v) = \langle g(t), v\rangle_{X^* \times X}, \\ &\langle u(0), v\rangle_{X^* \times X} = (u_0, v)_H. \end{aligned}$$

We would need the following definition:

Definition 2.11. *The operator $A : D(A) \subset X_0 \to X_0$ is called sectorial if and only if*

1. *A is linear, closed and densely defined in X_0,*
2. *If $\rho(A)$ denotes the resolvent set of A, there exists sectorial constant $\theta \in (0, \pi/2)$, such that*

$$G := \{z \in \mathbb{C} : \theta \leq |\arg z| \leq \pi\} \subset \rho(A),$$

and a positive constant $\tilde{C}$ such that

$$\|(z-A)^{-1} f\|_{X_0} \leq \frac{\tilde{C}}{|z|} \|f\|_{X_0}, \qquad \forall z \in G.$$

Now making a connection between the above definition and our variational equation, we have:

Lemma 2.12. *Define the linear operator $A : X \to X^*$ as*

$$\langle Av, w\rangle_{X^* \times X} := a(v, w), \qquad \forall v,\ w \in X,$$

where a is continuous and coercive bilinear form. Then $A : X \subset X^* \to X^*$ is sectorial operator. The operator $\tilde{A}$ equal to A on its domain

$$D(\tilde{A}) := \{f \in X; Af \in H\} \subset H$$

is also sectorial with the same sectorial constant θ. Moreover, there exists a positive constant $\tilde{C}$, such that for all $z \in \rho(A)$ and $v \in X$ we have:

$$|z|\|v\|_{L^2}^2 + \|v\|_X^2 \leq \tilde{C}|z\|v\|_{L^2}^2 - a(v, v)|.$$

Proof. See [7], Lemma 1.20 □

This lemma allows us to define the exponential operator of A and to derive some properties of the corresponding semigroup.

Lemma 2.13. *Let $A : X_0 \to X_0$ be a sectorial operator with $\theta \in (0, \pi/2)$ and Γ some piecewise smooth simple curve in G running from $\infty e^{i\theta}$ to $\infty e^{-i\theta}$. We can define a strongly continuous semigroup $T_t := e^{-tA}$ on X^* by $e^{0A} := I$ and*

$$e^{-tA} := \frac{1}{2\pi i}\int_\Gamma e^{-tz}(z - A)^{-1}dz, \qquad t > 0,$$

where $t \to e^{-tA}$ for $t \in (0, \infty)$ is analytical mapping in the operator topology of $X_0 \to X_0$. Further, we have for every $f \in X_0$ and $k \in \mathbb{N}_0$

1. *The following equality is well-defined:*

$$A^k e^{-tA} f = e^{-tA} A^k f.$$

2. *The function $t \to e^{-tA} f$ is in $C^\infty((0, \infty), X_0)$ with:*

$$\frac{\partial^k}{\partial t^k} e^{-tA} f = (-1)^k A^k e^{-tA} f.$$

3. *There exist a positive constant $\tilde{C}$ such that*

$$\|A^k e^{-tA} f\|_{X_0} \leq \tilde{C} t^{-k} \|f\|_{X_0}.$$

Proof. See [7], Lemma 1.3 □

We can use the above lemma, with our operators A and $\tilde{A}$ from Lemma 2.12, to show that their exponentials are bounded, and have bounded derivatives with respect to the norms in X^* and H. By Duhamel's principle we can write the solution of our variational equation as:

$$u(t) = e^{-tA}u_0 + \int_0^t e^{-(t-\tau)A}g(\tau)d\tau.$$

3 Hegding problem formulation

We will deal with the variance-optimal hedging problem for a European option in the case of single underlying stochastic process that is a martingale. Therefore we must minimize the variance-optimal hedging error

$$E\left((\tilde{H}(S_T) - v - \int_0^T \vartheta_s dSs)^2\right),$$

with payoff function $\tilde{H}$, discounted underlying asset price S, initial investment v and admissible hedging strategy ϑ. The minimization occurs over all ϑ and v.

3.1 Assumptions on the underlying

We start with a one-dimensional Lévy process $(\Omega, \mathcal{F}, P^x, (X_t)_{t\geq 0})_{x\in\mathbb{R}}$ and set $S_t := e^{X_t}$. We will assume $P := P^0$ if not explicitly stated otherwise. We define now the type of Lévy processes that we will use in the sequel - they provide a tractable subclass of Lévy processes from analytical point of view as outlined in [3] Section 1.2.

Definition 3.1. *A Lévy process X_t is called a regular Lévy process of exponential type $[-\eta, \eta]$ and of order $\rho \in (0, 2]$ (RLPE) if there exists a function $\Phi : \mathbb{C} \to \mathbb{C}$, holomorphic in the strip $\{z \in \mathbb{C}; \Im z \in (-\eta, \eta)\}$ and continuous in $\Im z \in [-\eta, \eta]$, such that for the characteristic exponent of X is true:*

$$\Psi^X(z) = -ibz + \Phi(z).$$

Moreover, there exist $\nu_1,\ \nu_2 < \rho$ and positive constants $C_1,\ C_2$ such that for all z with $\Im z \in (-\eta, \eta)$:

$$\Phi(z) = C_1|z|^\rho + O(|z|^{\nu_1}) \qquad \text{for } |z| \to \infty$$
$$|\Phi'(z)| \leq C_2(1 + |z|)^{\nu_2}.$$

From the above definition, for $\alpha \in [-\eta, \eta]$ the exponential moment $E[\mathrm{e}^{\alpha} X_t]$ exists. The following assumptions on the underlying driving process X will ensure that S is square-integrable martingale and also would alleviate the numerical treatment later:

(A1) X is RLPE of exponential type $[-\eta, \eta]$ and of order $\rho \in (0, 2]$ with Lévy-Khintchine triplet (b, σ^2, F) with respect to the truncation function $h(x) := x$.

(A2) $\eta \geq 2$, so that S is square-integrable

(A3) $b = -\frac{1}{2}\sigma^2 - \int_{\mathbb{R}}(e^x - 1 - x)F(dx)$, so that S is a martingale by Proposition 2.8.

(A4) To control the behavior of jumps, assume there exists $0 < \nu < 2$ such that

- F is absolutely continous with respect to the Lebesgue measure with kernel function k, i.e. $F(dy) = k(y)dy$.
- k satisfies a Calderón-Zygmund estimate: For all $\alpha \in \mathbb{N}_0$ and some positive δ there exists a positive constant $C(\alpha)$, such that

$$
\begin{aligned}
|k^{(\alpha)}(z)| &\leq C(\alpha)|z|^{-(1+\nu+\alpha)}, \qquad \forall z \in \mathbb{R} \setminus \{0\} \\
|k^{(\alpha)}(z)| &\leq C(\alpha)e^{-(\eta+\delta)|z|}, \qquad \forall z \in \mathbb{R} \setminus [-1, 1].
\end{aligned}
$$

- There exists positive C_-, such that

$$
\forall z \in (0, 1) \ : \qquad \frac{1}{2}(k(-z) + k(z)) \geq \frac{C_-}{|z|^{1+\nu}}.
$$

It should be noted that the order ρ of the process X is equal to ν if $\sigma^2 = 0$ and $\rho = 2$ otherwise. Thus ρ is an indicator of the small jump activity when there is no diffusion, or alternatively, of the path properties of the process X.

3.2 Definition of optimal strategy and corresponding hedging error J

We need to specify which hedging strategies and initial investments are admissible.

Definition 3.2. *A simple process is a finite linear combination of processes of the form $h\mathbb{1}_{(\tau_1,\tau_2]}$ with $\tau_1 \le \tau_2$ stopping times and h bounded $\mathcal{F}_{\tau_1}$ - measurable random variable. The investment-strategy pair $(v, \vartheta) \in L^0(\Omega, \mathcal{F}_0, P) \times L^1(S)$ is called admissible if:*

$$v^n + \vartheta^n \bullet S_t \to v + \vartheta \bullet S_t \text{ in probability for any } t \in [0,T] \text{ and}$$
$$v^n + \vartheta^n \bullet S_T \to v + \vartheta \bullet S_T \text{ in } L^2(P),$$

for some sequences $(v^n)_{n\in\mathbb{N}}$ in $L^2(\Omega, \mathcal{F}_0, P)$ and $(\vartheta^n)_{n\in\mathbb{N}}$ of simple processes, where $\bullet$ denotes stochastic integration.

From now on, we would work with the log-price. That means that we would work with the payoff function $H(x) := \tilde{H}(e^x)$. In the general case with driving process $x + X$ for some $x \in \mathbb{R}$ we set $S^x := e^{x+X}$. Define now the following payoff process:

$$V_t(x) := E^x(H(X_T)|\mathcal{F}_t).$$

By using Markov property and stationarity of increments we get

$$\begin{aligned} V_t(x) &= E^{X_t}(H(X_{T-t})) \qquad P^x - a.s. \\ &= E^{x+X_t}(H(X_{T-t})) \qquad P - a.s. \end{aligned}$$

By defining an option price function $V(t,x) := E(H(x + X_t))$ and noticing that $V_t(x) \overset{d}{=} V(T-t, x + X_t)$ our hedging problem is solved by finding the following quantities for fixed $x \in \mathbb{R}$:

$$(v^*(x), \vartheta^*(x)) := \underset{\text{admissible pairs}(v,\vartheta)}{\operatorname{argmin}} E\left(\left(V(0, x + X_T) - v - \int_0^T \vartheta_{s-} dS_s^x\right)^2\right)$$
$$J_0(x) := E\left(\left(V(0, x + X_T) - v^*(x) - \int_0^T \vartheta^*_{s-}(x) dS_s^x\right)^2\right)$$

We are interested in European call or put payoff functions. Without loss of generality our payoff function from now on would be $H(x) = (1 - e^x)^+$, i.e. the payoff of European put option with strike 1. Indeed, the payoff of European put option with strike K is given by $H^K(x) = KH(x - \log K)$ and thus the option price function is just scaled and translated version of the initial case. By the Call-Put parity it can further be derived a relationship between the needed quantities in the hedging problem for European call and put payoff functions with strike 1(see [13] Chapter 3 for details). Denoting the corresponding quantities with superscripts c and p respectively for the call and put case, we have:

$$J_0^c(x) = J_0^p(x),\ \vartheta^{*,c}(x) = \vartheta^{*,p}(x) + 1 \text{ and } v^{*,c}(x) = v^{*,p}(x) - 1 + \mathrm{e}^x.$$

3.3 Probabilistic solution of the hedging problem - formulas for the optimal strategy and hedging error via the carré-du-champ operator Γ and Ψ

For the analysis later, it would turn out that H is not sufficiently smooth and we will introduce a smooth approximation H^ϵ depending on some parameter ϵ. We denote the corresponding approximate option price, trading strategy, initial investment and hedging error functions by $V^\epsilon, \vartheta^{\epsilon,*}, v^{\epsilon,*}$ and J_0^ϵ.

Certain operator enters the expression for the solution of the hedging problem - the carré-du-champ operator. In [2] the authors represented the solution of the hedging problem via operator expressions involving the carré-du-champ under specific assumptions. Essentially, the carré-du-champ operator represents the density of the quadratic variation of the square-integrable martingales depending on a given Markov process with respect to the Lebesgue measure.

Let $B(\mathbb{R})$ denote the set of all Borel measurable functions $f : \mathbb{R} \to \mathbb{R}$ and let

$$\begin{aligned} D(\Gamma) :=& \{(f_1, f_2) \in C^0(\mathbb{R}) \times C^0(\mathbb{R}); f_1, f_2 \text{ have left-sided derivatives and} \\ & \forall x \in \mathbb{R} : \int_{\mathbb{R}} |f_1(x+y) - f_1(x)||f_2(x+y) - f_2(x)| F(dy) < \infty\}. \end{aligned}$$

Denote by $\exp : \mathbb{R} \to \mathbb{R}$ the function $exp(x) := \mathrm{e}^x$ and define the bilinear operators $\Gamma : D(\Gamma) \to B(\mathbb{R})$ and $\psi : D(\Gamma) \to B(\mathbb{R})$ as follows:

$$\Gamma(f_1, f_2)(x) := \sigma^2 f_1'(x) f_2'(x) + \int_{\mathbb{R}} (f_1(x+y) - f_1(x))(f_2(x+y) - f_2(x)) F(dy)$$

$$\psi(f_1, f_2)(x) := \left(\Gamma(f_1, f_2)(x) - \frac{\Gamma(\exp, f_1)\Gamma(\exp, f_2)}{\Gamma(\exp, \exp)}\right)(x).$$

Important is the fact, that the operator Γ coincides with the carré-du-champ operator for X, on the intersection of the respective domains, because the following expression is true (see [2] Proposition 4):

$$\Gamma(f, g) = A(fg) - fAg - gAf,$$

where A is the generator of X and f, g and fg lie in $D(A)$. The following theorem, in which we use results of [2] Section 2.2, is crucial:

Theorem 3.3. *The solution of the hedging problem with respect to the approximate payoff H^ϵ is given by*

$$\begin{aligned} v^{\epsilon,*}(x) &= V^\epsilon(T,x) \\ \vartheta_t^{\epsilon,*}(x) &= \vartheta^\epsilon(T-t, x+X_{t-}) \\ J_0^\epsilon(x) &= J^\epsilon(T,x), \end{aligned}$$

where

$$\begin{aligned} \vartheta^\epsilon(t,x) &:= \left(\frac{\Gamma(V^\epsilon(t,\cdot),\exp)}{\Gamma(\exp,\exp)}\right)(x), \\ J^\epsilon(t,x) &:= E\left(\int_0^t \psi\left(V^\epsilon(t-s,\cdot), V^\epsilon(t-s,\cdot)\right)(x+X_{s-})ds\right). \end{aligned}$$

Proof. We essentially have to check Assumption 5, Assumption 6 and Assumption 7 in [2].
First, since X is RLPE, it has stationary independent increments and using [2] Proposition 12, we have that it admits a carré-du-champs operator, and thus Assumption 5 is satisfied.
Second, the stock price S is martingale. Moreover, S is of the form $S = \exp(x+X)$ and exp lies in the domain of the generator of X and possesses a locally square-integrable martingale part due to $(A2)$ and the Calderón-Zygmund estimate $(A4)$. Thus Assumption 6 is satisfied.
Finally, since $H = (1-e^x)^+$, and S is square integrable we also have that H is in $L^2(\Omega, \mathcal{F}_T)$. The same is trivially true also for H^ϵ since it would be constructed as to be always bounded by H outside a neighborhood of 0. With this Assumption 7 is also satisfied.
The result now follows directly from [2], Theorem 8. □

4 Derivation of the PIDE

We mentioned in the previous section, that we will use approximate payoff function, due to the non-smoothness in the original one. Indeed, we would see that for the original payoff function, the norm bounds for the original option price function V tend to infinity as t tends to zero, which is undesirable as the expression for the hedging error involves integral with

0 lower limit in time. This is not the case for V^ϵ, whose norm estimates depend on the regularization parameter ϵ but not on t. In the sequel, the results are given without proof and taken from [13], unless stated otherwise.

4.1 Regularized (approximate) payoff function. Continuity of the option price and approximate option price functions. Error of the approximation of the hedging error J^ϵ

Fix some sufficiently large $M_p \in \mathbb{N}$ and let q be the unique polynomial of degree $2M_p + 1$ such that $q(-1) = 1$, $q'(-1) = -1$ and $\forall k \in \mathbb{N}_0$: $q^{(k+2)}(-1) = q^{(k)}(1) = 0$. Now given a regularization parameter $\epsilon > 0$ define

$$q^\epsilon(x) = \epsilon q\left(\frac{e^x - 1}{\epsilon}\right).$$

The regularized approximate payoff function H^ϵ is defined as:

$$H^\epsilon(x) = \begin{cases} q^\epsilon(x) & \text{, if } \log(1-\epsilon) \leq x \leq \log(1+\epsilon) \\ H(x) & \text{, otherwise.} \end{cases}$$

By definition H^ϵ is smooth. The corresponding functions would be denoted like $V^\epsilon, \vartheta^\epsilon, \psi^\epsilon$ and J^ϵ. Fix some positive constant ϵ_0 which we will use from now on as regularization parameter, with respect to which we will estimate norms and give stability estimates. Define for short

$$v^\epsilon(r) := 1 + \epsilon^{3/2-r}$$
$$v^t(r) := 1 + t^{\frac{3/2-r}{\rho}}.$$

We now state the properties of the option price function $V(t,x)$ which can be derived just from the distribution of X_t.

Lemma 4.1. *Assume that X_t satisfies (A1). Then we have that $V \in C^\infty((0,T] \times \mathbb{R})$, and for $s,t > 0$, $\omega \in [0,\eta]$ and $k \in \mathbb{N}_0$ we have:*

$$\|D_1^k(V(t,\cdot) - V^{\epsilon_0}(t,\cdot))\|_{H^s_\omega} \leq Cv^t(s + k(\rho \vee 1) + \delta).$$

Proof. See [13] Lemma 4.1.3. $\square$

Now the norm and error estimates for V^ϵ follow. They do not depend on time, but on the regularization parameters.

Lemma 4.2. *Let $\omega \in [-\eta, \eta]$ and X_t satisfies $(A1)$. Then for $s \in [0,1]$ and $t \in [0,T]$ we have the following error bound:*

$$\|V(t,\cdot) - V^\epsilon(t,\cdot)\|_{H^s_\omega} \leq C\epsilon^{3/2-s}.$$

Further, take $\omega_1 \in (0,\eta]$,$\omega_2 \in (-1,\eta]$, non-negative s and non-negative integer k such that $0 \leq s + k(\rho \vee 1) + 1/2 \leq M_p$. Then $V^\epsilon \in H^k((0,T); H_{\omega_1^s})$ and

$$\begin{aligned}
&\|D_1^k(V^\epsilon(t,\cdot) - V^{\epsilon_0}(t,\cdot))\|_{H^s_\omega} \leq Cv^\epsilon(s + k(\rho \vee 1)),\\
&\|D_1^k V^\epsilon(t,\cdot)\|_{H^s_{\omega_1}} \leq Cv^\epsilon(s + k(\rho \vee 1)),\\
&\|D_1^k D_2 V^\epsilon(t,\cdot)\|_{H^s_{\omega_2}} \leq Cv^\epsilon(s + 1 + k(\rho \vee 1)).
\end{aligned}$$

Proof. See [13] Lemma 4.1.4. □

Corollary 4.3. *We have that $V \in C_b^\infty((0,T] \times \mathbb{R})$ and $V^\epsilon \in C_b^{\tilde{M}_1,\tilde{M}_2}([0,T] \times \mathbb{R})$, where $\tilde{M}_1, \tilde{M}_2$ are arbitrary integers such that $\tilde{M}_2 + \tilde{M}_1(\rho \vee 1) + 1/2 + \delta \leq M_p$.*

Proof. See [13] Corollary 4.1.5. □

Notice in the above that V^ϵ is $\tilde{M}_1$-times continuously differentiable in time at 0. Finally we can obtain an error bound for the approximate hedging error.

Lemma 4.4. *For the error of the regularization of the hedging error J^ϵ we have the following bound:*

$$\|J(T,X) - J^\epsilon(T,x)\|_{L^2} \leq C\epsilon^{3/2}.$$

Proof. See [13] Lemma 4.1.6. □

4.2 Properties of Γ and ψ

For the numerical treatment we also need some properties of the operators Γ and ψ. First, by the definition of Γ in Section 3 we have:

$$\Gamma(\exp, \exp)(x) = (\sigma^2 + \int_{\mathbb{R}} (e^y - 1)^2 k(y)dy)e^{2x}.$$

A Fourier approach can be applied to compute norm estimates of the other terms involving Γ, as long as the arguments are integrable in space. This is not the case for V^ϵ which is needed for estimating $\Gamma(V^\epsilon, V^\epsilon)$. However, $V^\epsilon - V^{\epsilon_0} \in L^1$ and we can split Γ in terms involving only V^{ϵ_0} whose norm we can estimate independently of ϵ, and the others which we can compute with Fourier methods.

Now define the following weighted function space:

$$D^w_{s,\omega} := \{(f, \tilde{f}) \in C^1 \times C^1; \|f - \tilde{f}\|_{L^1_\omega} + \|f - \tilde{f}\|_{H^{s+\rho/2}_\omega} + \|\tilde{f}'\|_{H^s_\omega} < \infty\}$$

and the following norm estimator:

$$\|f, \tilde{f}\|^\omega_{s,\omega} := \|f - \tilde{f}\|_{H^{s+\rho/2}_\omega} + \|\tilde{f}'\|_{H^s_\omega}.$$

This is the kind of space in which the pair $(V^\epsilon, V^{\epsilon_0})$ will lie. Building on that, we further define:

$$\begin{aligned} D^{\Gamma_1}_{\omega_f,\omega} &:= D^\omega_{0,\omega_f} \times D^\omega_{0,\omega-\omega_f}, \\ D^{\Gamma_2}_{s,\omega_f,\omega} &:= (D^\omega_{s,\omega_f} \cap D^\omega_{1/2+\delta,\omega_f}) \times (D^\omega_{s,\omega-\omega_f} \cap D^\omega_{1/2+\delta,\omega-\omega_f}), \end{aligned}$$

with corresponding norm estimators:

$$\begin{aligned} &\|f, \tilde{f}, g, \tilde{g}\|^{\Gamma_1}_{(\omega_f,\omega)} := \|f, \tilde{f}\|^\omega_{(0,\omega_f)} \|g, \tilde{g}\|^\omega_{(0,\omega-\omega_f)}, \\ &\|f, \tilde{f}, g, \tilde{g}\|^{\Gamma_2}_{(s,\omega_f,\omega)} := \|f, \tilde{f}\|^\omega_{(s,\omega_f)} \|g, \tilde{g}\|^\omega_{(1/2+\delta,\omega-\omega_f)} + \|f, \tilde{f}\|^\omega_{(1/2+\delta,\omega_f)} \|g, \tilde{g}\|^\omega_{(s,\omega-\omega_f)}. \end{aligned}$$

The above spaces take into account that we subtract from each of the original arguments f, g of the operator expression $\Gamma(f, g)$ the functions $\tilde{f}$ and $\tilde{g}$ which lie in the same weighted spaces as f and g. Moreover, the norms estimators in $D^{\Gamma_1}_{\omega_f,\omega}$ will yield L^1 bound whereas those involved in $D^{\Gamma_2}_{s,\omega_f,\omega}$ will yield L^2 bound. The following lemma gives the properties of Γ.

Lemma 4.5. *Let $\omega_f \in [-\eta/2, \eta/2]$, $\omega \in [-\eta, \eta]$, $\omega - \omega_f \in [-\eta/2, \eta/2]$, s is nonnegative and $(f, \tilde{f}, g, \tilde{g}) \in D^{\Gamma_2}_{s,\omega_f,\omega}$. Then we have:*

$$\begin{aligned} \|\Gamma(f, g)\|_{H^s_\omega} &\le C\|f, \tilde{f}, g, \tilde{g}\|^{\Gamma_2}_{(s,\omega_f,\omega)}, \\ \|e^{-x}\Gamma(f, exp)\|_{H^s_\omega} &\le C\|f, \tilde{f}\|^\omega_{(s,\omega_f)}. \end{aligned}$$

If $(f, \tilde{f}, g, \tilde{g}) \in D^{\Gamma_1}_{\omega_f,\omega}$ we have:

$$\|\Gamma(f, g)\|_{L^1_\omega} \le C\|f, \tilde{f}, g, \tilde{g}\|^{\Gamma_1}_{(\omega_f,\omega)}.$$

Proof. See [13] Lemma 4.2.2. □

A consequence of the upper lemma is the following, giving the norm estimates for ψ:

Lemma 4.6. *Let $\omega_f \in [-\eta/2, \eta/2]$, $\omega \in [-\eta, \eta]$, $\omega - \omega_f \in [-\eta/2, \eta/2]$, s is nonnegative and $(f, \tilde{f}, g, \tilde{g}) \in D^{\Gamma_2}_{s,\omega_f,\omega}$. Then we have:*

$$\|\psi(f,g)\|_{H^s_\omega} \leq C\|f, \tilde{f}, g, \tilde{g}\|^{\Gamma_2}_{(s,\omega_f,\omega)}.$$

If $(f, \tilde{f}, g, \tilde{g}) \in D^{\Gamma_1}_{\omega_f,\omega}$ we have:

$$\|\psi(f,g)\|_{L^1_\omega} \leq C\|f, \tilde{f}, g, \tilde{g}\|^{\Gamma_1}_{(\omega_f,\omega)}.$$

Proof. See [13] Lemma 4.2.3. □

4.3 Continuity of Ψ and the hedging error

From the previous lemmas the following result can be derived:

Corollary 4.7. *For $\tilde{M} \in \mathbb{N}$ such that $3/2 + \tilde{M} + \tilde{M}(\rho \vee 1) + \rho/2 \leq M_p$ we have:*

$$\begin{aligned}
&\psi(V^\epsilon, V^\epsilon) \in C^{\tilde{M}}_b([0,T] \times \mathbb{R}) \cap H^{\tilde{M}}([0,T]; H^{\tilde{M}}), \\
&J^\epsilon \in C^{\tilde{M}}_b([0,T] \times \mathbb{R}) \cap H^{\tilde{M}}((0,T); H^{\tilde{M}}),
\end{aligned}$$

and for $m \in \mathbb{N}_0$ we have:

$$\begin{aligned}
&\psi(V, V) \in C^\infty_b((0,T] \times \mathbb{R}) \cap H^m([\delta,T]; H^m), \\
&J^\epsilon \in C^{0,1+\rho/2-\delta}_b([0,T] \times \mathbb{R}) \cap L^2([0,T]; H^{3/2+\rho/2-\delta}).
\end{aligned}$$

Proof. See [13] Lemma 4.3.3. □

The above result allows us to apply Proposition 2.9 for the approximate hedging error J^ϵ in order to derive the PIDE that it satisfies, since the Ito semimartingale formula requires twice continuously differentiable function in all arguments. As can be seen, J is not sufficiently differentiable in time for our purposes.

4.4 PIDE for the J^ϵ

Armed with the previous results, we are now able to derive the PIDE for the regularized hedging error J^ϵ. Define first the operator $A^X : D(A) \to L^1$

as:

$$A^X f(x) := \int_{\mathbb{R}} (f(x+y) - f(x) - (e^y - 1)Df(x))k(y)dy +$$
$$+ \frac{1}{2}\sigma^2(D^2 f(x) - Df(x))),$$

with domain:

$$D(A) := \{f \in C^2; \int_{\mathbb{R}} |f(\cdot + y) - f(\cdot) - (e^y - 1)Df(\cdot)|k(y)dy \in L^1$$
$$\text{and } (D^2 f(\cdot) - Df(\cdot)) \in L^1\}.$$

Define also the following auxiliary set which we use later in order to extend A^X to Sobolev space of fractional order so that we can derive the variational formulation of the PIDE:

$$D_{L^1}(A) := \{f \in D(A); f, Df, D^2 f \in L^1\}.$$

From Corollary 2.10 we have that A^X coincides with the generator of X_t on the intersection of their domains. From the properties of V^ϵ and J^ϵ in Lemma 4.2 and Lemma 4.4 it turns out that $J^\epsilon(t,\cdot) \in D_{L^1}(A)$ and $V^\epsilon(t,\cdot) - V^{\epsilon_0}(t,\cdot) \in D_{L^1}(A)$. We also have $H^\epsilon \in D_{L^1}(A)$ (see [13] Lemma 5.7.1). Thus, the PIDE will be well-defined for the extended operator.

Theorem 4.8 (PIDE for J^ϵ)**.** *If X_t satisfies $(A1) - (A4)$ then we have the following PIDE for the regularized hedging error function J^ϵ*

$$\begin{aligned} &D_1 J^\epsilon(t,x) - A^X J^\epsilon(t,x) = \psi(V^\epsilon, V^\epsilon)(t,x) \qquad \forall (t,x) \in (0,T] \times \mathbb{R} \\ &J^\epsilon(0,x) = 0 \qquad \forall x \in \mathbb{R} \end{aligned} \tag{1}$$

Proof. We follow [13] Theorem 4.4.1. For $x \in \mathbb{R}$ define the process

$$M_t^1 := E\left(\int_0^T \psi(V^\epsilon, V^\epsilon)(T-s, x+X_{s-})ds|\mathcal{F}_t\right) - J^\epsilon(T,x),$$

and notice that it is local martingale and by Theorem 3.3 it is with initial value 0. It is also bounded by Corollary 4.7 and therefore a martingale. Since X is Markov process we have

$$\begin{aligned} M_t^1 + J^\epsilon(T,x) &= E^{X_t}\left(\int_0^{T-t} \psi(V^\epsilon, V^\epsilon)(T-t-s, x+X_{s-})ds\right) \\ &+ \int_0^t \psi(V^\epsilon, V^\epsilon)(T-s, x+X_{s-})ds \\ &= J^\epsilon(T-t, x+X_t) + \int_0^t \psi(V^\epsilon, V^\epsilon)(T-s, x+X_{s-})ds. \end{aligned}$$

Note that, by Corollary 4.7, $J^\epsilon \in C^2([0,T] \times \mathbb{R})$ and we can apply Theorem 2.9 for the function J^ϵ and the process $\binom{T-\cdot}{x+X_\cdot}_{T\wedge t}$. Thus, we have that

$$\begin{aligned}
M_t^2 :=& J^\epsilon(T-t, x+X_t) - J^\epsilon(T, x+X_0) \\
&- \int_0^t \Big(- D_1 J^\epsilon(T-s, x+X_{s-}) + D_2 J^\epsilon(T-s, x+X_{s-}) b \\
&+ \frac{1}{2} D_{22} J^\epsilon(T-s, x+X_{s-})\sigma^2 \\
&- \int_{\mathbb{R}} (J^\epsilon(T-s, x+y+X_{s-}) - J^\epsilon(T-s, x+X_{s-}) \\
&- y D_2 J^\epsilon(T-s, x+X_{s-}))k(y)dy \Big) ds \\
&= J^\epsilon(T-t, x+X_t) - J^\epsilon(T, x+X_0) \\
&- \int_0^t \Big(-D_1 J^\epsilon(T-s, x+X_{s-}) + A^X J^\epsilon(T-s, x+X_{s-}) \Big) ds
\end{aligned}$$

is a local martingale. Again by Corollary 4.7 we have that $J^\epsilon, D_2 J^\epsilon, D_{22} J^\epsilon$ and $D_1 J^\epsilon$ are bounded. For the part $A^X J^\epsilon$ we have by triangle inequality and Taylor expansion:

$$\begin{aligned}
&\Big\| \int_{\mathbb{R}} (J^\epsilon(t, x+y) - J^\epsilon(t,x) - (e^y - 1) D_2 J^\epsilon(t,x)) k(y) dy \Big\|_{C([0,T]\times\mathbb{R})} \\
&\leq \Big\| \int_{\mathbb{R}} \int_0^1 \int_0^{\theta_1} D_{22} J^\epsilon(t, x+\theta_2 y) d\theta_2 d\theta_1 y^2 k(y) dy \Big\|_{C([0,T]\times\mathbb{R})} \\
&\quad + \int_{\mathbb{R}} |e^y - 1 - y| k(y) dy \|D_2 J^\epsilon\|_{C([0,T]\times\mathbb{R})} \\
&\leq \|D_{22} J^\epsilon\|_{C([0,T]\times\mathbb{R})} \int_{\mathbb{R}} y^2 k(y) dy + C \|D_2 J^\epsilon\|_{C([0,T]\times\mathbb{R})} \\
&\leq C(\|D_2 J^\epsilon\|_{C([0,T]\times\mathbb{R})} + \|D_{22} J^\epsilon\|_{C([0,T]\times\mathbb{R})}),
\end{aligned}$$

and again by the same argument M_t^2 is martingale. Setting

$$\begin{aligned}
Y_s :=& - D_1 J^\epsilon(T-s, x+X_{s-}) + A^X J^\epsilon(T-s, x+X_{s-}) \\
&+ \psi(V^\epsilon, V^\epsilon)(T-s, x+X_{s-}) \\
Z_t :=& \int_0^t Y_s ds,
\end{aligned}$$

we notice that $M_t^1 - M_t^2 = Z_t$, and therefore Z is also a martingale. By Corollary 4.7 and the previous argument, Y is bounded and thus we can write a finite upper bound for the variation of Z for all $\omega \in \Omega$

$$Var(Z)_t(\omega) \leq \int_0^t |Y_s(\omega)| ds.$$

Thus, Z is finite variation continuous martingale, and therefore pathwise 0 a.e. On the other hand, Y is cadlag by continuity of $D_1 J^\epsilon, A^X J^\epsilon$ and $\psi(V^\epsilon, V^\epsilon)$, meaning that for almost all $\omega \in \Omega$ and all $t \in (0, T]$ we have

$$0 = \lim_{h \to 0+} \frac{Z_{t-h}(\omega) - Z_t(\omega)}{h} = Y_t(\omega),$$

and Y is also pathwise 0 a.e. Now, for the range of X we can apply Theorem 24.10(i) from Sato and get that for $t \in (0, T]$ and almost all $x \in \mathbb{R}$ there exists $\omega \in \Omega$ such that $X_t(\omega) = x$. Thus we have that for $t \in (0, T]$ and almost all $x \in \mathbb{R}$

$$-D_1 J^\epsilon(t, x) + A^X J^\epsilon(t, x) + \psi(V^\epsilon, V^\epsilon)(t, x) = 0,$$

but from continuity of the functions involved, this holds true for every $(t, x) \in [0, T] \times \mathbb{R}$. □

5 Numerical solution of the PIDE

In this chapter we will develop the numerical method for solution of both the option price and hedging error functions PIDEs. To this end we would follow closely [8], and apply the numerical techniques there. More specifically, we would use matrix compression and approximate assembly, as well as some other standard methods when dealing numerically with PDEs. In the end we would make our numerical implementation dependent on only one parameter, namely the mesh width h. Thus the convergence rates and errors will be expressed only in terms of h. Stated results will again be taken directly from [13].

5.1 System of PIDEs for the approximate option price and hedging error

In order to end up in the setting of [8] we define:

$$c_1 := \begin{cases} \int_{\mathbb{R}} (e^y - 1) k(y) dy & \text{, if } 0 < \rho < 1 \\ 0 & \text{, otherwise.} \end{cases}$$

Now for $q > 0$ sufficiently large define:

$$\overline{H}^{\epsilon_0}(t,x) := \mathrm{e}^{-qt} H^{\epsilon_0}(x + (\sigma^2/2 + c_1)t)$$
$$\overline{V}^{\epsilon}(t,x) := \mathrm{e}^{-qt} V^{\epsilon}(t, x + (\sigma^2/2 + c_1)t)$$
$$\overline{\vartheta}^{\epsilon}(t,x) := \mathrm{e}^{-qt} \vartheta^{\epsilon}(t, x + (\sigma^2/2 + c_1)t)$$
$$\overline{J}^{\epsilon}(t,x) := \mathrm{e}^{-qt} J^{\epsilon}(t, x + (\sigma^2/2 + c_1)t).$$

The above transformations remove the drift in the case when small jumps have finite first moment, or equivalently when $0 < \rho < 1$. The removal of drift ensures the continuity of the bilinear form in the variational formulation of the PIDE (see [8] Remark 6). The exponential factor e^{-qt} ensures that the bilinear form is coercive (see [8] Equation (6a)). Define also the operator $A : D(A) \to L^1$ as:

$$\begin{aligned}
Af(x) &:= qf(x) - \Big(\int_{\mathbb{R}} (f(x+y) - f(x) - (e^y - 1)Df(x))k(y)dy + \\
&\quad + \frac{1}{2}\sigma^2 D^2 f(x) \Big), \qquad \text{for } \rho \geq 1 \\
Af(x) &:= qf(x) - \int_{\mathbb{R}} (f(x+y) - f(x))k(y)dy, \qquad \text{for } \rho < 1.
\end{aligned}$$

The operator A resembles A^X but corresponds to the transformations that we have applied and takes into consideration the fact that if $\rho < 1$ then $\sigma^2 = 0$ and we have removed the drift. Thus we end up with the following system of two PIDEs for all $(t,x) \in (0,T] \times \mathbb{R}$, where the first one is the familiar forward Kolmogorov equation for the option price subtracted by the regularized payoff (i.e. in excess to payoff form) for the fixed parameter ϵ_0:

$$\begin{aligned}
D_1(\overline{V}^{\epsilon}(t,\cdot) - \overline{H}^{\epsilon_0}(t,\cdot))(x) + A(\overline{V}^{\epsilon}(t,\cdot) - \overline{H}^{\epsilon_0}(t,\cdot))(x) &= -A^X \overline{H}^{\epsilon_0}(t,\cdot)(x), \\
(\overline{V}^{\epsilon}(0,\cdot) - H^{\epsilon_0}(\cdot))(x) &= (H^{\epsilon}(\cdot) - H^{\epsilon_0}(\cdot))(x), \\
D_1 \overline{J}^{\epsilon}(t,x) + A\overline{J}^{\epsilon}(t,\cdot)(x) &= e^{qt}\psi(\overline{V}^{\epsilon}, \overline{V}^{\epsilon})(t,x), \\
\overline{J}^{\epsilon}(0,x) &= 0, \qquad (2)
\end{aligned}$$

where for the trading strategy function we have

$$\overline{\vartheta}^{\epsilon}(t,x) = \mathrm{e}^{-2x} \frac{\Gamma(\overline{V}^{\epsilon}, \exp)}{\Gamma(\exp, \exp)}(t,x).$$

In order to unify the treatment of (2) we assume there exist functions g and u_0 such that

(G1) $\exists\lambda \in [0,\eta]$ such that $\forall\omega \in [-\lambda,\lambda]$ we have $u \in H^{p+1}_\omega$ and $g \in L^2([0,T]; H^{p+1}_\omega)$ for some $p \in \mathbb{N}_0$.

(G2) Assume there exist positive constant $\tilde{C}$ and $d > 1$ such that

$$\|D_t^k g\|_{H^{\rho/2}} \leq \tilde{C} d^k (k!), \qquad \forall t \in [0,T], \forall k \in \mathbb{N}_0.$$

Then it suffices to study the following generic PIDE:
Find $u \in C^{1,2}([0,T];\mathbb{R})$ with $u(t,\cdot) \in D_{L^1}(A)$ such that

$$\begin{aligned} \frac{d}{dt}u(t,x) + Au(t,\cdot)(x) = g(t,x) \qquad \forall (t,x) \in (0,T] \times \mathbb{R} \\ u(0,x) = u_0(x) \qquad \forall x \in \mathbb{R}. \end{aligned} \tag{3}$$

5.2 Variational formulation and unique solution for the J^ϵ PIDE

Now we want to extend the operator A, to a pseudo differential operator $\mathcal{A}$ in order to state the variational formulation of (3).
First we define the function $\hat{\Psi} : \{z \in \mathbb{C}; |\Im(z)| \leq \eta\} \to \{z \in \mathbb{C}; \Re(z) > 0\}$, which we would use later as a symbol of $\mathcal{A}$ as

$$\hat{\Psi}(z) := q - \left(\int_{\mathbb{R}} (e^{izy} - 1 - iz(e^y - 1))k(y)dy - \frac{1}{2}\sigma^2 z^2 \right), \qquad \rho \geq 1,$$
$$\hat{\Psi}(z) := q - \int_{\mathbb{R}} (e^{izy} - 1)k(y)dy, \qquad \rho < 1.$$

It turns out that this function is well-defined and satisfies certain bounds, which allow us to define $\mathcal{A}$ and even some weighted versions of it and to obtain continuity and coercivity of the resulting sesquilinear form.

Lemma 5.1. *There exist two constants $C_1, C_2 > 0$ such that for all $z \in \{y \in \mathbb{C}; |\Im(y)| \leq \eta\}$ we have*

$$|\hat{\Psi}(z)| \leq C_1(1 + |z|^2)^{\rho/2},$$
$$\Re\hat{\Psi}(z) \geq C_2(1 + |z|^2)^{\rho/2}.$$

Moreover, for all $\alpha \in \mathbb{N}_0$ there exists constant $C(\alpha)$ such that for all $z \in \{y \in \mathbb{C}; |\Im(y)| \leq \eta\}$ we have:

$$|D^\alpha \hat{\Psi}(z)| \leq C(\alpha)(1 + |z|)^{\rho - \alpha}$$

Proof. See [13] Lemma 5.1.1 □

Since $\Re\hat{\Psi}(z) > 0$ we can define by using the main branch of the complex logarithm

$$\tilde{\Psi}^r(z) := e^{r\log(\tilde{\Psi}(z))}.$$

Define also the so-called weighting operator E^ω as follows

$$\begin{aligned} E^\omega :& \mathcal{S}'_{\omega^*} \to \mathcal{S}'_{\omega^*-\omega} \\ & f \to e^{\omega \cdot} f \end{aligned}$$

No we can finally proceed with the general definition of the weighted and taken to some power version of the PDO $\mathcal{A}$.

Definition 5.2. *Take $\omega_1, \omega_2 \in [-\eta, \eta]$ such that $\omega_1+\omega_2 \in [-\eta, \eta]$ and define:*

$$H^{r,\tilde{\Psi}}_{\omega_1,\omega_2} := \left\{ f \in \mathcal{S}'_{\omega_2}; \left(\tilde{\Psi}^r(\cdot + i(\omega_1+\omega_2))\mathcal{F}f(\cdot + i\omega_2) \right) \in L^2 \right\}.$$

For $r, s \in \mathbb{R}$ define:

$$\begin{aligned} \mathcal{A}^{\omega_1,s} :& H^{r,\tilde{\Psi}}_{\omega_1,\omega_2} \to H^{r-s,\tilde{\Psi}}_{\omega_1,\omega_2} \\ & f \to E^{-\omega_2}\mathcal{F}^{-1}\left(\tilde{\Psi}^s(\cdot + i(\omega_1+\omega_2))\mathcal{F}f(\cdot + i\omega_2) \right). \end{aligned}$$

We denote $\mathcal{A}^\omega := \mathcal{A}^{\omega,1}$ and $\mathcal{A} := \mathcal{A}^{0,1}$.

By the stated properties of $\tilde{\Psi}$ the above definition is correct, i.e. the expressions in it are well-defined. The weighted generalization of the generator will be used later in the error estimates for the sparse assembly, while the fractional power generalization gives us an operator energy norm, which is equivalent to a particular Sobolev norm. Denote additionally $A^{\omega_1} := E^{\omega_1}AE^{-\omega_1}$. It turns out that $\mathcal{A}^{\omega_1} = A^{\omega_1}$ on $D_{L^1}(A)$. Formally we have:

Lemma 5.3. *Let $\omega_1, \omega_2 \in [-\eta, \eta]$ and $|\omega_1+\omega_2| \leq \eta$. Then $\mathcal{A}^{\omega_1}$ coincides with A^{ω_1} on $D_{L^1}(A)$. Furthermore, for $r \in \mathbb{R}$ and every $f_1 \in D(A) \cap H^{\rho+r}_{\omega_2}, f_2 \in H^{s\rho+r}_{\omega_2}$ we have*

$$\begin{aligned} \|A^X f_1\|_{H^r_{\omega_2}} &\leq \tilde{C}\|f_1\|_{H^{\rho+r}_{\omega_2}} \\ \|\mathcal{A}^{\omega_1,s} f_2\|_{H^r_{\omega_2}} &\sim \|f_2\|_{H^{s\rho+r}_{\omega_2}}, \end{aligned}$$

with some constant $\tilde{C}$.

Proof. See [13] Lemma 5.1.3. □

Now take $|\omega_2| \leq \lambda, |\omega_1| \leq \eta$ and $|\omega_1 + \omega_2| \leq \eta$, where λ is the constant from assumption $(G1)$. The above lemma implies that $H^{r,\tilde{\Psi}}_{\omega_1,\omega_2} = H^{r\rho}_{\omega_2}$ and this, in turn, means that $\mathcal{A}^{\omega_1}$ maps $H^{\rho/2}_{\omega_2}$ to $(H^{\rho/2}_{\omega_2})^*$. Therefore our Gelfand triple would be:

$$H^{\rho/2}_{\omega_2} \overset{d}{\hookrightarrow} L^2_{\omega_2} \overset{d}{\hookrightarrow} (H^{\rho/2}_{\omega_2})^*.$$

Define also the corresponding sesquilinear (not bilinear because functions in the domain can have complex values) form

$$a^{\omega_1}_{\omega_2}(v,w) := \langle \mathcal{A}^{\omega_1} v, w \rangle_{(H^{\rho/2}_{\omega_2})^* \times H^{\rho/2}_{\omega_2}}, \qquad \forall v, w \in H^{\rho/2}_{\omega_2},$$

and denote $a^{\omega} := a^{\omega}_0$. Our weighted variational formulation now reads: Find $u \in L^2([0,T]; H^{\rho/2}_{\omega_2}) \cap H^1([0,T]; (H^{\rho/2}_{\omega_2})^*)$ such that

$$\begin{aligned} \frac{d}{dt}(u(t), v)_{L^2_{\omega_2}} + a^{\omega_1}_{\omega_2}(u(t), v) = \langle g(t), v \rangle_{(H^{\rho/2}_{\omega_2})^* \times H^{\rho/2}_{\omega_2}}, \qquad \forall v \in H^{\rho/2}_{\omega_2} \\ u(0) = u_0. \end{aligned} \tag{4}$$

Every solution of (3) is also a solution of (4) for $\omega_1 = 0$ by extension. It can be proven that in the above formulation, the sesquilinear form $a^{\omega_1}_{\omega_2}$ is continuous and coercive (see [13] Lemma 5.1.4). Thus we have the standard a priori estimate of the solution u of (4) (see [13] Lemma 5.1.5 for details)

$$\begin{aligned} \|u\|_{L^2([0,T]; H^{\rho/2}_{\omega_2})} + \|\dot{u}\|_{L^2([0,T]; (H^{\rho/2}_{\omega_2})^*)} + \|u\|_{C([0,T]; L^2_{\omega_2})} \\ \leq C(\|u_0\|_{L^2_{\omega_2}} + \|g\|_{L^2([0,T]; (H^{\rho/2}_{\omega_2})^*)}). \end{aligned}$$

By using the properties of the approximate hedging error function J^ϵ from Corollary 4.7 we have that $J^\epsilon \in L^2([0,T]; H^{\rho/2})$. The same is true for the approximate option price function V^ϵ by Lemma 4.2. Finally, we can obtain the uniqueness of the solution of hedging error PIDE (1) by using the a priori estimate above:

Theorem 5.4. *The approximate hedging error function J^ϵ is the unique solution of the PIDE (1) in:*

$$L := \left\{ f \in C^{1,2}([0,T] \times \mathbb{R}) \cap L^2([0,T]; H^{\rho/2}); \forall t \in (0,T] : f(t,\cdot) \in D_{L^1}(A) \right\}.$$

Proof. See [13] Corollary 5.1.6. □

5.3 Localization – localized formulation

For the numerical implementation we truncate $\mathbb{R}$ to a finite interval $\Omega = (-R, R)$ for some truncation parameter $R > 0$. We would introduce some notation for $s \geq 0, k \in \mathbb{N}_0$ and $B \subset \mathbb{R}$:

$$\begin{aligned}\tilde{H}^s :=& \{u \in H^s; u|_{\mathbb{R}\setminus\Omega} = 0\},\\ \tilde{C}^k(B) :=& \{f : \mathbb{R} \to \mathbb{R}; f|_B \in C^k(B) \text{ and } f|_{\mathbb{R}\setminus\overline{B}} = 0\},\\ Y :=& \tilde{H}^{\rho/2}, \text{ with } \|\cdot\|_Y := \|\cdot\|_{H^{\rho/2}}.\end{aligned}$$

Define a truncation function $\Phi_r \in C^\infty$ for some $r > 0$ with $\Phi_r|_{(-r,r)} = 1$, supp $\Phi_r \subset (-r-\delta, r+\delta)$ and $\|\Phi_r\|_{C^{p+3}} \leq \infty$. Denote by $Y|_\Omega$ the restriction of Y to functions with domain Ω and take the Gelfand triple $Y|_\Omega \overset{d}{\hookrightarrow} L^2(\Omega) \overset{d}{\hookrightarrow} (Y|_\Omega)^*$. Consider the localized variational formulation
Find $u_R \in L^2([0,T]; Y) \cap H^1([0,T]; Y^*)$ such that

$$\begin{aligned}\frac{d}{dt}(u_R(t), v)_{L^2} + a(u_R(t), v) = \langle g(t), v\rangle_{Y^*\times Y}, \qquad \forall v \in Y\\ u_R(0) = \Phi_{R-\delta}u_0,\end{aligned}$$

where $a := a_0^0$. It is an extended version of the variational formulation corresponding to the chosen Gelfand triple, but we still have existence and uniqueness of the solution since a is continuous and coercive by Section 5.2, and the extended version is trivially solved outside Ω.
The additional error introduced by localization decays exponentially as $R \to \infty$.

Theorem 5.5. *For $|\omega| < \lambda$ and C independent of R we have the following localization error bounds:*

$$\begin{aligned}\|u(T,\cdot) - u_R(T,\cdot)\|_{L^2_\omega(\mathbb{R})} \leq C\mathrm{e}^{-(\lambda-|\omega|)R}\left(\|u_0\|_{L^2_{-\lambda,\lambda}} + \|g\|_{L^2([0,T];(H^{\rho/2}_{-\lambda,\lambda})^*)}\right)\\ \|u - u_R\|_{L^2([0,T];Y_\omega)} \leq C\mathrm{e}^{-(\lambda-|\omega|)R}\left(\|u_0\|_{L^2_{-\lambda,\lambda}} + \|g\|_{L^2([0,T];(H^{\rho/2}_{-\lambda,\lambda})^*)}\right),\end{aligned}$$

where Y_ω is the localized version of $H^{\rho/2}_\omega$.

Proof. See [13] Theorem 5.2.1. □

In order to have estimates depending on h, we usually choose $R = c_R|\log h|$ for some positive constant c_R.

5.4 Spatial semi-discretization – semi-discrete formulation. Properties of orthogonal and piecewise polynomial interpolation projections. Matrix compression and perturbed formulation. Error of the semi-discretization

In order to be consistent with the framework in [8] we define for $\omega \in [-\eta, \eta]$ the so-called Schwartz kernel $K_{\mathcal{A}^\omega} : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ as

$$K_{\mathcal{A}^\omega}(x,y) := \int_{\mathbb{R}} e^{i\xi(x-y)} \hat{\Psi}(\xi + i\omega) d\xi,$$

and from the properties of $\hat{\Psi}$ it follows (see [11] Lemma 3.0.2) that the following Calderón-Zygmund property for the Schwartz kernel holds

$$|D_x^\alpha D_y^\beta K_{\mathcal{A}^\omega}(x,y)| \le C(\alpha,\beta)|x-y|^{-1-\rho-\alpha-\beta}, \qquad \alpha,\beta \in \mathbb{N}_0, x \neq y.$$

Then by setting $K^\omega(x, x-y) := K_{\mathcal{A}^\omega}(x,y)$ we have for $u, v \in Y$

$$\langle \mathcal{A}^\omega u, v \rangle_{(H^{\rho/2})^* \times (H^{\rho/2}} = \int_{\Omega \times \Omega} K^\omega(x, x-y) u(y) \overline{v(x)} dy dx,$$

and we are in the setting of [8].We need the Calderón-Zygmund property for the matrix compression analysis. It basically means that for functions with increasing distance between their support, the duality pairing is decaying polynomially.

For the spatial semi-discretization we use linear Finite Element Method. To this end denote for $l \in \mathbb{N}$ and some $C > 0$

$$\mathcal{I}^l := \left\{ [x_k, x_{k+1}]; k = 0, \dots, C2^l - 1, x_k = -R + \frac{kR}{C2^{l-1}} \right\},$$

a partition of Ω. Let Y^l denote the space of all piecewise linear functions on $\mathcal{I}^l$ that are continuous and vanishing on the boundary of Ω. We have that $Y^l \subset H^{3/2-\delta}$.

Fix $L \in \mathbb{N}$ and define $Y_h := (Y^L, \|\cdot\|_Y)$, where $h = \frac{R}{C2^{L-1}}$ is the mesh width of the partition $\mathcal{I}^L$. Denote the orthogonal projection of Y_h^* onto Y_h by P_L, and $N = C2^L$ the number of elements in the partition $\mathcal{I}^L$. Then our semi-discrete variational formulation reads:

Find $u_{R,h} \in H^1([0,T]; Y_h)$ such that

$$\frac{d}{dt}(u_{R,h}(t), v_h)_{L^2} + a(u_{R,h}(t), v_h) = \langle g(t), v_h \rangle_{Y^* \times Y}, \qquad \forall v_h \in Y_h$$
$$u_{R,h}(0) = P_L(\Phi_{R-\delta} u_0).$$

Define now the projection $P_I : C(\Omega) \to Y_h$ to be the unique piecewise linear interpolation on Ω. In order to write this projection in explicit form we define for $i = 1, \ldots, N, j = 0, 1$:

$$
\begin{aligned}
&l_i := -R + ih,\\
&T_i := [-R + (i-1)h, -R + ih],\\
&t_i := \sup |T_i|,\\
&x_i^j := -R + (i+j)\,h,\\
&L_j(x) := \prod_{k=0,k\neq j}^{1} \frac{(k-x)}{(k-j)},\\
&q_i^j(x) := \mathbb{1}_{T_i}(x) L_j\left(\frac{1}{h}(x - x_i^0)\right),\\
&\varphi_i(x) := q_i^1(x) + q_{i+1}^0(x),\ \text{for } i < N.\ (\textit{local hat basis})
\end{aligned}
$$

Then we have the expression:

$$
P_I f(x) = \sum_{i=1}^{N-1} f(x_i^1)\varphi_i(x).
$$

Both operators P_L and P_I play an important role in the semidiscrete formulation, and we want to study their properties in order to obtain error estimates on the sparse assembly later.
Let $\Omega^* := \Omega \setminus l_i$ denote any of the resulting domains when $i = 1, \ldots, N$ and $\Omega_I := (-R+1, R-1)$. For $s \in [0, 3/2)$ and $\omega \in \mathbb{R}$ we also define a certain norm on Y_h:

$$
\|u\|_{H^{-s}_{\omega,\varphi}} := \max_i \frac{(|e^{\omega x} u, \varphi_i)|}{\|\varphi_i\|_{H^s}},
$$

and set $\|\cdot\|_{H^0_{\omega,\varphi}} := \|\cdot\|_{L^2_\omega}$. Now we are ready to give the properties of P_I.

Lemma 5.6. *For $f, g_1, g_2 \in \tilde{C}(\Omega)$ and g_2 non-vanishing in Ω*

$$
\begin{aligned}
P_I(fg_1) &= (P_I f).*(P_I g_1)\\
P_I(\frac{f}{g_2}) &= (P_I f).:(P_I g_2),
\end{aligned}
$$

where $.*$ and $.:$ denote pointwise multiplication resp. division of the components with respect to the local hat basis. Let $\omega \in \mathbb{R}$ and $f \in \tilde{H}^2$. Assume also $f \in \tilde{C}^3(\Omega^*)$.For $s \in [0, 1], t \in [1, 2]$

$$
\|(Id - P_I)f\|_{H^s_\omega} \le Ch^{t-s}\|f\|_{H^t_\omega}.
$$

For $r \in (0,1], r_1 \in (1/2, 2]$

$$\|(Id - P_I)f\|_{H^{-r}_{\omega,\varphi}} \le Ch^{r+r_1}\|f\|_{H^{r_1}_\omega}.$$

For $g \in C_b^2 \cap L^2$ *and* $m = 1, 2$

$$\begin{aligned}
&\|(Id - P_I)g\|_{L^\infty} \le C(h^m\|g\|_{C^m} + \|g\|_{L^\infty(\mathbb{R}\setminus\Omega_I)}),\\
&\|(Id - P_I)g\|_{L^2} \le C(h^t\|g\|_{H^t} + \|g\|_{L^2(\mathbb{R}\setminus\Omega_I)}).
\end{aligned}$$

Proof. See [13] Lemma 5.3.2. □

For investigating the properties of the orthogonal projection P_L we define for $s \in [0, 3/2), \omega \in \mathbb{R}$ the $(N-1)\times(N-1)$ matrices

$$\begin{aligned}
(M_s^\omega)_{(i,i')} &:= (\varphi_i, \varphi_{i'})_{H^s_\omega}\\
(D_\omega)_{(i,i')} &:= e^{\omega t_i}\delta_{ii'},
\end{aligned}$$

where $\delta_{ii'} \in \{0,1\}$ and $\delta_{ii'} = 1$ only when $i = i'$. Denote $M := M_0^0$ the so-called mass matrix.
We can now formally define $P_L : (Y_h)^* \to Y_h$ by

$$(P_L f, \varphi_i) = (f, \varphi_i) \qquad i = 1, \dots, N-1,$$

or equivalently by

$$P_L f = (\varphi_i)_i^\top M^{-1}(f, \varphi_i)_i,$$

where $(\varphi_i)_i$ denotes the vector $(\varphi_1, \dots, \varphi_{N-1})^\top$. The properties of P_L are as follows

Lemma 5.7. *Fix some non-arbitrary constant* $\alpha > 0$ *and take* ω *such that* $\omega h - \log\alpha < -\delta < 0$. *For* $s_1, t_1 \in [0, 3/2)$ *and* $f \in \tilde{H}^0$*:*

$$\|P_L f\|_{H^{s_1}_\omega} \le Ch^{-t_1-s_1}\|f\|_{H^{-t_1}_{\omega,\varphi}}.$$

The following approximation property holds for $s_1 \in [0,1], t_2 \in [1,2]$ *and* $f \in \tilde{H}^{t_2} \cap C^3(\Omega^*)$

$$\|(Id - P_L)f\|_{H^{s_2}_\omega} \le Ch^{t_2-s_2}\|f\|_{H^{t_2}_\omega}.$$

For $s_4 \in [0, 3/2), t_4 \in (1/2, 2]$ *and* $f \in \tilde{H}^{t_4} \cap C^3(\Omega^*)$

$$\|P_L f\|_{H^{s_4}_\omega} \le Ch^{-s_4}(h^{t_4}\|f\|_{H^{t_4}_\omega} + \|f\|_{L^2_\omega}).$$

Finally, for $g \in C_b^2$ and $m = 1, 2$

$$\|(Id - P_L)g\|_{L^\infty} \le C(h^m \|g\|_{C^m} + \|g\|_{L^\infty(\mathbb{R}\setminus\Omega_I)}),$$
$$\|P_L g\|_{L^\infty} \le C\|g\|_{L^\infty}.$$

Proof. See [13] Lemma 5.3.4. □

By using the local hat basis, our semi-discrete formulation can now be solved as a system of linear equations. However, the resulting matrix from the sesquilinear form a is densely populated due to the integral jump part in the generator of a Lévy process. Thus we will apply matrix compression techniques. We will choose another basis, so that the resulting matrix for a in this basis is sparse, and the loss of accuracy is acceptable. Note that we will also use the local hat basis for operations which involve the projection P_I as will be explained later.
Define the hierarchical biorthogonal wavelet basis (see also [15] 3.4.2),

$$\{\psi_j^l\}_{j,l}, \qquad l \in \mathbb{N}_0, j = 1, \dots, M^l,$$

where $M^l = \dim(Y^l) - \dim(Y^{l-1})$ as satisfying the following properties:

(W1) $Y_h = \text{span}\{\psi_j^l; l = 0, \dots, L, j = 1, \dots, M^l\}$

(W2) The diameter of the support S_j^l of ψ_j^l is bounded by $C2^{-l}$

(W3) If ψ_j^l is zero on $\partial\Omega$ and q is polynomial of degree less than 2 than $(\psi_j^l, q) = 0$

(W4) ψ_j^l can be obtained by scaling and translation from $\psi_j^{l_0}$ for $l \ge l_0$

(W5) For $v \in Y$

$$v = \sum_{l=0}^{\infty} \sum_{j=1}^{M^l} v_j^l \psi_j^l,$$

where $v_j^l = (v, \tilde{\psi}_j^l)$ with ψ_j^l the corresponding dual wavelets. The series converges in $\tilde{H}^s$ for $s \in [0, \rho/2]$

(W6) For $s \in [0, 3/2)$:

$$\sum_{l=0}^{\infty} \sum_{j=1}^{M^l} |v_j^l|^2 2^{2ls} \sim \|v\|_{H^s}^2,$$

and for $s \in (\rho/2, 2]$:

$$\sum_{l=0}^{\infty} \sum_{j=1}^{M^l} |v_j^l|^2 2^{2ls} \leq C L^{\kappa} \|v\|_{H^s}^2,$$

where $\kappa = 0$ if $s < 2$ and $\kappa = 1$ otherwise.

(W7) For $f \in \tilde{H}^t$ with $t \in [\rho/2, 2]$ and $s \in [0, \rho/2]$, the projection $Q_h : Y \to Y_h$ defined by

$$Q_h v := \sum_{l=0}^{L} \sum_{j=1}^{M^l} v_j^l \psi_j^l$$

satisfies

$$\|(Id - Q_h) f\|_{H^s} \leq C h^{t-s} \|f\|_{H^t}$$

For our purposes we define the so-called inner mother wavelet function $\psi(x)$ as the piecewise linear function vanishing outside $(0, 1)$ and taking values of $(0, -1/2, 1, -1/2, 0)$ at the points $(0, 1/4, 1/2, 3/4, 1)$. Then we define $\psi_j^l(x) := \psi(2^{l-1}x - (2j-1)2^{-2})$ for $1 \leq j \leq 2^l - 2$ and $l \geq 2$. The boundary wavelets can be obtained by using the mother wavelet $\psi_*(x)$ taking values $(0, 1, -1/2, 0)$ at $(0, 1/4, 1/2, 3/4)$ and setting $\psi_0^l(x) = \psi_*(2^{l-1}x)$, and the mother wavelet $\psi^*(x)$ taking values $(0, -1/2, 1, 0)$ at $(1/4, 1/2, 3/4, 1)$ and setting $\psi_{2^l-1}^l(x) = \psi^*(2^{l-1}x - 2^{l-1} + 1)$. By rescaling the wavelets to our domain Ω we can use them as a basis, since they satisfy all of the above properties - for $(W6)$ we refer to [4] Theorem 2.1, Remark 4.1 and [14] Proposition 4.2, while for $(W7)$ we can use $(W6)$ and apply the approximation in Lemma 5.6.

Now, for the weighted sesquilinear form in the semi-discrete setting $a_0^\omega : Y_h \times Y_h \to \mathbb{C}$ denote its corresponding stiffness matrix by $\mathbf{A}^\omega$, defined by $\mathbf{A}^\omega_{(l,j),(l',j')} = a_0^\omega(\psi_{j'}^{l'}, \psi_j^l)$. As in the local hat basis case, this matrix is densely populated. Therefore, we now substitute $\mathbf{A}^\omega$ with the matrix $\tilde{\mathbf{A}}^\omega$ defined as

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

where $\delta_{l,l'} := c_0 \max\{2^{-L+\hat{\alpha}(2L-l-l')}, 2^{-l}, 2^{-l'}\}$ for some compression factor $c_0 > 0$ and $\hat{\alpha} \in (\frac{4}{4+\rho}, 1]$. It is clear that the larger the compression factor c_0, the denser the matrix $\tilde{\mathbf{A}}^\omega$ is. The corresponding sesquilinear form is denoted by $\tilde{a}^\omega$ and for convenience we set $\tilde{a} := \tilde{a}^0, \mathbf{A} := \mathbf{A}^0, \tilde{\mathbf{A}} := \tilde{\mathbf{A}}^0$. We define the equivalent norm $\|u\|_{a^\omega} := \sqrt{a^\omega(u,u)} \sim \ \|u\|_Y$, by using that a^ω is continuous and coercive. We want coercivity and continuity also for $\tilde{a}^\omega$.

Lemma 5.8. *For large enough compression factor c_0 and $\omega \in [-\eta, \eta]$ there exist $\tilde{\alpha}, \tilde{\beta} > 0$ such that for all $v_h, u_h \in Y_h$:*

$$\begin{aligned}
&|\tilde{a}^\omega(u_h, v_h)| \leq \tilde{\alpha} \|u_h\|_{a^\omega} \|v_h\|_{a^\omega} \\
&Re\ \tilde{a}^\omega(u_h, u_h) \geq \tilde{\beta} \|u_h\|^2_{a^\omega}.
\end{aligned}$$

Proof. See [15], Proposition 3.2 □

Thus, we again have continuity and coercivity of the perturbed form $\tilde{a}$. Further we can quantify the compression effectiveness by the following lemma (see [8], Proposition 3)

Lemma 5.9. *The number of non-zero elements of $\tilde{\mathbf{A}}^\omega$ is of order $O(N \log N)$ if $\hat{\alpha} < 1$ and $O(N(\log N)^2)$ if $\hat{\alpha} = 1$ as $N \to \infty$.*

The perturbed variational formulation now reads:
Find $\tilde{u}_{R,h} \in H^1([0,T]; Y_h)$ such that

$$\begin{aligned}
\frac{d}{dt}(\tilde{u}_{R,h}(t), v_h)_{L^2} + \tilde{a}(\tilde{u}_{R,h}(t), v_h) &= \langle g(t), v_h \rangle_{Y^* \times Y}, \qquad \forall v_h \in Y_h \\
\tilde{u}_{R,h}(0) &= P_L(\Phi_{R-\delta} u_0),
\end{aligned}$$

and we still have unique solution due to Lemma 5.8.
Due to the expression for the hedging error in Theorem 3.3 and the definition of the operators Γ and ψ, we need a square integrable norm estimates in time. The following theorem makes use of the representation of the solutions of the variational and perturbed variational formulations by Duhamel's principle from Section 2.3 and shows that the error of semi-discretization is independent of time, but the price is that we must have higher order norms in the initial data and the right hand side (compare with [8] Theorem 1).

Theorem 5.10 (Error of the semi-discretization)**.** *If $\rho < 3/2$ and $\lambda R \geq (2+\rho)|\log h|$ then for $\theta \in [0,1]$ we have*

$$\begin{aligned}
\|u_R(T) - \tilde{u}_{R,h}(T)\|_{H^{\theta\rho/2}} \leq &C h^{2-\theta\rho/2} (\|u_0\|_{H^2 \cap H^\rho_{-\lambda,\lambda}} \\
&+ \|g\|_{L^\infty([0,T]; H^2 \cap H^\rho_{-\lambda,\lambda})}).
\end{aligned}$$

For $\theta \in (0,1]$ we have

$$\|u_R - \tilde{u}_{R,h}\|_{L^2([0,T];H^{(1-\delta)\theta\rho/2})} \leq Ch^{2-\theta\rho/2}(\|u_0\|_{H^2\cap H^\rho_{-\lambda,\lambda}} + \|g\|_{L^\infty([0,T];H^2\cap H^\rho_{-\lambda,\lambda})}).$$

If $u_0 = 0$ then for $\theta \in [0,1]$ we have

$$\|u_R - \tilde{u}_{R,h}\|_{L^2([0,T];H^{\theta\rho/2})} \leq Ch^{2-\theta\rho/2}\|g\|_{L^\infty([0,T];H^2\cap H^\rho_{-\lambda,\lambda})}.$$

Finally, if $\rho \geq 3/2$ and $\lambda R \geq 2|\log h|$ the above estimates hold for $\theta = 1$ and $H^{\rho/2}_{-\lambda,\lambda}$ instead of $H^\rho_{-\lambda,\lambda}$.

Proof. See [13] Theorem 5.3.10. □

5.5 Time discretization – discontinuous Galerkin (dG) formulation. Error of the dG formulation

We need to discretize in time the perturbed variational formulation and we use the discontinuous Galerkin (dG) scheme. The dG scheme takes into consideration the fact that the solution of the variational formulation is increasingly smooth as time to maturity increases and uses geometric mesh with increasing mesh width and local polynomial degree in time to maturity. More specifically, let a time mesh $0 = t_0 < t_1 < \cdots < t_{M_t} = T$ and vector of polynomial degrees $\underline{r} := (r_m)_{m=0}^{M_t}$ be given, and denote $I_m = (t_{m-1}, t_m)$, $k_m = t_m - t_{m-1}$ and $\mathscr{M} = (I_m)_{m=1}^{M_t}$. For function u in the following space

$$H^1(\mathscr{M}, Y_h) = \{v \in L^2([0,T]; Y_h) : v|_{I_m} \in H^1(I_m, Y_h), m = 1, 2, \ldots, M_t\},$$

define the one-sided limits and jump term

$$\begin{aligned} u_m^+ &:= \lim_{s\to 0^+} u(t_m + s), \qquad m = 0, 1, \ldots, M_t - 1, \\ u_m^- &:= \lim_{s\to 0^+} u(t_m - s), \qquad m = 1, 2, \ldots, M_t, \\ [\![u]\!]_m &:= u_m^+ - u_m^-, \qquad m = 1, 2, \ldots, M_t - 1. \end{aligned}$$

Let $\mathscr{P}_{r_m}(I_m, S)$ denote the space of all polynomials on I_m of degree r_m with values in some function space S and define our discrete space as:

$$\mathscr{S}^{\underline{r}}(\mathscr{M}, Y_h) := \left\{u \in L^2([0,T]; Y_h) : u|_{I_m} \in \mathscr{P}_{r_m}(I_m, Y_h), m = 1, 2, \ldots, M_t\right\}.$$

Note that $\mathscr{S}^{\underline{r}}(\mathscr{M}, Y_h) \subset H^1(\mathscr{M}, Y_h)$. Then our dG formulation reads: Find $\tilde{u}^{dG}_{R,h} \in \mathscr{S}^{\underline{r}}(\mathscr{M}, Y_h)$ such that for all $w \in \mathscr{S}^{\underline{r}}(\mathscr{M}, Y_h)$ we have

$$\tilde{B}_{dG}(\tilde{u}^{dG}_{R,h}, w) = F_{dG}(w), \tag{5}$$

where

$$\begin{aligned}\tilde{B}_{dG}(u, w) := &\sum_{m=1}^{M_t} \int_{I_m} ((\dot{u}, w) + \tilde{a}(u, w))\, dt \\ &\sum_{m=1}^{M_t} ([\![u]\!]_m, w_m^+) + (u_0^+, w_0^+),\end{aligned}$$

and

$$F_{dG}(w) := \sum_{m=1}^{M_t} \int_{I_m} \langle g(t), w(t)\rangle_{Y^* \times Y}\, dt + (P_L(\Phi_{R-\delta} u_0), w_0^+).$$

We have the following existence and uniqueness result (see [12], Proposition 2.6).

Lemma 5.11. *There exists a unique solution $\tilde{u}^{dG}_{R,h} \in \mathscr{S}^{\underline{r}}(\mathscr{M}, Y_h)$ of (5). The semi-discrete solution $\tilde{u}_{R,h}$ also solves the dG formulation and satisfies the following Galerkin orthogonality*

$$\tilde{B}_{dG}(\tilde{u}_{R,h} - \tilde{u}^{dG}_{R,h}, w) = 0$$

for all $w \in \mathscr{S}^{\underline{r}}(\mathscr{M}, Y_h)$.

We also need stability estimates, because later we use sparse assembly of the right hand side. For $v_h \in Y_h$ and $u \in \mathscr{S}^{\underline{r}}(\mathscr{M}, Y_h)$ define

$$\|v_h\|_{\tilde{a}} := \sqrt{|\tilde{a}(v_h, v_h)|} \sim \|v_h\|_Y$$
$$\|u\|^2_{dG} := \sum_{m=1} M_t \int_{I_m} \|u\|^2_{\tilde{a}} dt + \frac{1}{2}\left(\|u_0^+\|^2_{L^2} + \sum_{m=1}^{M_t - 1} \|[\![u]\!]_m\|^2_{L^2} + \|u^-_{M_t}\|^2_{L^2} \right).$$

The following lemma states the stability result.

Lemma 5.12. *The solution $\tilde{u}^{dG}_{R,h} \in \mathscr{S}^{\underline{r}}(\mathscr{M}, Y_h)$ of (5) satisfies:*

$$\|\tilde{u}^{dG}_{R,h}\|_{dG} \leq C(\|u_0\|_{L^2} + \|g\|_{L^2([0,T];(Y_h)^*)}).$$

Finally, we must take care of the error of the dG discretization. It turns out that we can derive exponential bounds with respect to the number of time mesh nodes M_t. For this purpose fix some $\sigma \in (0,1)$ and set $\mu := c_3 d|\log \sigma|$ where d is the constant from assumption $(G2)$ and c_3 is positive constant. The time mesh and the degree vector for some positive μ would then be

$$\begin{aligned} t_0 &:= 0, \qquad t_m := T\sigma^{M_t - m}, \qquad m = 1, \dots, M_t, \\ r_1 &:= 0, \qquad r_m := \lfloor \mu m \rfloor, \qquad m = 2, \dots, M_t. \end{aligned}$$

Slight adaptation from [12], Theorem 5.4 leads to the following error bound.

Lemma 5.13. *Take $\sigma \in (0,1)$ and $M_t = 2\frac{|\log h|}{|\log \sigma|}$. Then $\exists c_3 > 0$ such that:*

$$\|\tilde{u}_{R,h} - \tilde{u}_{R,h}^{dG}\|_{L^2([0,T];Y)} + \|\tilde{u}_{R,h}(T) - \tilde{u}_{R,h}^{dG}(T)\|_{L^2} \leq Cdh^2.$$

The number of spatial equations to be solved is bounded by $O(d|\log h|^2)$.

Proof. See [13] Theorem 5.4.3. □

5.6 Solution algorithm – system diagonalization and preconditioning, GMRES method and its error

This section is almost identical to [8], Section 4. Let L_j be the j-th Legendre polynomial on $(-1,1)$ normalized such that $L_j(1) = 1$ and define for $m = 1, \dots, M_t$

$$\left(\Phi_j := \sqrt{j + 1/2}\, L_j\right)_{j=0,\dots,r_m}, \qquad \text{with } \|\Phi_j\|_{L^2(-1,1)} = 1.$$

This is a basis of the polynomial space $\mathscr{P}_{r_m}(-1,1)$ which after rescaling of the domain will be used as basis at the m-th time step. For this purpose define the following interval mapping:

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

Now set $\tilde{u}_{R,h,m}^{dG} := \tilde{u}_{R,h}^{dG}|_{I_m}$ and $w_m = w|_{I_m}$ for $w \in \mathscr{S}^{\underline{r}}(\mathscr{M}, Y_h)$. Then we have the following basis representation on the time interval I_m of the dG solution and any other function in the solution space:

$$\begin{aligned} \tilde{u}_{R,h,m}^{dG}(t,x) &= \sum_{j=0}^{r_m} \tilde{u}_{R,h,m,j}^{dG}(x)(\Phi_j \circ F_m^{-1})(t) \\ w_m(t,x) &= \sum_{j=0}^{r_m} w_{m,j}(x)(\Phi_j \circ F_m^{-1})(t), \end{aligned}$$

where $\tilde{u}^{dG}_{R,h,m,j}, w_{m,j} \in Y_h$. Having fully discretized the perturbed variational formulation and fixed a basis with respect to time, we get the discrete formulation as:
For every $m = 1, \dots, M_t$, find $(\tilde{u}^{dG}_{R,h,m,j})_{j=0}^{r_m} \in (Y_h)^{r_m+1}$ such that for all $(w_{m,i})_{i=0}^{r_m} \in (Y_h)^{r_m+1}$ we have,

$$\sum_{i,j=0}^{r_m} C_{ij}(\tilde{u}^{dG}_{R,h,m,j}, w_{m,i}) + \frac{k_m}{2}\sum_{i=0}^{r_m} \tilde{a}(\tilde{u}^{dG}_{R,h,m,j}, w_{m,i}) = \sum_{i=0}^{r_m} f_{m,i}(w_{m,i}),$$

where for $i, j = 0, \dots, r_m$ we have

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

and

$$f_{m,i}(v) = \int_{I_m} \langle g(t), w_{m,i}\rangle_{Y^* \times Y} (\Phi_i \circ F_m^{-1})(t)dt + \Phi_i(-1)(\tilde{u}^{dG-}_{R,h,m-1}(t_{m-1}), v),$$

where we set $\tilde{u}^{dG-}_{R,h,0}(0) = P_L(\Phi_{R-\delta} u_0)$.
We fix now some $m = 1, \dots, M_t$ and drop the subscript for short when needed. Let C_{ij} form the matrix $\mathbf{C}$ and $\mathbf{M}$ the mass matrix of the wavelet basis. Then at time step m, with $\otimes$ denoting the Kronecker product, we must solve the following linear system:

$$\left(\mathbf{C} \otimes \mathbf{M} + \frac{k}{2}\mathbf{I} \otimes \tilde{\mathbf{A}}\right) \underline{u} = \underline{f}, \tag{6}$$

where $\underline{u}$ denotes the coefficient vector of $\tilde{u}^{dG}_{R,h,m} \in \mathscr{P}_{r_m}(I_m, Y_h)$ and the vector $\underline{f}$ is the right hand side evaluated at the wavelet basis of $(Y_h)^{r_m+1}$ obtained by taking all (r_m+1)-component vectors with r_m zero components and one element of the wavelet basis of Y_h.
If we set $\tilde{N} := \dim Y_h$, the above is a linear system of size $(r+1)\tilde{N}$ and we can decouple it into $(r+1)$ linear systems of size $\tilde{N}$. Indeed, let $\mathbf{C} = \mathbf{Q}\mathbf{T}\mathbf{Q}^H$ be the Schur decomposition with unitary $\mathbf{Q}$ and upper triangular $\mathbf{T}$, containing on the main diagonal the eigenvalues $\lambda_1, \dots, \lambda_{r+1}$ of $\mathbf{C}$. Multiply (6) from the left by $\mathbf{Q}^H \otimes \mathbf{I}$ to get:

$$\left(\mathbf{T} \otimes \mathbf{M} + \frac{k}{2}\mathbf{I} \otimes \tilde{\mathbf{A}}\right) \underline{w} = \underline{g}, \text{ with } \underline{w} = (\mathbf{Q}^H \otimes \mathbf{I})\underline{u}, \underline{g} = (\mathbf{Q}^H \otimes \mathbf{I})\underline{f}. \tag{7}$$

By denoting $\underline{w} = (\underline{w}_0, \dots, \underline{w}_r), \underline{j} \in \mathbb{C}^{\tilde{N}}$ we can solve (7) by iteratively solving

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

for $j = r, \ldots, 0$ where $\underline{s}_j = \underline{g}_j - \sum_{l=j+1}^{r} \mathbf{T}_{j+1,l+1} \mathbf{M} \underline{w}_j$.
We can also precondition the system at each step with the matrix:

$$\mathbf{S} = \left(\mathrm{Re}(\lambda_j)\mathbf{I} + \frac{k}{2}\mathbf{D} \right)^{1/2},$$

where $\mathbf{D}_{(i,l),(i,l)} = 2^{l\rho/2}$ is a diagonal matrix. Thus we must iteratively solve for $j = r, \ldots, 0$

$$\mathbf{S}^{-1} \left(\lambda_{j+1}\mathbf{M} + \frac{k}{2}\tilde{\mathbf{A}} \right) \mathbf{S}^{-1}(\mathbf{S}\underline{w}_j) = \mathbf{S}^{-1}\underline{s}_j,$$

with unknown $\mathbf{S}\underline{w}_j$ via n_G incomplete GMRES(m_0) iterations (restarted every m_0 iterations). We end up with an approximate solution $\tilde{u}_{R,h}^{dG,GMRes}$ for which the following error bound holds true (see [9], Section 5.5.4).

Theorem 5.14. *Take sufficiently large constants c_0, c_2 and c_3 and set $n_G = c_0 d^{4+\delta} |\log h|^5$, $M_t = c_2 |\log h|$, $\mu = c_3 d$. Then we have*

$$\|\tilde{u}_{R,h}^{dG}(T) - \tilde{u}_{R,h}^{dG,GMRes}(T)\|_{L^2} + \|\tilde{u}_{R,h}^{dG} - \tilde{u}_{R,h}^{dG,GMRes}\|_{L^2([0,T];Y)} \leq \\ \leq Ch^{p+1}(\|u_0\|_{L^2} + \|g\|_{L^2([0,T];Y_h^*)}).$$

The overall number of computation steps is bounded by $O(d^{6+\delta} N (\log N)^8)$.

5.7 Assembly of the right hand side. Sparse assembly of Ψ, representation of Γ via the infinitesimal generator. Wavelet transform, approximative operators and error of the sparse assembly

In the previous section we were finally able to describe a procedure for the assembly of the left hand side of the variational formulation of the PIDE for both the approximate option price and hedging error functions V^ϵ and J^ϵ. In this section, we would concentrate on the problem of computing the corresponding right hand sides of the equations. Moreover we assume $\sigma^2 = 0$ because the theoretical error estimates are valid only in this case.
First, notice that, up to basis transformations, the right hand side of the equation for $\overline{V}^\epsilon$ is reconstructed from the components $(A\overline{H}^{\epsilon_0}, \varphi_i)_i$. If we are working in the wavelet basis, then we need the inverse wavelet transform which results in $O(N)$ additional computation steps and does not have impact on the overall computation as can be seen from Theorem 5.14.
In the equation for the approximate hedging error $\overline{J}^\epsilon$, however, the right

hand side is given again up to basis transform by $(\psi(\overline{V}^\epsilon, \overline{V}^\epsilon), \varphi_i)_i$ where, instead of $\overline{V}^\epsilon$ we use $\tilde{\overline{V}}_{R,h}^{\epsilon,dG,GMRes}$, which is given as a coordinate vector with respect to the chosen basis. A naive approach, would be to compute the right hand side with respect to all basis components, but this would result in $O(N^3)$ computational steps. Therefore we use the special properties of Γ in order to come up with reasonable approximation of the right hand side. In [2] it was shown, that Γ can be represented in terms of the generator on a suitable restricted domain. Specifically, define:

$$D^2(\Gamma) := \{(f,g) \in D(\Gamma); f, g, fg \in D(A)\}\,,.$$

We have the following representation for $(f,g) \in D^2(\Gamma)$:

$$\Gamma(f,g) = A^X(fg) - fA^X g - gA^X f. \tag{8}$$

Indeed, from the definitions of A^X and Γ we have:

$$\begin{aligned}
(A^X(fg))(x) &:= \int_{\mathbb{R}} (f(x+y)g(x+y) - f(x)g(x) - (e^y-1)g(x)Df(x) \\
&\quad - (e^y-1)f(x)Dg(x))k(y)dy + \frac{1}{2}\sigma^2(g(x)D^2f(x) \\
&\quad + f(x)D^2g(x) + 2Df(x)Dg(x) - g(x)Df(x) - f(x)Dg(x)) \\
(f(A^X g))(x) &:= \int_{\mathbb{R}} (f(x)g(x+y) - f(x)g(x) - (e^y-1)f(x)Dg(x))k(y)dy \\
&\quad + \frac{1}{2}\sigma^2(f(x)D^2g(x) - f(x)Dg(x)) \\
(g(A^X f))(x) &:= \int_{\mathbb{R}} (g(x)f(x+y) - g(x)f(x) - (e^y-1)g(x)Df(x))k(y)dy \\
&\quad + \frac{1}{2}\sigma^2(g(x)D^2f(x) - g(x)Df(x)) \\
(A^X(fg))(x) - (f(A^X g))(x) - (g(A^X f))(x) &= \sigma^2 Df(x)Dg(x) \\
&\quad + \int_{\mathbb{R}} (f(x+y) - f(x))(g(x+y) - g(x))F(dy) \\
&= \Gamma(f,g)(x).
\end{aligned}$$

Note that the operators A^X and A have dense corresponding matrices due to the same reason - the integral part which corresponds to the jump part of the Lévy process X. The same line of reasoning used for introducing the sparse approximation matrix $\tilde{\mathbf{A}}$ of A in the perturbed variational formulation can be used to introduce sparse approximation of A^X. It can also be used to derive sparse approximation matrix $\tilde{\mathbf{A}}^\omega$ for the operator A^ω corresponding to the sesquilinear form a_0^ω. Thus we implicitly change the notation for short and identify A^X to be A for the rest of this section, since we are interested only in the jump part of the generator. It seems natural to substitute directly into the expression for Γ the already introduced sparse approximation for the generator. However, for the ensuing analysis we will further introduce the approximative operator A_d^ω. This operator is defined through the interpolation operator P_I, the approximate orthogonal projection $\tilde{P}_L$ introduced below and the already introduced sparse approximation operator $\tilde{A}^\omega$ corresponding to the matrix $\tilde{\mathbf{A}}^\omega$. It has good analytical properties and results in right hand side assembly done in $O(N(\log N))$ computation steps. Define first the approximate orthogonal projection operator $\tilde{P}_L$ as follows:

$$\tilde{P}_L := (\varphi_i)_i^\top B(M^{-1}, \gamma_\delta)(f, \varphi_i)_i,$$

where

$$\gamma_\delta = \left\lceil \frac{2|\log(1 - e^{-\delta})|}{\delta} \right\rceil,$$

and $B(M^{-1}, \gamma_\delta)$ is the matrix consisting of the main diagonal, γ_δ lower and γ_δ upper diagonals of M^{-1}. This is a sparse approximation of the orthogonal projection P_L because, in general, even if M is sparse, M^{-1} is not.
In order to exploit the multiplicative property of P_I we need to work in the local hat basis (φ_j^l). Thus we need not only the inverse wavelet transform, but also an efficient wavelet transform. For both coordinate transform directions we apply the multiscale wavelet transform described below.
Denote φ^l and ψ^l the local hat and wavelet bases column vectors respectively, for the partition $\mathcal{I}^l$. Then there exists a matrix $\mathbf{T} \in \mathbb{R}^{\dim(Y^{l+1}) \times \dim(Y^{l+1})}$ such that:

$$\mathbf{T}\varphi^{l+1} = P \begin{pmatrix} \varphi^l \\ \psi^l \end{pmatrix},$$

where P sorts the elements of a vector increasingly according to the infimum of the support of the its component functions. Specifically, P sorts both bases as in the vector $(\psi_0^l, \varphi_1^l, \psi_1^l, \dots, \varphi_{2^l-1}^l, \psi_{2^l-1}^l)$ and the matrix $\mathbf{T}$ is given by

$$\mathbf{T} = \begin{pmatrix} 1 & -1/2 & & & & \\ 1/2 & 1 & 1/2 & & & \\ & -1/2 & 1 & -1/2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1/2 & 1 & 1/2 \\ & & & & -1/2 & 1 \end{pmatrix}.$$

Take now $f \in Y_h$ and denote the coordinate vectors of f with respect to local hat and wavelet bases φ^l and ψ^l respectively by c^l and d^l and notice the trivial $c^0 := d^0$ since wavelet and hat bases coincide on partition of 1 interval. Then we have:

$$\mathbf{T}P\begin{pmatrix} c^l \\ d^l \end{pmatrix} = c^{l+1}. \tag{9}$$

Then the inverse wavelet transform is just finding c^{l+1} with given $d^0, d^1, \ldots, d^l$ which results in matrix vector multiplication with the tridiagonal matrix T and results in $O(N)$ steps. For the wavelet transform given c^{l+1} we must solve (9) for c^k and d^k for $k = 0, \ldots, l$ which can be done using band matrix solvers (see [5] Section 4.3) and results also in $O(N)$ steps.

We will need integrability in space for the argument functions of the sparse approximation of Γ so that we are able to derive error estimates. Therefore we again use the idea of subtracting a smooth approximation functions from its arguments. The terms which involve only the smooth approximation functions can be precomputed once using the operator Γ. To this end define the following spaces for $\omega \in \mathbb{R}$ and $r \geq 0$:

$$\begin{aligned} D_{\mathbb{R}^\omega} &:= \left\{ e^{\omega x} f \in C_b^{p+3} \cap D(A) \right\}, \\ D_\Omega^r &:= \left\{ f \in H^r \cap H^{1/2+\delta} \cap \tilde{C}^{p+2}(\Omega^*); f|_{\mathbb{R}\setminus(-R/2-\delta, R/2+\delta)} = 0 \right\}, \end{aligned}$$

and set $D_\Omega := D_\Omega^0$. The smooth functions that we subtract from the original arguments will lie in $D_{\mathbb{R}^\omega}$ while the resulting differences will be in D_Ω^r. Finally, the approximative operator $A_d^\omega : D_\Omega \to Y_h$ is defined as:

$$A_d^\omega f := E^{-\omega} \tilde{P}_L \tilde{A}^\omega P_I(E^\omega f),$$

where $\tilde{A}^\omega$ is the operator with corresponding matrix $\tilde{\mathbf{A}}^\omega$ in the wavelet basis, and all other operators on the right hand side are taken with respect to the local hat basis. From the definition, it can be seen that computing the right hand side needs basis transformations ($O(N)$), pointwise vector multiplication ($O(N)$) and multiplication of sparse matrices $\tilde{\mathbf{A}}^\omega$ and $B(M^{-1}, \gamma_\delta)$ with vectors and assembly ($O(N(\log N))$). Our ultimate goal is deriving error bounds for the sparse assembly of the right hand side. For shorter notation, define for $r, s \geq 0$:

$$\|f\|^{A_d}_{(r,s,\omega)} := h^{-1/2-\nu}(\|f\|_{H^r_\omega} + h^s \|f\|_{H^{r+s}_\omega}).$$

Assume also that $\omega, \omega_f, \omega - \omega_f \in (-\eta, \eta)$ and for $\omega^* = |\omega| \vee |\omega_f| \vee |\omega - \omega_f|$ we set $c_R \geq 2\frac{3/2-\nu}{\eta-\omega^*}$. On the space

$$D(\hat{\Gamma}_d) := \{(f,g) \in D_\Omega \times D_\Omega; fg \in D_\Omega\},$$

define the approximative carré-du-champ operator $\hat{\Gamma}^{\omega_f,\omega}_d : D(\hat{\Gamma}_d) \to Y_h$ as:

$$\hat{\Gamma}^{\omega_f,\omega}_d(f,g) := A^\omega_d(fg) - fA^{\omega-\omega_f}_d g - gA^{\omega_f}_d f.$$

Notice that the above definition is similar to the representation (8). In order to compare Γ and $\hat{\Gamma}^{\omega_f,\omega}_d$, we define an intersection of their domains depending on $r \geq 0$:

$$D^{\hat{\Gamma}^d}_r := \{(f,g) \in D^2(\Gamma) \cap (D^r_\Omega \times D^r_\Omega); fg \in D^r_\Omega\},$$

and for notational convenience introduce the norm estimate for $r, s \geq 0$

$$\|f,g\|^{\hat{\Gamma}^d}_{(r,s,\omega_f,\omega)} := \|f\|^{A_d}_{(r,s,\omega_f)} \|g\|_{H^{1/2+\delta}_{\omega-\omega_f}} + \|g\|^{A_d}_{(r,s,\omega-\omega_f)} \|f\|_{H^{1/2+\delta}_{\omega_f}}.$$

Now we can define the approximation of Γ that we would use in the numerical implementation and that represents the idea of subtracting smooth functions from the argument functions. To this end for

$$D(\Gamma_d) := \Big\{(f, \tilde{f}, g, \tilde{g}) \in \bigcup_{\omega_1,\omega_2 \in [-\eta,\eta]} (C(\mathbb{R}) \times D^{\omega_1}_{\mathbb{R}} \times C(\mathbb{R}) \times D^{\omega_2-\omega_1}_{\mathbb{R}};$$
$$f - \tilde{f}, g - \tilde{g} \in D_\Omega)\Big\},$$

define $\Gamma^{\omega_f,\omega}_d : D(\Gamma_d) \to C_b$ as

$$\begin{aligned}\Gamma^{\omega_f,\omega}_d(f,\tilde{f},g,\tilde{g}) :=& \hat{\Gamma}^{\omega_f,\omega}_d(f-\tilde{f}, g-\tilde{g}) \\ &+ \Big(A^\omega_d(\tilde{f}(g-\tilde{g})) - \tilde{f}A^{\omega-\omega_f}_d(g-\tilde{g}) - (g-\tilde{g})A\tilde{f}\Big) \\ &+ \Big(A^\omega_d(\tilde{g}(f-\tilde{f})) - \tilde{g}A^{\omega_f}_d(f-\tilde{f}) - (f-\tilde{f})A\tilde{g}\Big) + \Gamma(\tilde{f},\tilde{g}).\end{aligned}$$

We again define the respective intersection of domains of Γ and $\Gamma_d^{\omega_f,\omega}$ for $r \geq 0$ as:

$$D^{\Gamma_d}_{r,\omega_f,\omega} := \Big\{ (f,\tilde{f},g,\tilde{g}) \in C(\mathbb{R}) \times D^{\omega_f}_{\mathbb{R}} \times C(\mathbb{R}) \times D^{\omega-\omega_f}_{\mathbb{R}}; (f-\tilde{f}, g-\tilde{g}) \in D^r_\Omega \\ \text{and } (f,g),(\tilde{f},g),(f,\tilde{g}),(\tilde{f},\tilde{g}) \in D^2(\Gamma) \Big\},$$

and for shorter notation the norm estimate for $r,s \geq 0$

$$\|f,\tilde{f},g,\tilde{g}\|^{\Gamma_d}_{(r,s,\omega_f,\omega)} := \|f-\tilde{f},g-\tilde{g}\|^{\hat{\Gamma}_d}_{(r,s,\omega_f,\omega)} + h^{-\delta}\|g-\tilde{g}\|^{A_d}_{(r,s,\omega-\omega_f)}\|\tilde{f}\|_{C^{1+p}_{\omega_f}} \\ + h^{-\delta}\|f-\tilde{f}\|^{A_d}_{(r,s,\omega_f)}\|\tilde{g}\|_{C^{1+p}_{\omega-\omega_f}}.$$

The arguments of $\Gamma_d^{\omega_f,\omega}$ are both original arguments of Γ as well as the smooth approximations that we subtract from them. As mentioned, and as can be seen from the definition, we use the operator Γ and the generator A for computing the terms involving only the smooth approximation functions. Finally we can define the approximative operator $\psi_d^{\omega_f,\omega} : D(\Gamma_d) \to C_b$ as

$$\psi_d^{\omega_f,\omega}(f,\tilde{f},g,\tilde{g}) := \Gamma_d^{\omega_f,\omega}(f,\tilde{f},g,\tilde{g}) - \frac{1}{ce^{2x}}\Gamma_d^{\omega_f,\omega_f-1}(f,\tilde{f},\exp,\exp)\times \\ \times \Gamma_d^{\omega-\omega_f,\omega-\omega_f-1}(g,\tilde{g},\exp,\exp),$$

and the corresponding notation of the norm estimate for $r,s \geq 0$

$$\|f,\tilde{f},g,\tilde{g}\|^{\psi_d}_{(r,s,\omega_f,\omega)} := \|f,\tilde{f},g,\tilde{g}\|^{\Gamma_d}_{(r,s,\omega_f,\omega)} + \|f-\tilde{f}\|^{A_d}_{(r,s,\omega_f)}\|g,\tilde{g}\|^{\omega}_{(1/2+\delta,\omega-\omega_f)} \\ + \|g-\tilde{g}\|^{A_d}_{(r,s,\omega-\omega_f)}(h^{1/2+\nu}\|f-\tilde{f}\|^{A_d}_{(1/2+\nu+\delta,\nu+1/2,\omega_f)} \\ + \|f,\tilde{f}\|^{\omega}_{(1/2+\delta,\omega_f)}).$$

Notice that in the expression for $\psi_d^{\omega_f,\omega}$ the terms $\Gamma_d^{\omega_f,\omega_f-1}(f,\tilde{f},\exp,\exp)$ and $\Gamma_d^{\omega-\omega_f,\omega-\omega_f-1}(g,\tilde{g},\exp,\exp)$ are well-defined because $\exp \in D^{-1}_{\mathbb{R}}$ and therefore $(f,\tilde{f},\exp,\exp),(g,\tilde{g},\exp,\exp) \in D(\Gamma_d)$. In our actual implementation we substitute $\psi(f,g)$ with $P_I\psi_d^{0,0}(f,\tilde{f},g,\tilde{g})$ - we use weighting only for theoretical purposes. The main result for the error and stability of the sparse assembly follows.

Theorem 5.15. *For $m, m_1 = 1,2$,$(f,\tilde{f},g,\tilde{g}) \in D^{\Gamma_d}_{(m+\nu+1/2,0,0)}$ and $\psi(f,g) \in L^\infty$ we have*

$$\|\psi(f,g) - P_I\psi_d^{0,0}(f,\tilde{f},g,\tilde{g})\|_{Y_h^*} \\ \leq Ch^m\|f,\tilde{f},g,\tilde{g}\|^{\psi_d}_{(m,\nu+1/2,0,0)} + h^{m_1}\|\psi(f,g)\|_{C^{m_1}} + \|\psi(f,g)\|_{L^\infty(\mathbb{R}\setminus\Omega_I)}.$$

Moreover, for $(f,\tilde{f},g,\tilde{g}) \in D(\Gamma_d)$ *we have*

$$\begin{aligned}
&\|P_I\psi_d^{0,0}(f,\tilde{f},g,\tilde{g})\|_{Y_h^*} \\
&\quad \le Ch^{-\delta}\Big(\|f,\tilde{f},g,\tilde{g}\|^{\Gamma_d}_{(m,\delta+1/2,0,0)} + \|g-\tilde{g}\|_{H^{1/2+\delta}}\|A\tilde{f}\|_{L^2} \\
&\quad + \|f-\tilde{f}\|_{H^{1/2+\delta}}\|A\tilde{g}\|_{L^2} + \|\Gamma(\tilde{f},\tilde{g})\|_{H^{1/2+\delta}} \\
&\quad + (\|f-\tilde{f}\|^{A_d}_{(0,1/2+\delta,0)} + \|f,\tilde{f}\|^{\omega}_{(1/2+\delta,0)})(\|g-\tilde{g}\|^{A_d}_{(0,1/2+\delta,0)} + \|g,\tilde{g}\|^{\omega}_{(1/2+\delta,0)})\Big)
\end{aligned}$$

Proof. See [13] Theorem 5.6.8. □

5.8 Final variational formulation. Overall error of the option price, optimal strategy and hedging error

Taking into consideration the right hand side we now have the following variational formulation for the option price:
Find $\tilde{u}^{\epsilon,dG}_{R,h} \in \mathscr{S}^{\underline{r}}(\mathscr{M},Y_h)$ such that for all $w \in \mathscr{S}^{\underline{r}}(\mathscr{M},Y_h)$ we have

$$\tilde{B}_{dG}(\tilde{u}^{\epsilon,dG}_{R,h},w) = \int_0^T \langle A\overline{H}^{\epsilon_0}(t),w(t)\rangle_{Y^*\times Y}dt + (P_L(\Phi_{R-\delta}(H^{\epsilon}-H^{\epsilon_0})),w_0^+), \tag{10}$$

with the respective solution computed by the GMRES method denoted by $\tilde{u}^{\epsilon,dG,GMRes}_{R,h}$. Thus, the approximation of the option price function is given by:

$$\tilde{\tilde{V}} := \tilde{u}^{\epsilon,dG,GMRes}_{R,h} + \overline{H}^{\epsilon_0}.$$

For the application of $\psi_d^{0,0}$ we need its arguments to be functions in D_Ω but $\tilde{\tilde{V}} - \overline{H}^{\epsilon_0} \notin D_\Omega$ because it has larger support. Therefore for $r>0$ we introduce

$$\tilde{\tilde{V}}_r := \Phi_r\tilde{u}^{\epsilon,dG,GMRes}_{R,h} + \overline{H}^{\epsilon_0},$$

and we note that $\tilde{\tilde{V}}_R = \tilde{\tilde{V}}$. The approximate trading strategy is now given by

$$\tilde{\tilde{\vartheta}} = P_I\left(\frac{1}{ce^{2x}}\Gamma_d^{0,-1}(\tilde{\tilde{V}}_{R/2}-\overline{H}^{\epsilon_0},0,\exp,\exp)\right) + \frac{1}{ce^{2x}}\Gamma(\overline{H}^{\epsilon_0},\exp).$$

The hedging error variational formulation is:
Find $\tilde{\tilde{J}}^{\epsilon,dG,\Delta}_{R,h} \in \mathscr{S}^{\underline{r}}(\mathscr{M},Y_h)$ such that for all $w \in \mathscr{S}^{\underline{r}}(\mathscr{M},Y_h)$ we have

$$\tilde{B}_{dG}(\tilde{\tilde{J}}^{\epsilon,dG,\Delta}_{R,h},w) = \int_0^T \langle P_I\psi_d^{0,0}(\tilde{\tilde{V}}_{R/2},\overline{H}^{\epsilon_0},\tilde{\tilde{V}}_{R/2},\overline{H}^{\epsilon_0})(t),w(t)\rangle_{Y^*\times Y}dt, \tag{11}$$

and is well-defined due to $(\tilde{\tilde{V}}_{R/2}, \overline{H}^{\epsilon_0}, \tilde{\tilde{V}}_{R/2}, \overline{H}^{\epsilon_0}) \in D(\Gamma_d)$. Denote the approximate solution of (11) via the GMRES method by $\tilde{\tilde{J}}$ and introduce the following indicator function:

$$\kappa := \begin{cases} 1 & , \rho \geq 3/2 \\ 0 & \text{otherwise.} \end{cases}$$

We have the following result on the approximate solutions of the variance-optimal hedging problem where we have error bounds depending only on h, or equivalently on the number of computation steps.

Theorem 5.16. *Undoing the transformations that we have previously applied, the approximate solutions are given by*

$$\begin{aligned} \tilde{V}(T, x) &:= Ke^{qT} \tilde{\tilde{V}}(T, x - (\sigma^2/2 + c_1)T + \log K), \\ \tilde{\vartheta}(t, x) &:= Ke^{qt} \tilde{\tilde{\vartheta}}(t, x - (\sigma^2/2 + c_1)t + \log K) \\ \tilde{J}(T, x) &:= Ke^{qT} \tilde{\tilde{J}}(T, x - (\sigma^2/2 + c_1)T + \log K). \end{aligned}$$

Letting $\epsilon := Ch^s$, *denoting* M *the number of computation steps and defining the following constants*

$$\begin{aligned} s^V &:= 1, \\ s^\vartheta &:= \frac{3/2 - (1 + 1/2\kappa)\nu - c_R/2 - 2\delta}{2}, \\ s^J &:= \frac{1 - (1 + 1/2\kappa)\nu - \delta}{2}, \end{aligned}$$

we have the error bounds:

$$\begin{aligned} \|V(T) - \tilde{V}(T)\|_{L^2} &\leq CM^{(-1+\delta)3/2s^V}, \\ \|\vartheta - \tilde{\vartheta}\|_{L^2([0,T];L^2)} &\leq CM^{(-1+\delta)3/2s^\vartheta}, \\ \|J(T) - \tilde{J}(T)\|_{L^2} &\leq CM^{-\frac{3/2s^J}{1+(6+\delta)\nu s+\delta}} \end{aligned}$$

Proof. See [13] Corollary 5.7.7. □

5.9 Implementation issues – computation of certain Γ expressions and assembly of the approximative operators through functions of the kernel. Implementation for the CGMY kernel

We still have to deal with the terms that we have subtracted from the right hand side of both PIDE equations in order to have numerical tractability, i.e. terms involving $\overline{H}^{\epsilon_0}$, or equivalently just H^{ϵ_0}, because the operators of interest do not act on the time variable. Indeed, we must compute $\Gamma(H^{\epsilon_0}, H^{\epsilon_0}), \Gamma(H^{\epsilon_0}, \exp)$ and assemble the sparse stiffness matrix $\tilde{\mathbf{A}}^{\omega}$ of the wavelet basis, where the weighted version at $\omega = -1$ is needed for the approximative operator $\psi_d^{0,0}$. Notice that the terms which are not straightforward to compute correspond to the jump part of these operators and matrices, so we are interested only in them. For that purpose assume $\sigma^2 = 0$ and ignore the part of A corresponding to coercivity ensuring parameter q. The implementation would be devised, so that it is easy to implement anew when we change the distributional assumptions.

We introduce now the functions that will contain the dependance on the distributional assumptions (i.e. the kernel) and with the help of which, we will compute the right hand side expressions. They are the so-called antiderivatives of the kernel function $k(x)$ corresponding to the Lévy measure $F(x)$ of the underlying process X. For $k^{(0)}(x) := k(x)$ and $i \in \mathbb{N}$

$$k^{(-i)}(x) := \begin{cases} -\int_x^\infty k^{(-i+1)}(y)dy & , x > 0 \\ \int_{-\infty}^x k^{(-i+1)}(y)dy & , x < 0 \end{cases}$$

By assumption $(A4)$ we get the same exponential bound as $z \to +-\infty$ for the antiderivatives as well as the derivatives. Denote also the weighted function $k_\omega^{(-2)}(x) := \mathrm{e}^{\omega x} k^{(-2)}(x)$.

In the computations, we consider a regularized payoff function of the form

$$H^{\epsilon_0}(x) := (1 - e^x) 1_{(-\infty, -\epsilon_0)}(x) + \tilde{q}(x) 1_{[-\epsilon_0, \epsilon_0]}(x),$$

where $\epsilon_0 > 0$ and $\tilde{q}$ is polynomial such that H^{ϵ_0} is at least in $C^2(\mathbb{R})$. We need to compute

$$\Gamma(f, g) = A(fg) - fAg - gAf.$$

We assume now that we have applied the removal of drift even in the case $\rho \geq 1$. This does not change our results up until now, since we never assumed $\rho \geq 1$. However this simplifies case study because in this case the jump part of the operator A can be written in the following general form, by subtracting from A^X the expression for the drift:

$$Af(x) = \int_{\mathbb{R}} (f(x+y) - f(x) - yf'(x)) k(y) dy.$$

Therefore, it is sufficient to compute Ag where

$$g(x) = e^{mx}1_{(-\infty,-\epsilon_0)}(x) \text{ or } g(x) = e^{mx}p(x)1_{[-\epsilon_0,\epsilon_0]}(x),$$

for $m \in \mathbb{N}, m \leq \eta$ (so that Ag is finite) and p some polynomial. H^{ϵ_0} is sufficiently smooth, so that we can compute the integrals involved in Ag only on the respective domains, without worrying about jumps at $-\epsilon_0$ due to taking derivatives of indicator function. Notice that for any function $f \in \tilde{H}^2$ we have by twice integrating by parts

$$Af(x) = \int_0^\infty (f(x+y) - f(x) - yf'(x))k(y)dy = \int_0^\infty f''(x+y)k^{(-2)}(y)dy$$
$$Af(x) = \int_{-\infty}^0 (f(x+y) - f(x) - yf'(x))k(y)dy = \int_{-\infty}^0 f''(x+y)k^{(-2)}(y)dy,$$

and thus

$$Af(x) = \int_{\mathbb{R}} (f(x+y) - f(x) - yf'(x))k(y)dy = \int_{\mathbb{R}} f''(x+y)k^{(-2)}(y)dy. \tag{12}$$

Ignoring derivative jump terms as explained above, for $g(x) = e^{mx}1_{(-\infty,-\epsilon_0)}(x)$ we obtain

$$Ag(x) = \int_{-\infty}^{-\epsilon_0} (\mathrm{e}^{m(x+y)} - \mathrm{e}^{mx} - my\mathrm{e}^{mx})k(y)dy, \tag{13}$$

and if $g(x) = e^{mx}p(x)1_{[-\epsilon_0,\epsilon_0]}(x)$

$$Ag(x) = \mathrm{e}^{mx}\int_{-\epsilon_0}^{\epsilon_0} k_m^{(-2)}(y)(m^2p(x+y) + 2mp'(x+y) + p''(x+y))dy. \tag{14}$$

We can work out an explicit formula for the integral (13) as explained at the end of this section, while for (14) we choose a gaussian quadrature rule for simplicity and because the integral is on a finite interval.

Now, we are interested in assembling the sparse stiffness matrix $\tilde{\mathbf{A}}^\omega$ or equivalently the computation of the terms $a^\omega(\psi_i^l, \psi_{i'}^{l'})$ for the wavelet basis. Our approach is to compute $a^\omega(\varphi_i, \varphi_{i'})$ and then we can use the wavelet transform and the bilinearity of a^ω to obtain $a^\omega(\psi_i^l, \psi_{i'}^{l'})$. After that we can set to 0 the entries corresponding to basis wavelet functions with large distance between their supports. To this end, since $\varphi_i \in \tilde{H}^2$ we get by integrating by parts (12) the unweighted version

$$\begin{aligned} a(\varphi_i, \varphi_j) &:= (-A\varphi_i, \varphi_j)_{L^2} = -\int_\Omega\int_\Omega \varphi_i''(x+y)k^{(-2)}(y)dy\varphi_j(x)dx \\ &= \int_{x_{j-1}}^{x_{j+1}}\int_{x_{i-1}}^{x_{i+1}} \varphi_i'(y)\varphi_j'(x)k^{(-2)}(y-x)dydx, \end{aligned}$$

where $i, j = 1, \ldots, N-1$ and $x_0 = -R$. We want to obtain more explicit expression. Denote for $j \in \mathbb{N}_0$

$$
\begin{aligned}
k_j^+ &:= \int_0^h \int_{jh}^{(j+1)h} k^{(-2)}(y-x)dydx \\
k_j^- &:= \int_{jh}^{(j+1)h} \int_0^h k^{(-2)}(y-x)dydx.
\end{aligned}
$$

Now for $j \geq i$,$i, j = 1, \ldots, N-1$ by using that the first derivatives of the local hat basis functions are piecewise constants we obtain

$$
\begin{aligned}
&a(\varphi_j, \varphi_i) = \int_{x_{i-1}}^{x_{i+1}} \int_{x_{j-1}}^{x_{j+1}} \varphi_j'(y)\varphi_i'(x)k^{(-2)}(y-x)dydx \\
&= \frac{1}{h^2}\left(2\int_0^h \int_{(j-i)h}^{(j-i+1)h} k^{(-2)}(y-x)dydx - \int_0^h \int_{(j-i+1)h}^{(j-i+2)h} k^{(-2)}(y-x)dydx \right. \\
&\left. - \int_0^h \int_{(j-i-1)h}^{(j-i)h} k^{(-2)}(y-x)dydx\right).
\end{aligned}
$$

Analogous result can be obtained when $j < i$. Write the above in compact form for the hat basis stiffness matrix $\mathbf{A}$ and $i = 1, \ldots, N-1$

$$
\begin{aligned}
\mathbf{A}_{i,i} &:= a(\varphi_i, \varphi_i) = \frac{1}{h^2}(2k_0^+ - k_1^+ - k_1^-), \\
\mathbf{A}_{i,i+j} &:= a(\varphi_{i+j}, \varphi_i) = \frac{1}{h^2}(2k_j^+ - k_{j+1}^+ - k_{j-1}^+), \ j = 1, \ldots, N-i-1 \\
\mathbf{A}_{i,i-j} &:= a(\varphi_{i-j}, \varphi_i) = \frac{1}{h^2}(2k_j^- - k_{j+1}^- - k_{j-1}^-), \ j = 1, \ldots, i-1.
\end{aligned}
$$

We can compute k_j^+ and k_j^- exactly by using antiderivatives of order -3 and -4 like

$$
\begin{aligned}
&k_0^+ = k_0^- := \int_0^h \int_0^h k^{(-2)}(y-x)dydx \\
&= \int_0^h (k^{(-3)}(0^-) - k^{(-3)}(-x) + k^{(-3)}(h-x) - k^{(-3)}(0^+))dx \\
&= h(k^{(-3)}(0^-) - k^{(-3)}(0^+)) + k^{(-4)}(-h) - k^{(-4)}(0^-) - k^{(-4)}(0^+) + k^{(-4)}(h),
\end{aligned}
$$

where $k^{(-i)}(0^-), k^{(-i)}(0^+)$ are respectively the left and right limits at 0. By analogy we obtain

$$
\begin{aligned}
k_j^+ &= -2k^{(-4)}(jh) + k^{(-4)}((j-1)h) + k^{(-4)}((j+1)h), \ j = 1, \ldots, N-1 \\
k_j^- &= -2k^{(-4)}(-jh) + k^{(-4)}(-(j-1)h) + k^{(-4)}(-(j+1)h), \ j = 1, \ldots, N-1,
\end{aligned}
$$

where if $j = 1$ we have $k^{(-4)}(-(j-1)h) = k^{(-4)}(0^-)$ and $k^{(-4)}((j-1)h) = k^{(-4)}(0^+)$.

For the assembly of the weighted hat basis stiffness matrix $\mathbf{A}^\omega$ we use numerical quadrature after noting that $E^{-\omega}\varphi_i \in \tilde{H}^2$ and we have

$$
\begin{aligned}
a^\omega(\varphi_i, \varphi_j) &:= (E^\omega(-A)E^{-\omega}\varphi_i, \varphi_j)_{L^2} \\
&= -\int_{\mathbb{R}}\int_{\mathbb{R}} (E^{-\omega}\varphi_i)''(x+y)k^{(-2)}(y)dy\varphi_j(x)\mathrm{e}^{\omega x}dx \\
&= -\omega^2 \int_{\mathbb{R}}\int_{\mathbb{R}} \varphi_i(x+y)\mathrm{e}^{-\omega y}k^{(-2)}(y)dy\varphi_j(x)dx \\
&+ 2\omega \int_{\mathbb{R}}\int_{\mathbb{R}} \varphi_i'(x+y)\mathrm{e}^{-\omega y}k^{(-2)}(y)dy\varphi_j(x)dx \\
&- \int_{\mathbb{R}}\int_{\mathbb{R}} \varphi_i''(x+y)\mathrm{e}^{-\omega y}k^{(-2)}(y)dy\varphi_j(x)dx \\
&= -\omega^2 \int_{x_{j-1}}^{x_{j+1}}\int_{x_{i-1}}^{x_{i+1}} \varphi_i(y)k_{-\omega}^{(-2)}(y-x)dy\varphi_j(x)dx \\
&+ 2\omega \int_{x_{j-1}}^{x_{j+1}}\int_{x_{i-1}}^{x_{i+1}} \varphi_i'(y)k_{-\omega}^{(-2)}(y-x)dy\varphi_j(x)dx \\
&+ \int_{x_{j-1}}^{x_{j+1}}\int_{x_{i-1}}^{x_{i+1}} \varphi_i'(y)k_{-\omega}^{(-2)}(y-x)dy\varphi_j'(x)dx.
\end{aligned}
$$

The operator $A_d^\omega f := E^{-\omega}\tilde{P}_L\tilde{A}^\omega P_I(E^\omega f)$ which is used in $\psi_d^{0,0}$ can now be computed, because $\tilde{A}^\omega$ acts on element in the basis space (i.e. $P_I(E^\omega f)$) and thus for computing $\tilde{P}_L\tilde{A}^\omega P_I(E^\omega f)$ we only need the corresponding sparse stiffness matrix $\tilde{\mathbf{A}}^\omega$.
To give a practical example we mention the case when X is CGMY process. This means X is pure jump (i.e. $\sigma^2 = 0$) Lévy process with kernel function $k(x)$ corresponding to the jump measure $F(x)$ given by

$$
k(x) := C \begin{cases} \frac{e^{-Mx}}{x^{1+Y}} & \text{, if } x > 0, \\ \frac{e^{Gx}}{(-x)^{1+Y}} & \text{, if } x < 0, \end{cases}
$$

where $C, G > 0, M > 1$ and $Y < 2$. If $G > 2, M > 2$ and $Y > 0$, then X satisfies assumptions $(A1)-(A4)$. It turns out that this form of kernel allows exact integration and expression for its antiderivatives - for this purpose symbolic integration software can be used.

6 Numerical results

In the numerical experiments we use CGMY underlying process with parameters

$$
C = 1, G = 9, M = 15, Y = 0.2.
$$

We consider European put option with strike $K = 1$ and maturity 1. We further set the coercivity ensuring parameter $q = 0$ and the slope of the dG scheme $\mu = 2$. We do not subtract $\overline{H}^{\epsilon_0}$ before solving the option price equation, i.e. we do not solve for the excess to payoff. As reference solution we use the finer mesh level that we were able to compute in reasonable time, that is $L = 5$. The errors at strike price at time to maturity for levels $L = 3$ and $L = 4$ are 0.0109 and 0.0067 respectively so we obtain convergence.
For the trading strategy we do not get reasonable results. Below are graphs depicting the strategy.
The hedging error is expensive to compute because of the right hand side function and the necessity to do numerical integration and scalar product in each time interval of the dG scheme, resulting in evaluations of the approximative operator $\psi_d^{0,0}$.

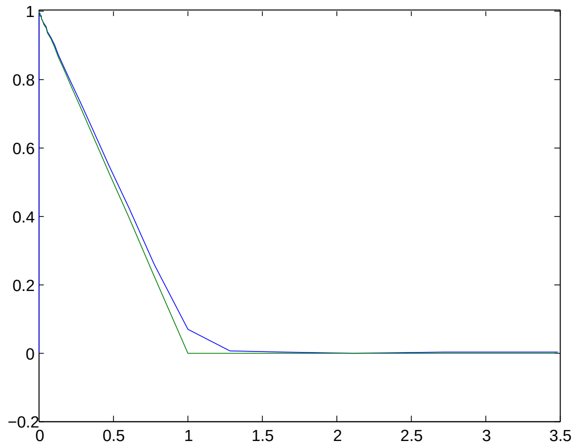

Figure 1: Put option price $\tilde{V}(t, x)$ at time $t = 1$ and $L = 3$

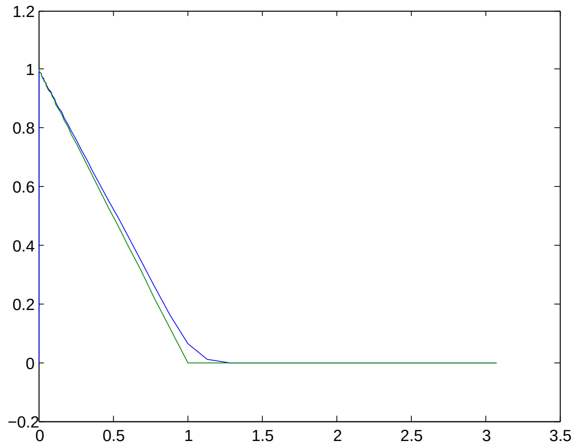

Figure 2: Put option price $\tilde{V}(t, x)$ at time $t = 1$ and $L = 4$

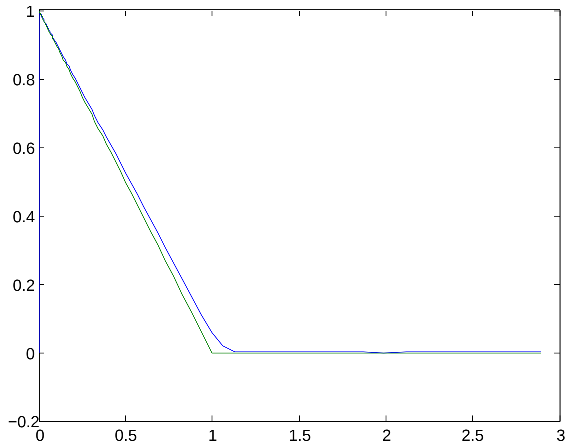

Figure 3: Put option price $\tilde{V}(t,x)$ at time $t=1$ and $L=5$

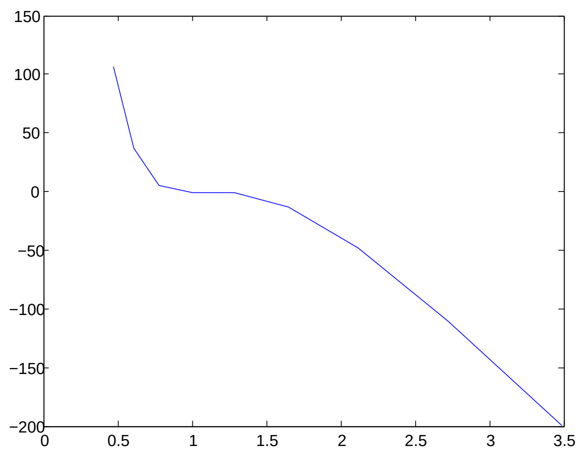

Figure 4: Trading strategy $\tilde{\vartheta}(t,x)$ at time $t=1$ and $L=3$

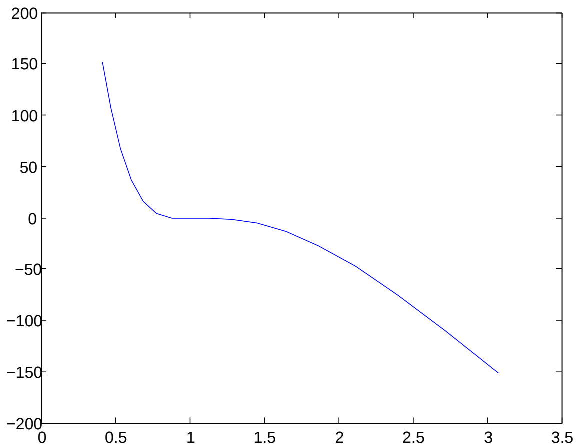

Figure 5: Trading strategy $\tilde{\vartheta}(t, x)$ at time $t = 1$ and $L = 4$

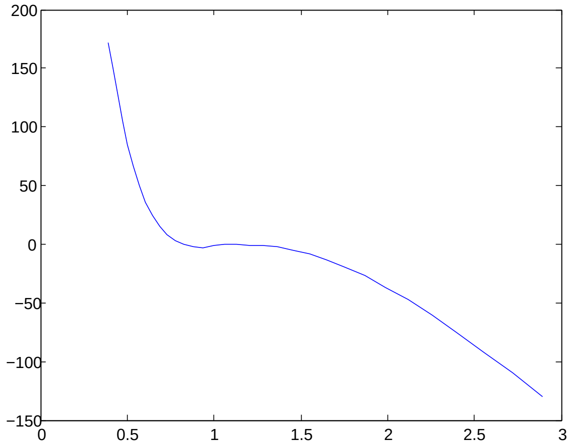

Figure 6: Trading strategy $\tilde{\vartheta}(t, x)$ at time $t = 1$ and $L = 5$

References

[1] Robert A. Adams and John J. F. Fournier, *Sobolev spaces. (second ed.)*, Pure and Applied Mathematics 140. New York, NY: Academic Press., 2003.

[2] N. Bouleau and D. Lamberton, *Residual risks and hedging strategies in markovian markets*, Stochastic Processes Appl. **33** (1989), no. 1, 131–150.

[3] Svetlana Boyarchenko and Sergei Levendorskii, *Non-gaussian merton-black-scholes theory*, World Scientific, New Jersey-London-Singapore-Hong Kong, 2002.

[4] Wolfgang Dahmen and Rob Stevenson, *Element-by-element construction of wavelets satisfying stability and moment conditions*, SIAM J. Numerical Analysis **37** (1999), 319–352.

[5] Gene Golub and Charles F. Van Loan, *Matrix computations. (third ed.)*, Texts and Readings in Mathematics 43. New Delhi: Hindustan Book Agency. 722p., 2007.

[6] J. Jacod and A. Shiryaev, *Limit theorems for stochastic processes (second ed.)*, Springer, Berlin, 2003.

[7] Markus Jürgens, *A semigroup approach to the numerical solution of parabolic differential equations*, PhD thesis, TH Aachen, 2005.

[8] A.-M. Matache, C. Schwab, and Thomas P. Wihler, *Linear complexity solution of parabolic integro-differential equations*, Numer. Math. **104** (2006), no. 1, 69–102.

[9] T. von Petersdorff and C. Schwab, *Numerical solution of parabolic equations in high dimensions*, M2AN **38** (2004), no. 1, 93–127.

[10] Walter Rudin, *Functional analysis. 2nd ed.*, International Series in Pure and Applied Mathematics. New York, NY:McGraw-Hill. xviii, 424p., 1991.

[11] Reinhold Schneider, *Multiscale and wavelet matrix compression.(multiskalen- und wavelet-matrixkompression.)*, PhD thesis, TH Darmstadt, 1998.

[12] D. Schoetzau and C. Schwab, *Time discretization of parabolic problems by the hp-version of the discontinuous galerkin finite element method*, SIAM J. Num. Anal. **38** (2000), 837–875.

[13] Bernhard Antonin Johannes Vesenmayer, *Efficient numerical solution of the variance-optimal hedging problem in geometric lévy models*, PhD thesis, Technische Universität München, 2009.

[14] T. von Petersdorff, R. Schneider, and C. Schwab, *Multiwavelets for second kind integral equations*, SIAM J. Num. Anal. **34** (1997), 2212–2227.

[15] T. von Petersdorff and C. Schwab, *Wavelet discretizations of parabolic integro-differential equations*, SIAM J. Num. Anal. **41** (2003), no. 1, 159–180.