Option pricing in Hilbert space-valued jump-diffusion models using partial integro-differential equations

Peter Hepperger[*]

Hilbert space-valued jump-diffusion models are employed for various markets and derivatives. Examples include swaptions, which depend on continuous forward curves, and basket options on stocks. Usually, no analytical pricing formulas are available for such products. Numerical methods, on the other hand, suffer from exponentially increasing computational effort with increasing dimension of the problem, the "curse of dimension." In this paper, we present an efficient approach using partial integro-differential equations. The key to this method is a dimension reduction technique based on a Karhunen–Loève expansion, which is also known as proper orthogonal decomposition. Using the eigenvectors of a covariance operator, the differential equation is projected to a low-dimensional problem. Convergence results for the projection are given, and the numerical aspects of the implementation are discussed. An approximate solution is computed using a sparse grid combination technique and discontinuous Galerkin discretization. The main goal of this article is to combine the different analytical and numerical techniques needed, presenting a computationally feasible method for pricing European options. Numerical experiments show the effectiveness of the algorithm.

1 Introduction

PRICING WITH PARTIAL DIFFERENTIAL EQUATIONS   The pricing of options written on a single asset by means of partial differential equations (PDEs) is a well-studied problem. Efficient numerical solutions have been proposed not only for the classical Black–Scholes setting but also for Lévy-driven processes (cf., e.g., [24, 12, 13] and the references therein). Due to the jump part, an additional nonlocal term occurs when using

Lévy models. Thus, partial integro-differential equations (PIDEs) have to be solved instead of plain PDEs.

The pricing problem becomes much more involved when the derivative (e.g., a basket option) depends on more than one asset. PIDE methods have recently been extended to such multidimensional settings [28, 36]. The main problem here is the exponentially increasing computational effort for high-dimensional problems. While this *curse of dimension* can be reduced effectively by applying sparse grid discretizations, numerically feasible dimensions n are still moderate, usually $n \leq 10$. On the other hand, real world products often imply much higher dimensional equations. In fact, derivatives (e.g., swaptions) may even depend on a continuum of "assets" and thus be infinite dimensional. This makes dimension reduction techniques an interesting tool. To motivate the subsequently presented techniques, consider the following example, which naturally arises in energy markets.

MOTIVATING EXAMPLE: ELECTRICITY SWAPTIONS   Since the liberalization of many European electricity markets during the 1990s, producers, consumers, and speculators have traded in electricity on energy exchanges. The Scandinavian Nordpool, the European Energy Exchange (EEX) in Germany, and the Amsterdam Power Exchange (APX) are the largest European trading centers for electricity. Traded products include spot, forward, and futures contracts and options on these. The most liquidly traded underlyings are contracts of futures type. These are agreements traded at time t for a constant delivery of 1 MW of electricity over a certain future period of time $[T_1, T_2]$, while in return a fixed rate $F(t; T_1, T_2)$ is paid during this delivery period. These products are also called *electricity swaps*. The relation of spot and forward prices is not clearly defined for electricity due to its nonstorability [3, 4, 6]. This difficulty can be avoided by directly modeling the forward curve under a risk neutral measure [1, 20], similar to the Heath–Jarrow–Morton approach for bond markets. For every maturity $u \in [T_1, T_2]$, let

$$
(1) \qquad S(t, u) := \lim_{v \to u} F(t; u, v)
$$

be the corresponding value of the forward curve at time $t \leq u$. In practice, the forward curve is constructed from a discrete set of available market prices [5]. Several authors propose exponential additive forward curve models of diffusion or jump-diffusion type [7, 20], i.e.,

$$
(2) \qquad \begin{aligned} X(t, u) &= \int_0^t \gamma(s, u)ds + \sum_{k=1}^{n_W} \int_0^t \sigma_k(s, u)dW_k(s) + \sum_{k=1}^{n_J} \int_0^t \int_{\mathbb{R}} \eta_k(s, u)y\ \widetilde{M}_k(dy, ds), \\ S(t, u) &= S(0, u)\exp\left(X(t, u)\right), \end{aligned}
$$

where W_k are scalar Brownian motions and $\widetilde{M}_k$ are compensated random jump measures, all of them independent. Assuming a constant interest rate $r \geq 0$, a European call option on the swap $F(T; T_1, T_2)$ with strike rate K and maturity $T \leq T_1$ has the

discounted price

$$V(t) = e^{-rT}E\left[(T_2 - T_1)\left(F(T; T_1, T_2) - K\right)^+ \Big| \mathcal{F}_t\right]$$

(3)

$$= e^{-rT}(T_2 - T_1)E\left[\left(\int_{T_1}^{T_2} w(u; T_1, T_2)S(T, u)du - K\right)^+ \Big| \mathcal{F}_t\right],$$

at time $t \leq T$, where

(4)
$$w(u; T_1, T_2) = e^{-ru} \Big/ \int_{T_1}^{T_2} e^{-rs}ds$$

is a deterministic, nonnegative discounting factor. In general, there is no explicit representation for the probability density of the integral term inside the conditional expectation. This makes pricing more difficult than in the one-dimensional standard setting. One possible valuation method is to use log-normal approximation formulas [9]. These are fast but sometimes give poor results in the presence of jumps. Numerical Monte Carlo simulations, on the other hand, are precise but rather time consuming since they converge slowly. In the present work, we will discuss an efficient numerical approximation method which is based on solving Hilbert space-valued PIDEs and dimension reduction techniques.

OUTLINE OF THE ARTICLE Both basket options and options depending on forward curves can be viewed as special cases of Hilbert space-valued problems. The main goal of this article is to present an efficient method for pricing European options on such derivatives. In Section 2, we state the general model and hypotheses for our approach. The corresponding Hilbert space-valued PIDE for pricing European options is derived in Section 2.3. This PIDE can be approximated with finite-dimensional equations. To this end, we apply Karhunen–Loève approximation, which is closely related to proper orthogonal decomposition and factor analysis. In Section 3, the dimension reduction method is described. We show how to construct an optimal set of approximating basis functions by solving an eigenvector problem. This basis set is then used to transform the PIDE, and the main results of this work are presented. After proving existence and uniqueness, convergence results and error bounds are given for the solutions of the approximating problems. In Section 4, we discuss how to perform the dimension reduction numerically, using a Galerkin approach and the sparse grid combination technique. Finally, we apply the pricing method to test problems (a basket option and an electricity swaption) and demonstrate the efficiency of the method with numerical experiments in Section 5.

2 HILBERT SPACE-VALUED JUMPDIFFUSION

We now state the Hilbert space-valued model used throughout this article. For a definition of stochastic processes and integration in Hilbert spaces with respect to

Brownian motion, see, e.g., [14, 22]. An overview of Poisson random measures in Hilbert spaces can be found in [18], and the Lévy case is treated in [25]. Infinite-dimensional stochastic analysis and its applications to interest-rate theory and Heath–Jarrow–Morton models are presented in [10].

2.1 EXPONENTIAL ADDITIVE MODEL

Let $D \subset \mathbb{R}^m$ and let μ_D be a measure on D. Then

$$(5) \qquad H := L^2(D, \mu_D)$$

is a separable Hilbert space. For every $h \in H$, we denote the corresponding norm by

$$(6) \qquad \|h\|_H := \sqrt{\int_D \left[h(u)\right]^2 \mu_D(u)}.$$

For electricity swaptions, we choose $D = [T_1, T_2]$, and $\mu_D = \lambda_D$ is the Lebesgue measure. Then H can be interpreted as the space of forward curves on the delivery period D. For basket options, we choose D to be a finite (though possibly large) index set with n entries, each index corresponding to one asset. The measure μ_D is then simply the counting measure, and the norm $\|\cdot\|_H$ is the Euclidean norm on $\mathbb{R}^n$.

Consider the H-valued exponential-additive stochastic model

$$(7) \qquad \begin{aligned} X(t) &= \int_0^t \gamma(s)\, ds + \int_0^t \sigma(s)\, dW(s) + \int_0^t \int_H [\eta(s)](\xi)\, \widetilde{M}(d\xi, ds), \\ S(t) &= S(0) \exp\left(X(t)\right) \end{aligned}$$

for $t \in [0, T]$, with initial value $S(0) \in H$. Subsequently, we will write $f(t, u) := [f(t)](u)$ for every $f : [0, T] \to H$, $u \in D$, and similarly $g(t, h) := [g(t)](h)$ for every $g : [0, T] \to L(H, H)$, $h \in H$. The exponential function in the model is defined pointwise, i.e., $S(t, u) = S(0, u) \exp\left(X(t, u)\right)$ for a.e. $u \in D$. The diffusion part is driven by an H-valued Wiener process W whose covariance is described by a symmetric, non-negative definite trace class operator Q. The driving process $\widetilde{M}$ for the jump part is the compensated random measure of an H-valued compound Poisson process

$$(8) \qquad J(t) = \sum_{i=1}^{N(t)} Y_i,$$

which is independent of W. Here, N denotes a Poisson process with intensity λ and $Y_i \sim P^Y$ ($i = 1, 2, \ldots$) are iid on H. The corresponding Lévy measure is denoted by $\nu = \lambda P^Y$. Further, denote by $L(H, H)$ the space of all bounded linear operators on H and let $\gamma : [0, T] \to H$, $\sigma : [0, T] \to L(H, H)$, and $\eta : [0, T] \to L(H, H)$ be deterministic functions. In particular, this model generalizes the electricity swaption model (2) driven by a sum of scalar processes.

The following hypothesis is assumed to hold throughout this article. For an introduction to time-dependent Bochner spaces, such as $L^2(0, T; H)$, see [16, Chap. 5.9].

Assumption 2.1. *Suppose that the second exponential moment of the jump distribution Y exists:*

(9) $$E[e^{2\|Y\|_H}] = \int_H e^{2\|\xi\|_H} P^Y(d\xi) < \infty.$$

Assume further that $\|\eta(t)\|_{L(H,H)} \le 1$ for a.e. $t \in [0, T]$,

(10) $$\gamma \in L^2(0, T; H), \quad and \quad \sigma \in L^2(0, T; L(H, H)).$$

Under Assumption 2.1, $\big(X(t)\big)_{t \ge 0}$ is an additive process with finite activity and finite expectation. This simplifies notation, since no truncation of large jumps is needed in the characteristic function.

Theorem 2.2. *The process $\big(X(t)\big)_{t \ge 0}$ is square-integrable:*

(11) $$\sup_{0 \le t \le T} E\big[\|X(t)\|_H^2\big] < \infty.$$

Proof. The definition of the process $\big(X(t)\big)_{t \ge 0}$ yields
(12)
$$E\big[\|X(t)\|_H^2\big] \le 3E\bigg[\bigg\|\int_0^t \gamma(s)\, ds\bigg\|_H^2 + \bigg\|\int_0^t \sigma(s)\, dW(s)\bigg\|_H^2 + \bigg\|\int_0^t \int_H \eta(s, \xi)\, \widetilde{M}(d\xi, ds)\bigg\|_H^2\bigg]$$

We now apply three different results to the three integrals on the right-hand side. For the first one, we use the basic properties of Bochner integrals and Jensen's inequality to obtain

(13) $$\bigg\|\int_0^t \gamma(s)\, ds\bigg\|_H^2 \le \int_0^t \|\gamma(s)\|_H^2\, ds \le \|\gamma\|_{L^2(0,T;H)}^2.$$

By the definition of the integral with respect to H-valued Gaussian processes and the results from [14, eqs. (4.8) and (4.10)], we have

(14) $$E\bigg\|\int_0^t \sigma(s)\, dW(s)\bigg\|_H^2 \le (\mathrm{Tr}\, Q)\, E \int_0^t \|\sigma(s)\|_{L(H,H)}^2\, ds = (\mathrm{Tr}\, Q)\, \|\sigma\|_{L^2(0,T;L(H,H))}^2,$$

where $\mathrm{Tr}\, Q$ denotes the trace of the covariance operator of W. Finally, from Young's inequality and [17, Prop. 3.3] we get

(15)
$$E\bigg[\bigg\|\int_0^t \int_H \eta(s, \xi)\, \widetilde{M}(d\xi, ds)\bigg\|_H^2\bigg] \le C \int_0^t \int_H \|\eta(s, \xi)\|_H^2\, \nu(d\xi) ds$$
$$\le C\lambda T \int_H \|y\|_H^2\, P^Y(dy).$$

Combining the above estimates and employing Assumption 2.1 yields (11), since the right-hand side in each estimate is independent of t. $\square$

Theorem 2.3. *The characteristic function of $X(t)$ is given by*

$$(16) \quad E\big[e^{i\langle X(t),h\rangle_H}\big] = \exp\left[i\left\langle \int_0^t \gamma(s)ds, h\right\rangle_H - \frac{1}{2}\left\langle \left[\int_0^t \sigma(s)Q\sigma^*(s)ds\right](h), h\right\rangle_H\right.$$
$$\left. + \int_0^t \int_H \left(e^{i\langle \eta(s,\xi),h\rangle_H} - 1 - i\left\langle \eta(s,\xi), h\right\rangle_H\right) \nu(d\xi)ds\right]$$

for every $h \in H$, where $\sigma^*(s)$ is the adjoint operator of $\sigma(s)$.

Proof. The drift γ is deterministic, and so the first term on the right-hand side of (16) is trivial. Since we have finite second moments by Theorem 2.2, we may apply [22, Thm. 4] to obtain the characteristic function of the diffusion and jump parts. It remains to verify the expression for the covariance operator of the diffusion. Applying [14, Prop. 4.13] yields

$$(17)$$
$$E\left[\left\langle \int_0^t \sigma(s)dW(s), h_1\right\rangle_H \left\langle \int_0^t \sigma(s)dW(s), h_2\right\rangle_H\right] = \left\langle \left[\int_0^t \sigma(s)Q\sigma^*(s)\,ds\right](h_1), h_2\right\rangle_H$$

for every $h_1, h_2 \in H$. The integral on the right-hand side is a Bochner integral with values in $L(H, H)$. $\qquad\square$

2.2 Equivalent exponential Lévy model

As already stated, the main goal of this article is pricing European options depending on $S(T) = S_0 \exp\big(X(T)\big)$. Since by definition the payoff of such an option depends only on the value of X at time $t = T$, the characteristic function of $X(T)$ completely determines the price of the product. Therefore, it is possible to construct a (time homogeneous) Lévy process $\big(X_L(t)\big)_{t\geq 0}$ which produces the same terminal distribution and thus the same price (cf. [31, Chap. 11]). We denote the Borel sets on H by $\mathcal{B}(H)$ and the indicator function of a set $B \in \mathcal{B}(H)$ by χ_B. The Lévy–Khinchin triplet of the Lévy process is given by

$$(18) \qquad A_L(h_1, h_2) = \frac{1}{T}\left\langle \left[\int_0^T \sigma(s)Q\sigma^*(s)\,ds\right](h_1), h_2\right\rangle_H,$$

$$(19) \qquad \gamma_L = \frac{1}{T}\int_0^T \gamma(t)dt,$$

$$(20) \qquad \nu_L(B) = \frac{1}{\lambda T}\int_0^T \int_H \chi_B\big(\eta(t,\xi)\big)\nu(d\xi)\,dt \quad \text{for } B \in \mathcal{B}(H),$$

where $A_L : H \times H \to \mathbb{R}$ is a bilinear covariance operator, $\gamma_L \in H$, and ν_L is a finite activity Lévy measure on H. Note that the resulting characteristic function of X_L at time $t = T$ is identical to (16).

Let $I = \{1, 2, \ldots, \dim H\}$, if H has finite dimension, or $I = \mathbb{N}$ otherwise. In order to obtain an explicit Lévy representation for $X(T)$, define

$$(21) \qquad Q_L : \begin{cases} H \to H, \\ h \mapsto \frac{1}{T}\left[\int_0^T \sigma(s)Q\sigma^*(s)\,ds\right](h). \end{cases}$$

This is a symmetric nonnegative definite operator with finite trace, since, by construction, and by the proof of Theorem 2.2, we have

$$
\begin{aligned}
\sum_{l \in I} \langle Q_L p_l, p_l \rangle_H &= \sum_{l \in I} \frac{1}{T} E\left[\left\langle \int_0^t \sigma(s) dW(s), p_l \right\rangle_H \left\langle \int_0^t \sigma(s) dW(s), p_l \right\rangle_H \right] \\
&= \frac{1}{T} E \left\| \int_0^t \sigma(s) dW(s) \right\|_H^2 \leq (\mathrm{Tr}\, Q) \, \|\sigma\|_{L^2(0,T;L(H,H))}^2
\end{aligned}
$$
(22)

for every orthonormal basis $(p_l)_{l \in I}$ of H. Consequently, the operator Q_L is compact by [14, Prop. C.3] and, in particular, a trace class operator. By [15, Thm. 1.2.1], there is a unique Gaussian probability measure with mean 0 and covariance operator Q_L. Moreover, there is a corresponding Q_L-Wiener process by [14, Prop. 4.2], which we denote by W_L.

In addition, we introduce the H-valued compound Poisson process

$$
J_L(t) = \sum_{i=1}^{N(t)} Y_{L,i},
$$
(23)

where $Y_{L,i} \sim P^{Y_L}$ are iid random variables with values in H, and

$$
P^{Y_L}(Y_{L,i} \in B) = \begin{cases} \frac{1}{\lambda T} \int_0^T \int_H \chi_B(\eta_t(y)) \nu(dy)\, dt & \text{if } \lambda > 0, \\ \chi_B(0) & \text{if } \lambda = 0, \end{cases}
$$
(24)

for every Borel set $B \in \mathcal{B}(H)$. The random measure corresponding to J_L is denoted by M_L and its compensated version by $\widetilde{M}_L$. Combined, we obtain the Lévy process

$$
X_L(t) = \gamma_L t + W_L(t) + \int_0^t \int_H \zeta \widetilde{M}_L(d\zeta, ds),
$$
(25)

with the same distribution as X at $t = T$. The last two summands are H-martingales. We will use this model subsequently in place of (7), since it is completely equivalent with respect to European option pricing.

Due to the existence of second moments (Theorem 2.2), the bounded linear covariance operator

$$
\mathcal{C}_X : \begin{cases} H \to H' \cong H, \\ h \mapsto E\big[\langle X_L(T) - E[X_L(T)], h \rangle_H \langle X_L(T) - E[X_L(T)], \cdot \rangle_H \big] \end{cases}
$$
(26)

is well defined, and $E[X_L(T)] = \gamma_L T$. For notational convenience, define the centered process

$$
Z_L(t) := X_L(t) - E[X_L(t)] = X_L(t) - \gamma_L t, \quad t \in [0, T].
$$
(27)

The following theorem gives a summary of the properties of $\mathcal{C}_X$.

Theorem 2.4. *The operator $\mathcal{C}_X$ defined in (26) is a symmetric and nonnegative definite trace class operator (and thus compact).*

Proof. Since $\mathcal{C}_X$ is a covariance operator, it is symmetric and nonnegative definite by definition. It remains to prove that the trace of $\mathcal{C}_X$ is finite. To this end, let $(p_l)_{l \in I}$, be an arbitrary orthonormal base of H. Then, by dominated convergence,

$$
\begin{aligned}
(28) \quad \operatorname{Tr} \mathcal{C}_X &= \sum_{l \in I} \langle \mathcal{C}_X p_l, p_l \rangle_H = \sum_{l \in I} E\Big[\langle Z_L(T), p_l \rangle_H^2 \Big] = E\Big[\sum_{l \in I} \langle Z_L(T), p_l \rangle_H^2 \Big] \\
&= E\Big[\|Z_L(T)\|_H^2 \Big] < \infty
\end{aligned}
$$

holds. We use [14, Prop. C.3] again to conclude that $\mathcal{C}_X$ is compact and thus a trace class operator. $\qquad\square$

We denote the eigenspace of $\mathcal{C}_X$ corresponding to eigenvalue 0 by $E_0(\mathcal{C}_X)$. Its orthogonal complement $E_0(\mathcal{C}_X)^\perp$ is the linear span of all eigenvectors corresponding to positive eigenvalues. This is the subspace of H to which the centered process Z_L is restricted almost surely, since

$$
(29) \qquad E\big[\langle Z_L(t), h \rangle_H^2 \big] = 0 \quad \text{for every } h \in E_0(\mathcal{C}_X) \text{ and a.e. } t \geq 0.
$$

Finally, in analogy to the requirement of a nonvanishing volatility in one-dimensional jump-diffusion models, an assumption on the Q_L-Wiener process W_L is needed.

Assumption 2.5. *Assume that the restriction of Q_L to the subspace $E_0(\mathcal{C}_X)^\perp \subset H$ is positive definite, i.e.,*

$$
(30) \qquad \langle Q_L h, h \rangle_H > 0 \quad \text{for every } h \in E_0(\mathcal{C}_X)^\perp \backslash \{0\}.
$$

This means that Z_L has a nonvanishing Brownian component for all directions in H, which are *not* almost surely orthogonal to the trajectory of the process. This is necessary for the coercivity property (Gårding's inequality) of the PIDE which we are going to derive in the next section.

2.3 HILBERT SPACE-VALUED PIDE

We consider a European option depending on the value of S at time $t = T$. Since $S(T)$ is a deterministic function of $X(T)$, and $X(T)$ is distributed like $X_L(T) = \gamma_L T + Z_L(T)$, we may equivalently price a European option whose payoff G is a function of Z_L. Hence, if $z \in H$ is the value of Z_L at time t, the value V of the option at time $t \leq T$, discounted to time 0, is described by

$$
(31) \qquad V(t, z) := e^{-rT} E\big[G(z + Z_L(T - t)) \big].
$$

Note that V is a martingale under the risk neutral measure. In this section, an Itô formula for Hilbert space-valued random variables and the martingale property of V are employed to derive a PIDE for V.

First, let us recall the definition of derivatives on Hilbert spaces. We denote by $D_z V(t,z) \in L(H, \mathbb{R})$ and $D_z^2 V(t,z) \in L(H, H)$ continuous linear operators such that

(32) $\qquad V(t, z + \zeta) = V(t,z) + [D_z V(t,z)](\zeta) + \frac{1}{2} \left\langle [D_z^2 V(t,z)](\zeta), \zeta \right\rangle_H + o(\|\zeta\|_H^2)$

for every $\zeta \in H$. It is often convenient to identify $D_z^2 V(t,z)$ with a bilinear form on $H \times H$, setting

(33) $\qquad\qquad [D_z^2 V(t,z)](\zeta_1, \zeta_2) := \left\langle [D_z^2 V(t,z)](\zeta_1), \zeta_2 \right\rangle_H .$

The partial derivative with respect to time is denoted with $\partial_t V(t,z)$. We assume the following regularity condition for V, which is, in particular, a prerequisite for Itô's formula. However, note that this hypothesis is not necessary for the convergence results in Section 3.4.

Assumption 2.6. *Suppose that $V \in C^{1,2}((0,T] \times H, \mathbb{R}) \cap C([0,T] \times H, \mathbb{R})$; i.e., V is continuously differentiable with respect to t and twice continuously differentiable with respect to z. Moreover, assume that the operator norms $\left\| D_z^2 V(t,z) \right\|$, $\left\| D_z V(t,z) \right\|$, and $\left\| \partial_t V(t,z) \right\|$ are bounded.*

As a direct consequence of this assumption, V satisfies the Lipschitz condition

(34) $\qquad\qquad |V(t,z) - V(t, z + \zeta)| \leq K_V \|\zeta\|_H \quad \text{for every } \zeta \in H$

with constant $K_V := \sup_{(s,y) \in [0,T] \times H} \| D_z V(s,y) \|$. We are now able to calculate the stochastic dynamics of V using Itô's formula.

Theorem 2.7. *The discounted price V given by (31) satisfies*

$$dV(t, Z_L(t)) =$$
$$\partial_t V(t, Z_L(t-)) \, dt + \frac{1}{2} \operatorname{Tr} \left([D_z^2 V(t, Z_L(t-))] Q_L \right) dt$$
(35)
$$+ \int_H \left\{ V(t, Z_L(t-) + \zeta) - V(t, Z_L(t-)) - [D_z V(t, Z_L(t-))](\zeta) \right\} \nu_L(d\zeta) \, dt$$
$$+ D_z V(t, Z_L(t-)) \, dW_L(t) + \int_H \left[V(t, Z_L(t-) + \zeta) - V(t, Z_L(t-)) \right] \widetilde{M}_L(d\zeta, dt).$$

Proof. By Itô's formula [22, Thm. 3], we obtain

(36)
$$V(t, Z_L(t)) =$$
$$V(0, Z_L(0)) + \int_0^t \partial_t V(s, Z_L(s-)) ds$$
$$+ \frac{1}{2} \left\langle\!\!\left\langle \int_0^\cdot D_z^2 V(s, Z_L(s-)) \, dW_L(s) \, ; \, W_L(\cdot) \right\rangle\!\!\right\rangle_t$$
$$+ \int_0^t D_z V(s, Z_L(s-)) \, dZ_L(s)$$
$$+ \sum_{0 \leq s \leq t} \left[V(s, Z_L(s-) + \Delta Z_L(s)) - V(s, Z_L(s-)) - [D_z V(s, Z_L(s-))](\Delta Z_L(s)) \right],$$

9

where $\langle\!\langle X ; Y \rangle\!\rangle$ denotes the predictable quadratic covariation of $\langle X, Y \rangle_H$.

We first calculate the covariation. From [14, Cor. 4.14], we know that

(37) $\quad E\Big\langle \int_{t_1}^{t_2} D_z^2 V(s, Z_L(s-)) dW_L(s), W_L(t_2) - W_L(t_1) \Big\rangle_H$

$$= \int_{t_1}^{t_2} \mathrm{Tr}\left(\left[D_z^2 V(s, Z_L(s-)) \right] Q_L \right) ds$$

for every $0 \le t_1 \le t_2$. Consequently, by independence of the increments of W_L, we get

(38) $\quad \Big\langle\!\!\Big\langle \int_0^{\cdot} D_z^2 V(s, Z_L(s-)) \, dW_L(s) \, ; \, W_L(\cdot) \Big\rangle\!\!\Big\rangle_t = \int_0^t \mathrm{Tr}\left(\left[D_z^2 V(s, Z_L(s-)) \right] Q_L \right) ds.$

For the next term in (36), we use the dynamics of Z_L to obtain

$$\int_0^t D_z V(s, Z_L(s-)) dZ_L(s)$$

$$= \int_0^t D_z V(s, Z_L(s-)) d\Big[W_L(s) + \int_0^s \int_H \zeta \widetilde{M}_L(d\zeta, ds_2) \Big]$$

(39)
$$= \int_0^t D_z V(s, Z_L(s-)) \, dW_L(s)$$

$$+ \int_0^t D_z V(s, Z_L(s-)) \, d\Big[\sum_{0 \le s_2 \le s} \Delta Z_L(s_2) - \int_0^s \int_H \zeta \, \nu_L(d\zeta) \, ds_2 \Big].$$

A theorem for interchanging linear operators and Bochner integrals [16, App. E, Thm. 8] yields
(40)
$$\int_0^t D_z V(s, Z_L(s-)) \, d\Big[\int_0^s \int_H \zeta \, \nu_L(d\zeta) \, ds_2 \Big] = \int_0^t \int_H \left[D_z V(s, Z_L(s-)) \right](\zeta) \, \nu_L(d\zeta) \, ds.$$

Plugging (38), (39), and (40) into the Itô dynamics (36) finishes the proof. $\qquad\square$

In order to get a slightly more explicit form of the trace expression in (35), let $(p_l)_{l \in I}$ be an arbitrary orthonormal basis of H. By definition of the trace, this yields

(41) $\quad \mathrm{Tr}\left(\left[D_z^2 V(t, Z_L(t-)) \right] Q_L \right) = \sum_{l \in I} \left[D_z^2 V(t, Z_L(t-)) \right] (Q_L p_l, p_l).$

Theorem 2.8. *The discounted price V of a European option with payoff $G(Z_L(T))$ at maturity T satisfies the PIDE*

(42)
$$-\partial_t V(t, z) = \frac{1}{2} \sum_{l \in I} \left[D_z^2 V(t, z) \right] (Q_L p_l, p_l)$$

$$+ \int_H \Big\{ V(t, z + \zeta) - V(t, z) - \left[D_z V(t, z) \right](\zeta) \Big\} \nu_L(d\zeta),$$

10

with terminal condition

$$(43) \qquad V(T, z) = e^{-rT} G(z),$$

for a.e. $t \in (0, T)$, $z \in E_0(\mathcal{C}_X)^\perp$.

Proof. We employ Theorem 2.7. The penultimate term in (35),

$$(44) \qquad \int_0^t D_z V(s, Z_L(s-)) \, dW_L(s),$$

is a martingale by [14, Thm. 4.12]. In order to show that the integral with respect to the compensated Poisson measure is a martingale too, we apply [29, Thm. 3.11]. The prerequisite for this theorem is a strong integrability condition, which is satisfied due to [29, Thm. 3.12], since

$$(45) \quad \int_0^t \int_H E \left| V(s, Z_L(s-) + \zeta) - V(s, Z_L(s-)) \right| \nu_L(d\zeta) ds \leq t \int_H K_V \|\zeta\|_H \, \nu_L(d\zeta) < \infty.$$

The remaining integral terms in (35) are continuous in t and of finite variation. Consequently, the martingale property of V, together with the fact that continuous martingales of finite variation are almost surely constant [27, Th. 27], yields the PIDE. $\qquad \square$

3 Dimension reduction for the PIDE

The main goal of this section is to introduce a low-dimensional approximation of the H-valued process X_L. To this end, the *Karhunen–Loève expansion* of $X_L(T)$ is used, which is in fact identical to *proper orthogonal decomposition* (POD). It is also closely related to *principal component analysis* (which is commonly used for data analysis) and *factor analysis* (which uses additional error terms in the decomposition). All of these methods are based on the construction of a small set of orthogonal basis elements which can be used to approximate X_L in some L^2-norm. For an overview of POD methods in the context of deterministic differential equations, see [21]. An introduction to Karhunen–Loève expansions of stochastic processes can be found in [23, Chap. 37]. Numerical aspects of the method and most of the theory needed here are presented in [32]. After deriving the low-dimensional approximating PIDE, we will show existence and uniqueness of a solution. Finally, we study convergence of the calculated option prices and give error estimates.

3.1 Karhunen–Loève approximation for jump-diffusion

While principal component analysis and factor analysis are usually applied to analyze empirical data, we apply the Karhunen–Loève method directly to our model. The dimension reduction takes place in the state space of the previously H-valued process. Let us now give a mathematically precise formulation of what is meant by "approximating $X_L(T)$."

Definition 3.1. *A sequence of orthonormal elements $p_l \in H = L^2(D)$, $l \in I$, is called a* POD basis *for $X_L(T)$, if it solves the minimization problem*

$$(46) \qquad \min_{\langle p_i, p_j \rangle_H = \delta_{ij}} E\left[\left\| X_L(T) - \left(E[X_L(T)] + \sum_{l=1}^{d} p_l \langle Z_L(T), p_l \rangle_H \right) \right\|_H^2 \right]$$

for every $d \in I$.

In other words, a POD basis is a set of deterministic orthonormal functions such that we expect the projection of the random vector $Z_L(T) = X_L(T) - E[X_L(T)] \in H$ onto the first d elements of this basis to be a good approximation.

Remark 3.2. *Since we are interested in pricing European options, we include only the value of X_L at time $t = T$ in definition 3.1. However, we could approximate the whole trajectory of X as well by using the difference to its projection in the space $L^2(0, T; H)$. This may be useful for pricing path-dependent derivatives, which we will study separately in future work.*

Approximation with a POD basis is equivalent to using the partial sum of the first d elements of a Karhunen–Loève expansion, which itself is closely connected to the eigenvector problem of the covariance operator $\mathcal{C}_X$ defined in (26). The following theorem shows that the eigenvectors of $\mathcal{C}_X$ are indeed the POD basis we are looking for.

Theorem 3.3. *A sequence of orthonormal eigenvectors $(p_l)_{l \in I}$ of the operator $\mathcal{C}_X$, ordered by the size of the corresponding eigenvalues $\mu_1 \geq \mu_2 \geq \ldots \geq 0$, solves the maximization problem*

$$(47) \qquad \max_{\langle p_i, p_j \rangle_H = \delta_{ij}} \sum_{l=1}^{d} \langle \mathcal{C}_X p_l, p_l \rangle_H$$

for every $d \in I$. The maximum value is

$$(48) \qquad \sum_{l=1}^{d} \langle \mathcal{C}_X p_l, p_l \rangle_H = \sum_{l=1}^{d} \mu_l.$$

Moreover, the eigenvectors are a POD basis in the sense of definition 3.1, and the expectation of the projection error is

$$(49) \qquad E\left[\left\| Z_L(T) - \sum_{l=1}^{d} p_l \langle Z_L(T), p_l \rangle_H \right\|_H^2 \right] = \sum_{l=d+1}^{\dim H} \mu_l.$$

Proof. This is an application of [32, Thm. 2.7 and Prop. 2.8]. $\qquad \square$

3.2 Projection on d-dimensional subspace

Subsequently, let $(p_l)_{l \in I}$ and $(\mu_l)_{l \in I}$ denote the orthonormal basis and eigenvalues from Theorem 3.3. Further, let

$$
(50) \qquad\qquad U_d := \mathrm{span}\{p_1, p_2, \ldots, p_d\} \subset H
$$

be the d-dimensional subspace spanned by the eigenvectors corresponding to the largest eigenvalues. We will assume that $d \leq \dim(E_0(\mathcal{C}_X)^\perp)$, i.e., $\mu_1 \geq \ldots \geq \mu_d > 0$, as there is no need to include eigenvectors of the covariance operator corresponding to eigenvalue 0. Indeed, one may project the PIDE to $E_0(\mathcal{C}_X)^\perp$ without any error, since the projection of Z_L on $E_0(\mathcal{C}_X)$ is almost surely 0. Define the projection operator

$$
(51) \qquad\qquad \mathcal{P}_d : \begin{cases} H \to U_d \cong \mathbb{R}^d, \\ z \mapsto x := \sum_{l=1}^d \langle z, p_l \rangle_H \, p_l. \end{cases}
$$

We identify U_d with $\mathbb{R}^d$ via the isometry

$$
(52) \qquad\qquad \iota : \begin{cases} (U_d, \|\cdot\|_H) \to (\mathbb{R}^d, \|\cdot\|), \\ x \mapsto (\langle x, p_l \rangle_H)_{l=1}^d. \end{cases}
$$

In particular, we identify $\mathcal{P}_d z$ with the sequence $(\langle z, p_l \rangle_H)_{l=1}^d$.

We introduce the finite-dimensional approximation

$$
(53) \qquad\qquad V_d(t, x) := e^{-rT} E\big[G(x + \mathcal{P}_d Z_L(T - t))\big]
$$

for $x \in U_d \cong \mathbb{R}^d$. We do not assume a finite-dimensional analogue of the regularity assumption (Assumption 2.6), which is hard to verify in practice. Instead, we impose a simple condition on the payoff function.

Assumption 3.4. *Suppose that the payoff function G is Lipschitz continuous on H with Lipschitz constant K_G.*

Remark 3.5. *Assumption 3.4 is not necessarily satisfied for payoffs depending on the exponential of $Z_L(T)$, e.g., a plain call option depending on $S(T)$. However, this can be remedied easily. In the specific case of a call, we can apply a put-call parity. More generally, every payoff can be truncated to a bounded domain (e.g., by multiplying with a smooth cutoff function). A payoff function has finite expectation; hence the error introduced by truncation is arbitrarily small. Since we have to localize the computational domain for any numerical calculation anyway (compare Section 4.2), Assumption 3.4 is no substantial restriction.*

The following theorem shows that Assumption 3.4 is actually enough to recover regularity of V_d.

Theorem 3.6. *The finite-dimensional approximation $V_d : [0, T] \times U_d \to \mathbb{R}$ defined in (53) satisfies $V_d \in C^{1,2}((0, T) \times U_d, \mathbb{R}) \cap C([0, T] \times U_d, \mathbb{R})$. Moreover, the partial derivatives $\partial_{x_i} V_d(t, x)$, $\partial_{x_i} \partial_{x_j} V_d(t, x)$, and $\partial_t V_d(t, x)$ are functions of at most linear growth in $\|x\|$ for $i, j = 1, \ldots, d$.*

Proof. The first step of the proof is to show the existence of a smooth density for the random variable $\mathcal{P}_d Z_L(t)$ for a.e. $t \geq 0$. To achieve this, a fast decay condition for its characteristic function

(54) $$\hat{\mu}_t(x) := E\left(e^{i\langle x, \mathcal{P}_d Z_L(t)\rangle}\right) \in \mathbb{C}$$

is needed. We have

(55) $$E\left[\langle \mathcal{P}_d W_L(t), x_1\rangle_{U_d} \langle \mathcal{P}_d W_L(t), x_2\rangle_{U_d}\right]$$
$$= \sum_{k,l=1}^{d} \langle x_1, p_k\rangle_{U_d} \langle x_2, p_l\rangle_{U_d} E\left[\langle W_L(t), p_k\rangle_H \langle W_L(t), p_l\rangle_H\right] = \langle Q_L x_1, x_2\rangle_H$$

for every $x_1, x_2 \in U_d$. Thus, the covariance operator for the diffusion part of $\mathcal{P}_d Z_L$ is given by $\mathcal{P}_d Q_L \mathcal{P}_d$. The same arguments as in the proof of Theorem 2.3 yield

(56) $$\hat{\mu}_t(x) = \exp\left(-\frac{1}{2}t\langle \mathcal{P}_d Q_L x, x\rangle_{U_d}\right.$$
$$\left. + \int_0^t \int_H \left(e^{i\langle \mathcal{P}_d \zeta, x\rangle_{U_d}} - 1 - i\langle \mathcal{P}_d \zeta, x\rangle_{U_d}\right) \nu_L(d\zeta)ds\right)$$

for every $x \in U_d$. Using Assumptions 2.1 and 2.5, this implies

(57) $$|\hat{\mu}_t(x)| \leq \exp\left(-\frac{1}{2}t\langle \mathcal{P}_d Q_L x, x\rangle_H\right.$$
$$\left. + \int_0^t \int_H \left|e^{i\langle \mathcal{P}_d \zeta, x\rangle_H} - 1 - i\langle \mathcal{P}_d \zeta, x\rangle_H\right| \nu_L(d\zeta)ds\right)$$
$$\leq \exp\left[t\left(-\frac{1}{2}C_1 \|x\|^2 + C_2 \|x\| + C_3\right)\right],$$

with positive constants C_1, C_2, and C_3 depending on d. In particular, we have

(58) $$\lim_{\|x\| \to \infty} \|x\|^n \hat{\mu}_t(x) = 0 \quad \text{for every } n \in \mathbb{N}.$$

Similarly, we obtain

(59) $$|\partial_x^\alpha \hat{\mu}_t(x)| \leq p_\alpha(t, \|x\|) |\hat{\mu}_t(x)| \quad \text{and} \quad |\partial_x^\alpha \partial_t \hat{\mu}_t(x)| \leq q_\alpha(t, \|x\|) |\hat{\mu}_t(x)|$$

for every multiindex $\alpha \in \mathbb{N}_0^d$, where p_α and q_α are polynomials. Consequently, for every $t \in (0, T)$, $\hat{\mu}_t$ and $\partial_t \hat{\mu}_t$ are elements of the Schwartz space

(60) $$\mathcal{S}(\mathbb{R}^d) = \left\{f \in C^\infty(\mathbb{R}^d) : \lim_{\|x\| \to \infty} \|x\|^n \partial^\alpha f(x) = 0 \text{ for every } \alpha \in \mathbb{N}_0^d, n \in \mathbb{N}_0\right\}.$$

From [31, Prop. 28.1], we know that $\mathcal{P}_d Z_L(t)$ has a density $g_t \in C^\infty(\mathbb{R}^d)$, given by

(61) $$g_t(y) := (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-i\langle y, x\rangle} \hat{\mu}_t(x)\, dx.$$

14

Moreover, by the properties of $\hat{\mu}_t$ and [35, Thm. V.2.8], we obtain $g_t \in \mathcal{S}(\mathbb{R}^d)$ and $\partial_t g_t \in \mathcal{S}(\mathbb{R}^d)$. Finally, note that V_d can be written as a convolution of the payoff and the density:

$$(62) \qquad V_d(t,x) = e^{-rT} \int_{\mathbb{R}^d} G(x+y)g_{T-t}(y)dy = e^{-rT} \int_{\mathbb{R}^d} G(y)g_{T-t}(y-x)dy.$$

Due to Assumption 3.4, we have $|G(y)| \leq |G(0)| + K_G \|y\|$. Hence, for $x \in U_d$, $t \in (0,T)$, we may compute

$$(63) \qquad \partial_x^\alpha V_d(t,x) = -e^{-rT} \int_{\mathbb{R}^d} G(x+y)\, \partial_x^\alpha g_{T-t}(y)\, dy$$

and

$$(64) \qquad \partial_t V_d(t,x) = -e^{-rT} \int_{\mathbb{R}^d} G(x+y)\, \partial_t g_{T-t}(y)\, dy$$

for every $\alpha \in \mathbb{N}_0^d$. This proves continuity of the derivatives. In addition, we obtain

$$(65) \qquad \begin{aligned} |\partial_x^\alpha V_d(t,x)| &\leq e^{-rT} \int_{\mathbb{R}^d} |G(y+x)|\, |\partial_x^\alpha g_{T-t}(y)|\, dy \\ &\leq \int_{\mathbb{R}^d} (|G(0)| + K_G \|x\| + K_G \|y\|)\, |\partial_x^\alpha g_{T-t}(y)|\, dy \end{aligned}$$

for every $\alpha \in \mathbb{N}_0^d$. Similarly,

$$(66) \qquad \begin{aligned} |\partial_t V_d(t,x)| &= e^{-rT} \left| \int_{\mathbb{R}^d} G(y)\partial_t g_{T-t}(y-x)dy \right| \\ &\leq \int_{\mathbb{R}^d} (|G(0)| + K_G \|x\| + K_G \|y\|)\, |\partial_t g_{T-t}(y)|\, dy. \end{aligned}$$

Thus, the growth condition is shown.

It remains to prove that V_d is also continuous for $t \to T$. This is, however, a direct consequence of the fact that

$$(67) \qquad \lim_{t \to 0} E\big[\, \|Z_L(t)\|_H \,\big] = 0,$$

and thus

$$(68) \qquad \begin{aligned} |V_d(t,x) - V_d(T,x)| &\leq e^{-rT} E\big[\, |G(x + \mathcal{P}_d Z_L(T-t)) - G(x)| \,\big] \\ &\leq C\, E\big[\, \|Z_L(T-t)\|_H \,\big] \to 0 \quad \text{for } t \to T. \end{aligned}$$

$\square$

Theorem 3.7. *The function V_d defined in (53) is a classical solution of the finite-dimensional PIDE*

$$(69) \qquad \begin{aligned} -\partial_t V_d(t,x) = {}&\frac{1}{2} \sum_{l=1}^d \big[D_x^2 V_d(t,x) \big] (\mathcal{P}_d Q_L p_l, \mathcal{P}_d p_l) \\ &+ \int_H \Big\{ V_d(t, x + \mathcal{P}_d \zeta) - V_d(t,x) - \big[D_x V_d(t,x) \big](\mathcal{P}_d \zeta) \Big\} \nu_L(d\zeta), \end{aligned}$$

with terminal condition

(70)
$$V_d(T, x) = V(T, x) = e^{-rT} G(x),$$

for $t \in (0, T)$, $x \in U_d$.

Proof. The stochastic dynamics of $\mathcal{P}_d Z_L(t)$ are given by

(71)
$$d(\mathcal{P}_d Z_L(t)) = d(\mathcal{P}_d W_L(t)) + \int_H \mathcal{P}_d \zeta \, \widetilde{M}_L(d\zeta, ds).$$

The process $\mathcal{P}_d W_L(t) = \sum_{i=1}^d \langle W_L(t), p_l \rangle_H p_l$ is a d-dimensional Wiener process with correlation operator $\mathcal{P}_d Q_L \mathcal{P}_d$. The integral with respect to $\widetilde{M}_L$ can easily be rewritten as an integral over U_d, since the integrand depends only on the projection $\mathcal{P}_d \zeta \in U_d$.

We apply the finite-dimensional version of Itô's formula (cf., e.g., [11, Thm. 8.18]) to $V_d(t, \mathcal{P}_d Z_L(t))$. In contrast to the Hilbert space-valued case, bounded derivatives are not needed here. The properties of V_d shown in Theorem 3.6 are sufficient. By the same arguments as in the proof of Theorem 2.7, we obtain the following:

(72)
$$
\begin{aligned}
dV_d(t, &\mathcal{P}_d Z_L(t)) = \\
&\partial_t V_d \, dt + \frac{1}{2} \operatorname{Tr}\left([D_x^2 V_d] \mathcal{P}_d Q_L \mathcal{P}_d \right) dt \\
&+ \int_H \left\{ V_d(t, \mathcal{P}_d Z_L(t-) + \mathcal{P}_d \zeta) - V_d(t, \mathcal{P}_d Z_L(t-)) - [D_x V_d](\mathcal{P}_d \zeta) \right\} \nu_L(d\zeta) \, dt \\
&+ D_x V_d \, d(\mathcal{P}_d W_L(t)) + \int_H \left[V_d(t, \mathcal{P}_d Z_L(t-) + \mathcal{P}_d \zeta) - V_d(t, \mathcal{P}_d Z_L(t-)) \right] \widetilde{M}_L(d\zeta, dt).
\end{aligned}
$$

Proceeding exactly as in the proof of Theorem 2.8, we obtain (69). □

The PIDE in Theorem 3.7 is of course nothing more than a projected version of the PIDE (42) for V. The derivatives $D_x V_d(t, x) \in L(\mathbb{R}^d, \mathbb{R})$ and $D_x^2 V_d(t, x) \in L(\mathbb{R}^d, \mathbb{R}^d)$ can be interpreted as a vector and a matrix, respectively. In particular, we have

(73)
$$
\begin{aligned}
\sum_{l=1}^d [D_x^2 V_d(t, x)](\mathcal{P}_d Q_L p_l, \mathcal{P}_d p_l) &= \sum_{l=1}^d \sum_{i,j=1}^d \partial_{x_i} \partial_{x_j} V_d(t, x) \langle \mathcal{P}_d Q_L p_l, p_j \rangle_H \langle \mathcal{P}_d p_l, p_i \rangle_H \\
&= \sum_{i,j=1}^d \langle Q_L p_i, p_j \rangle_H \partial_{x_i} \partial_{x_j} V_d(t, x).
\end{aligned}
$$

To simplify notation, we define coefficients

(74)
$$a_{ij} := \frac{1}{2} \langle Q_L p_i, p_j \rangle_H.$$

The PIDE can thus be written as

(75)
$$
\begin{aligned}
-\partial_t V_d(t, x) = &\sum_{i,j=1}^d a_{ij} \, \partial_{x_i} \partial_{x_j} V_d(t, x) \\
&+ \int_H \left\{ V_d(t, x + \mathcal{P}_d \zeta) - V_d(t, x) - \sum_{i=1}^d \langle \zeta, p_i \rangle_H \, \partial_{x_i} V_d(t, x) \right\} \nu_L(d\zeta).
\end{aligned}
$$

Moreover, we have the following ellipticity property.

Theorem 3.8. *The matrix $(a_{ij})_{i,j=1}^{d}$ is symmetric positive definite.*

Proof. This is a direct consequence of Assumption 2.5, since

$$(76) \qquad \sum_{i,j=1}^{d} y_i\, a_{ij}\, y_j = \frac{1}{2}\left\langle Q_L \sum_{i=1}^{d} y_i p_i, \sum_{i=1}^{d} y_i p_i \right\rangle_H$$

for every $y \in \mathbb{R}^d$, and $\|\sum_{i=1}^{d} y_i p_i\|_H = \|y\|$ due to the isometry (52). $\qquad\square$

3.3 Variational formulation and uniqueness

We have already shown that the approximation $V_d(t,x)$ is a classical solution of the finite-dimensional PIDE (75). In this section, we introduce the corresponding variational formulation in appropriate Hilbert spaces and show uniqueness of the weak solution. Since the payoff is not necessarily bounded, and thus not an element of $L^2(\mathbb{R}^d)$, we use weighted Sobolev spaces instead. Let ρ_θ be the weight function with exponential decay defined by

$$(77) \qquad \rho_\theta : \begin{cases} \mathbb{R}^d \to \mathbb{R}, \\ x \mapsto e^{-\theta\sqrt{1+\|x\|^2}}, \end{cases}$$

with a parameter $\theta > 0$. We define the scalar products

$$(78) \qquad \langle \psi, \varphi \rangle_{L^{2,\theta}} := \int_{\mathbb{R}^d} \psi(x)\varphi(x)\rho_\theta(x)\, dx$$

and

$$(79) \qquad \langle \psi, \varphi \rangle_{H^{k,\theta}} := \sum_{\alpha \in \mathbb{N}_0^d, |\alpha| \le k} \langle \partial^\alpha \psi, \partial^\alpha \varphi \rangle_{L^{2,\theta}}$$

for functions $\psi, \varphi : \mathbb{R}^d \to \mathbb{R}$. The corresponding Hilbert spaces are denoted by $L^{2,\theta}(\mathbb{R}^d)$ and $H^{k,\theta}(\mathbb{R}^d)$ (cf., e.g., [2, Chap. 3.1]). In particular, we consider the Gelfand triplet

$$(80) \qquad H^{1,\theta}(\mathbb{R}^d) \hookrightarrow L^{2,\theta}(\mathbb{R}^d) \hookrightarrow \left(H^{1,\theta}(\mathbb{R}^d)\right)'.$$

Finally, we define the bilinear form

(81)

$$a(\psi, \varphi) := \int_{\mathbb{R}^d} \sum_{i,j=1}^{d} a_{ij}\, \partial_{x_i}\psi(x)\, \partial_{x_j}\varphi(x)\, \rho_\theta(x)\, dx + \int_{\mathbb{R}^d} \sum_{i=1}^{d} b_i(x)\, \partial_{x_i}\psi(x)\, \varphi(x)\, \rho_\theta(x)\, dx$$

$$- \int_H \int_{\mathbb{R}^d} \left[\psi(x + \mathcal{P}_d\zeta) - \psi(x) - \sum_{i=1}^{d} \langle \zeta, p_i \rangle_H\, \partial_{x_i}\psi(x) \right] \varphi(x)\, \rho_\theta(x)\, dx\, \nu_L(d\zeta)$$

for $\psi, \varphi \in \mathrm{H}^{1,\theta}(\mathbb{R}^d)$, where

$$(82) \qquad b_i(x) := -\frac{\theta \sum_{j=1}^d a_{ij} x_j}{\sqrt{1 + \|x\|^2}} \quad \text{for } i = 1, \dots, d.$$

The coefficients $b_i(x)$ satisfy

$$(83) \qquad |b_i(x)| \leq \theta\, d \max_{j=1,\dots,d} |a_{ij}|$$

and are therefore bounded on $\mathbb{R}^d$.

We can now state a variational form of (75).

Theorem 3.9. *The function V_d defined in (53) satisfies*

$$(84) \qquad -\langle \partial_t V_d(t, \cdot), \varphi \rangle_{\mathrm{L}^{2,\theta}} + a(V_d(t, \cdot), \varphi) = 0$$

for every $\varphi \in \mathrm{H}^{1,\theta}(\mathbb{R}^d)$ and a.e. $t \in (0, T)$, with terminal condition

$$(85) \qquad V_d(T, x) = G(x).$$

Proof. We first note that $V_d(t, \cdot) \in \mathrm{H}^{2,\theta}(\mathbb{R}^d)$ and $\partial_t V_d(t, \cdot) \in \mathrm{L}^{2,\theta}(\mathbb{R}^d)$ hold for every $t \in (0, T)$ due to Theorem 3.6. Starting with (75), partial integration yields

(86)
$$-\langle \partial_t V_d(t, x), \varphi \rangle_{\mathrm{L}^{2,\theta}}$$
$$= -\sum_{i,j=1}^d a_{ij} \int_{\mathbb{R}^d} \partial_{x_i} V_d(t, x)\, \partial_{x_j}(\varphi \rho_\theta)(x)\, dx$$
$$+ \int_H \int_{\mathbb{R}^d} \left[V_d(t, x + \mathcal{P}_d \zeta) - V_d(t, x) - \sum_{i=1}^d \langle \zeta, p_i \rangle_H \partial_{x_i} V_d(t, x) \right] \varphi(x) \rho_\theta(x)\, dx\, \nu_L(d\zeta).$$

Using the product rule, we obtain

$$(87) \qquad \sum_{j=1}^d a_{ij} \partial_{x_j}(\varphi \rho_\theta)(x) = \sum_{j=1}^d a_{ij} \partial_{x_j} \varphi(x)\, \rho_\theta(x) + \varphi(x)\, \frac{-\theta \sum_{j=1}^d a_{ij} x_j}{\sqrt{1 + \|x\|^2}}\, \rho_\theta(x).$$

$\square$

Theorem 3.10. *The bilinear form a defined in (81) is continuous and satisfies Gårding's inequality. More precisely, there are constants $C > 0$, $c_1 \geq 0$, and $c_2 > 0$ (possibly depending on d) such that*

$$(88) \qquad |a(\psi, \varphi)| \leq C \|\psi\|_{\mathrm{H}^{1,\theta}} \|\varphi\|_{\mathrm{H}^{1,\theta}}$$

and

$$(89) \qquad a(\psi, \psi) + c_1 \|\psi\|_{\mathrm{L}^{2,\theta}}^2 \geq c_2 \|\psi\|_{\mathrm{H}^{1,\theta}}^2$$

hold for every $\psi, \varphi \in \mathrm{H}^{1,\theta}(\mathbb{R}^d)$.

Proof. First, we show continuity. From the definition of a, we obtain

(90)
$$|a(\psi, \varphi)|$$
$$\leq \sum_{i,j=1}^{d} |a_{ij}| \int_{\mathbb{R}^d} |\partial_{x_i}\psi(x)||\partial_{x_j}\varphi(x)|\rho_\theta(x)\,dx + \sum_{i=1}^{d} \int_{\mathbb{R}^d} |b_i(x)||\partial_{x_i}\psi(x)||\varphi(x)|\rho_\theta(x)\,dx$$
$$+ \int_H \int_{\mathbb{R}^d} |\psi(x + \mathcal{P}_d\zeta)||\varphi(x)|\rho_\theta(x)\,dx\,\nu_L(d\zeta) + \int_H \int_{\mathbb{R}^d} |\psi(x)||\varphi(x)|\rho_\theta(x)\,dx\,\nu_L(d\zeta)$$
$$+ \sum_{i=1}^{d} \int_H \int_{\mathbb{R}^d} |\langle \zeta, p_i \rangle_H||\partial_{x_i}\psi(x)||\varphi(x)|\rho_\theta(x)\,dx\,\nu_L(d\zeta).$$

The Cauchy–Schwarz inequality yields

(91)
$$|a(\psi, \varphi)| \leq \max_{i,j=1,\dots,d} |a_{ij}| \sum_{i,j=1}^{d} \|\partial_{x_i}\psi\|_{L^{2,\theta}} \|\partial_{x_j}\varphi\|_{L^{2,\theta}}$$
$$+ \max_{i=1,\dots,d} \|b_i\|_{L^\infty} \sum_{i=1}^{d} \|\partial_{x_i}\psi\|_{L^{2,\theta}} \|\varphi\|_{L^{2,\theta}} + 2 \int_H \|\psi\|_{L^{2,\theta}} \|\varphi\|_{L^{2,\theta}}\,\nu_L(d\zeta)$$
$$+ \int_H \sum_{i=1}^{d} |\langle \zeta, p_i \rangle_H| \|\partial_{x_i}\psi\|_{L^{2,\theta}} \|\varphi\|_{L^{2,\theta}}\,\nu_L(d\zeta).$$

Due to

(92)
$$\int_H \nu_L(d\zeta) < \infty \quad \text{and} \quad \int_H \|\zeta\|_H\,\nu_L(d\zeta) < \infty,$$

this proves (88).

For the proof of Gårding's inequality, we start with the ellipticity property from Theorem 3.8. For every $\zeta \in \mathbb{R}^d$,

(93)
$$\sum_{i,j=1}^{d} a_{ij}\,\zeta_i\,\zeta_j \geq c \sum_{i=1}^{d} \zeta_i^2$$

holds, with a constant $c > 0$. Hence

(94)
$$c \int_{\mathbb{R}^d} \sum_{i=1}^{d} |\partial_{x_i}\psi(x)|^2 \rho_\theta(x)\,dx \leq \int_{\mathbb{R}^d} \sum_{i,j=1}^{d} a_{ij}\,\partial_{x_i}\psi(x)\partial_{x_j}\psi(x)\,\rho_\theta(x)\,dx$$
$$= a(\psi, \psi) - \int_{\mathbb{R}^d} \sum_{i=1}^{d} b_i(x)\,\partial_{x_i}\psi(x)\,\psi(x)\,\rho_\theta(x)\,dx$$
$$+ \int_H \int_{\mathbb{R}^d} \left[\psi(x + \mathcal{P}_d\zeta) - \psi(x) - \sum_{i=1}^{d} \langle \zeta, p_i \rangle_H\,\partial_{x_i}\psi(x) \right] \psi(x)\,\rho_\theta(x)\,dx\,\nu_L(d\zeta).$$

19

The same calculations as in the proof of continuity above yield

$$
(95) \qquad
\begin{aligned}
c \sum_{i=1}^{d} \|\partial_{x_i} \psi\|_{L^{2,\theta}}^2 &\le a(\psi, \psi) + C_1 \sum_{i=1}^{d} \|\partial_{x_i} \psi\|_{L^{2,\theta}} \|\psi\|_{L^{2,\theta}} + C_2 \|\psi\|_{L^{2,\theta}}^2 \\
&\le a(\psi, \psi) + C_1 \left(\frac{\varepsilon}{2} \sum_{i=1}^{d} \|\partial_{x_i} \psi\|_{L^{2,\theta}}^2 + \frac{d}{2\varepsilon} \|\psi\|_{L^{2,\theta}}^2 \right) + C_2 \|\psi\|_{L^{2,\theta}}^2,
\end{aligned}
$$

where we have used Young's inequality in the last estimate. Choosing ε so small that

$$
(96) \qquad C_1 \frac{\varepsilon}{2} \le \frac{1}{2} c
$$

and setting

$$
(97) \qquad c_1 = \frac{C_1 d}{2\varepsilon} + C_2 \quad \text{and} \quad c_2 = \frac{1}{2} c
$$

yields (89). □

With the following two theorems, we show that V_d is indeed the unique solution of the PIDE. We start with a lemma requiring stronger regularity hypotheses for G. Afterwards, we give a result for arbitrary Lipschitz continuous payoffs.

Lemma 3.11. *Suppose that $G \in H^{2,\theta}(\mathbb{R}^d)$ has bounded first and second derivatives. Then V_d is the unique solution of (84), with terminal condition (85), in the space $\mathcal{W}(0,T)$ defined by*

$$
(98) \qquad \mathcal{W}(0,T) := \left\{ f : \mathbb{R}^d \to \mathbb{R} : f \in L^2(0,T; H^{1,\theta}), \, \partial_t f \in L^2(0,T; (H^{1,\theta}(\mathbb{R}^d))') \right\}.
$$

Proof. The bilinear form a is continuous and satisfies Gårding's inequality by Theorem 3.10. Therefore, the PIDE has a unique solution in the space $\mathcal{W}(0,T)$ by [37, Thm. 26.1]. On the other hand, we know from Theorem 3.7 that V_d satisfies (84). It remains to prove that $V_d \in \mathcal{W}(0,T)$.

Using the same notation as in the proof of Theorem 3.6, we have

$$
(99) \qquad V_d(t,x) = e^{-rT} \int_{\mathbb{R}^d} G(x+y) g_{T-t}(y) dy,
$$

where $g_{T-t} \in \mathcal{S}(\mathbb{R}^d)$ is the density of $\mathcal{P}_d Z_L(T-t)$. Consequently,

$$
(100) \qquad \int_{\mathbb{R}^d} V_d^2(t,x) \rho_\theta(x) \, dx \le \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} (C_1 + C_2 \|x\| + C_3 \|y\|) g_{T-t}(y) \, dy \right)^2 \rho_\theta(x) \, dx.
$$

Since

$$
(101) \qquad \int_{\mathbb{R}^d} \|x\|^i \rho_\theta(x) \, dx < \infty \quad (i \in \{1,2\}), \qquad \int_{\mathbb{R}^d} g_{T-t}(y) \, dy = 1,
$$

and

$$
(102) \qquad \int_{\mathbb{R}^d} \|y\| \, g_{T-t}(y) \, dy = E\left[\|\mathcal{P}_d Z_L(T-t)\| \right] < \infty
$$

hold for every $t \in (0,T)$, this implies $V_d \in L^2(0,T;L^{2,\theta}(\mathbb{R}^d))$. Moreover, we have

$$(103) \qquad \partial_x^\alpha V_d(t,x) = e^{-rT} \int_{\mathbb{R}^d} \partial^\alpha G(x+y)\, g_{T-t}(y)\, dy$$

for every $\alpha \in \mathbb{N}_0^d$, $|\alpha| \in \{1,2\}$. Due to the boundedness of the derivatives of G, the following holds:

$$(104) \qquad \int_{\mathbb{R}^d} \left(\partial_x^\alpha V_d(t,x) \right)^2 \rho_\theta(x)\, dx \leq C \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} g_{T-t}(y)\, dy \right)^2 \rho_\theta(x)\, dx = C \int_{\mathbb{R}^d} \rho_\theta(x)\, dx.$$

Consequently, $V_d \in L^2(0,T;H^{2,\theta}(\mathbb{R}^d))$. Since V_d satisfies the PIDE (75), its time derivative $\partial_t V_d$ can be expressed in terms of its first and second spatial derivatives. Thus, we also have $\partial_t V_d \in L^2(0,T;L^{2,\theta}(\mathbb{R}^d))$, and the proof is finished. $\qquad \square$

Theorem 3.12. *Let G satisfy Assumption 3.4. Then V_d is an element of the space $\mathcal{W}(0,T)$ defined in Lemma 3.11 and the unique solution of (84) with terminal condition (85).*

Proof. The key of the proof is to approximate G with a sequence of smooth functions to which Lemma 3.11 can be applied. To this end, let $(\psi_n)_{n \in \mathbb{N}} \subset C_0^\infty(\mathbb{R}^d)$ be a sequence of standard mollifiers with compact support. Define the convolution

$$(105) \qquad G_n(x) := \int_{\mathbb{R}^d} G(x-y)\psi_n(y)dy \in C^\infty(\mathbb{R}^d), \quad x \in \mathbb{R}^d.$$

Since G is by assumption Lipschitz, and thus uniformly continuous, the following uniform approximation property holds by [16, Thm. C.6]:

$$(106) \qquad \forall \varepsilon > 0 \ \exists N \in \mathbb{N} : |G_n(x) - G(x)| \leq \varepsilon \quad \text{for every } n \geq N, x \in \mathbb{R}^d.$$

Moreover, for every $\xi \in \mathbb{R}^d$ and every multiindex $\alpha \in \mathbb{N}_0^d$, we have

$$(107) \qquad \begin{aligned} |\partial^\alpha G_n(x+\xi) - \partial^\alpha G_n(x)| &\leq \int_{\mathbb{R}^d} |G(x+\xi-y) - G(x-y)|\, |\partial^\alpha \psi_n(y)|\, dy \\ &\leq K_G \|\xi\| \int_{\mathbb{R}^d} |\partial^\alpha \psi_n(y)|\, dy. \end{aligned}$$

Thus, in particular, the first and second derivatives of G_n are bounded for every $n \in \mathbb{N}$.

Now let

$$(108) \qquad V_d^n(t,x) := e^{-rT} E\left[G_n(x + \mathcal{P}_d Z_L(T-t)) \right]$$

be the price function associated with payoff G_n. By Lemma 3.11, $V_d^n \in \mathcal{W}(0,T)$ is the unique solution of (75) with terminal condition $V_d^n(T,x) = G_n(x)$. Moreover, the PIDE with terminal value $G(x)$ also has a unique solution, which we denote by $\widetilde{V}_d \in \mathcal{W}(0,T)$. From [37, Thm. 26.1], we obtain

$$(109) \qquad \left\| V_d^n - \widetilde{V}_d \right\|_{L^2(0,T;L^{2,\theta})} \leq C \left\| G_n - G \right\|_{L^{2,\theta}} \to 0 \quad \text{for } n \to \infty.$$

On the other hand, by the proof of Lemma 3.11, we have $V_d \in L^2(0, T; L^{2,\theta})$ for every payoff G satisfying Assumption 3.4. Thus, using the notation from the proof of Theorem 3.6, we get

$$
\begin{aligned}
(110) \quad & \|V_d^n - V_d\|_{L^2(0,T;L^{2,\theta})}^2 \\
&= e^{-rT} \int_0^T \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \left(G(x+y) - G_n(x+y) \right) g_{T-t}(y) dy \right)^2 \rho_\theta(x) \, dx \, dt \\
&\leq e^{-rT} \int_0^T \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |G(x+y) - G_n(x+y)| \, g_{T-t}(y) dy \right)^2 \rho_\theta(x) \, dx \, dt \\
&\to 0 \quad \text{for } n \to \infty.
\end{aligned}
$$

Combining (109) and (110), we obtain $V_d = \widetilde{V}_d$. $\qquad\square$

3.4 CONVERGENCE OF FINITE-DIMENSIONAL APPROXIMATION

We are interested in the convergence of the finite-dimensional approximation V_d to the true price function V. The fair price of the option is given by $V(0,0)$, since we assume the Lévy process Z_L starts in 0. In particular, we are looking for a pointwise convergence result.

Theorem 3.13. *Let $\mu_1 \geq \mu_2 \geq ... \geq 0$ be the eigenvalues of the covariance operator $\mathcal{C}_X$ defined in (26). Then there exists a constant $C > 0$ such that*

$$
(111) \qquad |V_d(0,0) - V(0,0)| \leq C \sqrt{\sum_{l=d+1}^{\dim H} \mu_l}.
$$

Proof. By definition of V and V_d, we have

$$
\begin{aligned}
(112) \quad |V_d(0,0) - V(0,0)| &= e^{-rT} \left| E \left[G(Z_L(T)) - G(\mathcal{P}_d Z_L(T)) \right] \right| \\
&\leq e^{-rT} E \left[K_G \| Z_L(T) - \mathcal{P}_d Z_L(T) \|_H \right].
\end{aligned}
$$

Since $\|\cdot\|_{L^1} \leq C \|\cdot\|_{L^2}$ for finite measure spaces, we may apply Theorem 3.3 to obtain

$$
(113) \quad |V_d(0,0) - V(0,0)| \leq C \sqrt{E \left[\| Z_L(T) - \mathcal{P}_d Z_L(T) \|_H^2 \right]} = C \sqrt{\sum_{l=d+1}^{\dim H} \mu_l}.
$$

$\qquad\square$

Depending on the properties of the covariance operator $\mathcal{C}_X$, bounds for the decay of the eigenvalues μ_l can be found. To this end, we define the kernel

$$
(114) \qquad K : \begin{cases} D \times D \to \mathbb{R}, \\ (u,v) \mapsto E \left[Z_L(T,u) Z_L(T,v) \right], \end{cases}
$$

where as before $D \subset \mathbb{R}^m$ and $H = L^2(D, \mu_D)$. Then K is indeed the kernel of the covariance operator $\mathcal{C}_X$, since by Fubini's theorem

$$
\begin{aligned}
(115) \quad \int_D \int_D K(u,v) h_1(u)\, \mu_D(du)\, h_2(v)\, \mu_D(dv) &= E\big[\langle Z_L(T), h_1\rangle_H \langle Z_L(T), h_2\rangle_H\big] \\
&= \langle \mathcal{C}_X h_1, h_2 \rangle_H
\end{aligned}
$$

for every $h_1, h_2 \in H$. Consequently,

$$
(116) \qquad \mathcal{C}_X h_1(\cdot) = \int_D K(\cdot, v) h_1(v)\, \mu_D(dv).
$$

Moreover, K is an element of $L^2(D \times D)$, since

$$
\begin{aligned}
(117) \quad \int_D \int_D K^2(u,v)\, \mu_D(du)\, \mu_D(du) &\leq \int_D \int_D E\big[Z_L(T,u)^2\big] E\big[Z_L(T,v)^2\big]\, \mu_D(du)\, \mu_D(du) \\
&= E\big[\, \|X_L(T) - \gamma_L\, T\|_H^2\,\big]^2 < \infty
\end{aligned}
$$

by Theorem 2.2. We now give a result for the eigenvalue decay, depending on the properties of K.

Theorem 3.14. *Let $D \subset \mathbb{R}^m$ be a bounded Borel set and μ_D be the Lebesgue measure. If the kernel K is piecewise $H^k \otimes L^2$ on $D \times D$ for a $k \in \mathbb{N}$, then there exists a constant C such that*

$$
(118) \qquad \mu_l \leq C\, l^{-\frac{k}{m}} \quad \text{for every } l \in I.
$$

Moreover, if the kernel K is piecewise analytic, then there are constants C_1, C_2 such that

$$
(119) \qquad \mu_l \leq C_1 e^{-C_2 l^{\frac{1}{m}}} \quad \text{for every } l \in I.
$$

Proof. These results are shown in [32, Props. 2.18 and 2.21]. $\qquad\square$

Remark 3.15. *A precise definition of "piecewise $H^k \otimes L^2$" as used in the hypothesis of Theorem 3.14 is given in [34, Def. 3.1]. The hypothesis of Theorem 3.14 is fulfilled for the jump-diffusion model if the drift γ, the volatility σ, the Brownian covariance operator Q, and the jump-dampening factors η satisfy corresponding piecewise smoothness criteria.*

4 NUMERICAL METHODS

In this section, methods for the numerical solution of the European option pricing problem are described. In particular, the computation of a suitable basis for the dimension reduction described in the previous section is addressed. Moreover, we give a give a short overview of the techniques needed to solve the resulting finite-dimensional PIDE.

4.1 COMPUTATION OF THE POD BASIS

As we have seen in Section 3.1, the construction of a POD basis for $X(T)$ can be reduced to the eigenvalue problem

$$(120) \qquad \mathcal{C}_X p_l = \mu_l p_l, \quad l \in I,$$

with the covariance operator $\mathcal{C}_X : H \to H' \cong H = L^2(D, \mu_D)$ defined in (26). In general, the eigenvectors p_l are not known analytically. However, it is possible to compute good approximations numerically. For finite index sets D (basket options), the solution of the symmetric eigenvalue problem can be performed by standard methods (e.g., QR-algorithm). For all subsequent results, we will therefore assume the more complicated setting that $D \subset \mathbb{R}^m$ is a bounded Borel set and μ_D is the Lebesgue measure. This holds, e.g., for electricity swaptions (with $m = 1$). In this case, some further approximation is needed to obtain a finite-dimensional problem eventually. To this end, we employ a finite element discretization and solve a projected eigenvalue problem. Convergence results for this technique can be shown under certain regularity conditions on the covariance kernel K defined in (114). The general theory of Galerkin approximations of Karhunen–Loève expansions is discussed in [32].

Let $U_{\Delta x} \subset H = L^2(D, \mu_D)$ be a finite element subspace with discretization parameter Δx. For $l = 1, \ldots, \dim U_{\Delta x}$, the Galerkin approximations $(\mu_l^{\Delta x}, p_l^{\Delta x}) \subset \mathbb{R} \times U_h$ of the eigenpairs $(\mu_l, p_l) \subset \mathbb{R} \times H$ are, by definition, solutions of the following problem:

$$(121) \qquad \begin{aligned} \forall \varphi \in U_{\Delta x} : \left\langle \mathcal{C}_X p_l^{\Delta x}, \varphi \right\rangle_H &= \int_D \left(\int_D K(u, v) p_l^{\Delta x}(v)\, \mu_D(dv) \right) \varphi(u)\, \mu_D(du) \\ &\overset{!}{=} \mu_l^{\Delta x} \left\langle p_l^{\Delta x}, \varphi \right\rangle_H. \end{aligned}$$

This is equivalent to the eigenvalue problem

$$(122) \qquad \mathcal{P}_{\Delta x} \mathcal{C}_X \mathcal{P}_{\Delta x}\, p_l^{\Delta x} = \mu_l^{\Delta x} p_l^{\Delta x},$$

where $\mathcal{P}_{\Delta x} : H \to U_{\Delta x}$ is the projection operator onto the finite element subspace. The operator $\mathcal{P}_{\Delta x} \mathcal{C}_X \mathcal{P}_{\Delta x}$ is self-adjoint and compact due to the properties of the projection and Theorem 2.4. The following theorem gives an error bound for the approximation of $Z_L(T)$ obtained with this Galerkin method.

Theorem 4.1. *Let the covariance kernel K defined in (114) be a piecewise smooth function. Further, let $U_{\Delta x} \subset H$ be a finite element space of piecewise polynomials of degree $q \in \mathbb{N}$, where Δx denotes the mesh width of the regular triangulation. Finally, let μ_l be the true eigenvalues of the covariance operator $\mathcal{C}_X$, and let $p_l^{\Delta x}$ be orthonormal solutions of the projected eigenvalue problem (122). Then there exists a constant C such that*

$$(123) \qquad E\left[\left\| Z_L(T) - \sum_{l=1}^d \langle Z_L(T), p_l^{\Delta x} \rangle_H p_l^{\Delta x} \right\|_H^2 \right] \leq C \Delta x^{2q+1} + \sum_{l=d+1}^{\dim H} \mu_l$$

holds for every $d \leq \dim U_{\Delta x}$.

Proof. The estimate is taken from [32, Prop. 3.3]. The necessary assumption [32, Ass. 3.1] is satisfied due to the piecewise smoothness of the kernel and [34, Thm. 1.5] □

Note that all the results from Sections 3.2 and 3.3 are still valid when we use $(p_l^{\Delta x})_{l=1}^d$ instead of $(p_l)_{l=1}^d$ if we replace Assumption 2.5 with the following hypothesis:

$$\tag{124} \langle Q_L h, h \rangle_H > 0 \quad \text{for every } h \in \text{span}\{p_l^{\Delta x}, l = 1, \dots, d\} \setminus \{0\}.$$

In this case, we denote the unique solution of the corresponding finite-dimensional PIDE (projected to $\text{span}\{p_l^{\Delta x}, l = 1, \dots, d\}$) by $V_d^{\Delta x}$.

Corollary 4.2. *Let the hypotheses of Theorem 4.1 and (124) hold. Then there exists a constant $C > 0$ such that*

$$\tag{125} \left| V_d^{\Delta x}(0,0) - V(0,0) \right| \le C \sqrt{(\Delta x)^{2q+1} + \sum_{l=d+1}^{\dim H} \mu_l}.$$

Proof. The estimate follows from Theorem 4.1 by exactly the same arguments as in the proof of Theorem 3.13. □

4.2 Numerical PIDE solution

In this subsection, we give an overview of the numerical methods for solving the finite-dimensional PIDE (75). We will not go into details about convergence results for PIDE solvers but refer to recent literature instead. Finite difference methods for integro-differential equations are analyzed, e.g., in [12, 30], finite elements and wavelet compression techniques are described in [24, 36, 28].

Localization   We consider the PIDE (75) whose spatial domain is the whole of $\mathbb{R}^d$. The first step towards a numerical solution is therefore the localization to a bounded domain. To this end, we restrict the spatial part of the equation to a d-dimensional cuboid

$$\tag{126} \Omega_R := [-R_1, R_1] \times [-R_2, R_2] \times \dots \times [-R_d, R_d].$$

This simple domain can be described by a single vector $R = (R_1, \dots, R_d) \in \mathbb{R}^d$. The probabilistic interpretation of this truncation is that we price a barrier option whose value is set to 0 if the process Z_L leaves Ω_R at any time before maturity. Under polynomial growth conditions for the payoff function G, one can show that the truncation error decreases exponentially with $\|R\|$ (cf., e.g., [36, Thm. 3.3.2]). The error is of course higher if the solution is evaluated closer to the boundary of Ω_R.

Analysis of the dependence of the localization error on the truncation parameter in one-dimensional problems shows that good results are achieved for R_1 greater than a

certain multiple of the standard deviation of the jump-diffusion process at time T (cf., e.g., [11, Fig. 12.1]). By construction of the POD basis $(p_l)_{l=1}^d$, we have

(127) $$\langle \mathcal{C}_X p_l, p_l \rangle_H = \mu_l \quad \text{for } l = 1, \ldots, d.$$

Hence, the eigenvalues μ_l describe the variance in direction of the POD vectors. This suggests an adaptive truncation strategy, setting

(128) $$R_l = C\sqrt{\mu_l}, \quad l = 1, \ldots, d,$$

with a constant $C > 0$. This heuristic choice accounts for the decreasing variance in the coordinate directions (compare Theorem 3.14). It results in smaller domains Ω_R (compared to cubic domains) and allows for more accurate discretization results using the same number of grid points. We introduce artificial homogeneous Dirichlet boundary conditions and set

(129) $$V_d^h(t, x) = 0 \quad \text{for every } t \in [0, T], \ x \in \partial\Omega_R,$$

where $\partial\Omega_R$ is the boundary of the domain.

DISCRETIZATION   We will solve the PIDE using a vertical method of lines; i.e., we first discretize the spatial operators and apply some time stepping for ordinary differential equations afterwards. Since we have already derived the variational formulation of the PIDE in Theorem 3.9, we can directly apply a finite element method to approximate the spatial derivatives in (84). Usually, finite elements have several advantages compared to finite differences. In particular, they allow for an easy discretization of geometrically complex domains, adaptive refinement, and higher-order approximations. Moreover, the theory of weak solutions allows for lower regularity assumptions than the classical differential operator.

However, in the specific setting of option pricing, these arguments are only partly valid. First, we may choose the shape of our computational domain arbitrarily due to localization. As described in the previous paragraph, we truncate the domain to a d-dimensional cuboid. Second, we have already proven that V_d is a smooth classical solution to (75). Moreover, despite the simple shape of the domain, a significant overhead is needed to compute and store the geometric information about finite elements.

Hence, we apply finite differences for the numerical experiments in Section 5. We define a regular but anisotropic grid G_α on Ω_R. The grid is described by a multiindex $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, and the mesh size in each coordinate is given by $h_i := 2R_i/2^{\alpha_i}$, $i = 1, \ldots, d$. The grid contains the points

(130) $$x(\beta) := \left(-R_i + \beta_i h_i\right)_{i=1}^d \in \mathbb{R}^d, \quad \beta \in \{0, 1, \ldots, 2^{\alpha_1}\} \times \cdots \times \{0, 1, \ldots, 2^{\alpha_d}\}.$$

The corresponding discretized subspace is denoted by $U_d^h \subset U_d \subset H$. The partial derivatives are approximated by central differences as follows:

(131) $$\partial_{x_i} V_d(t, x(\beta)) \approx \frac{1}{2h_i}\left[V_d\big(t, x(\beta + e_i)\big) - V_d\big(t, x(\beta - e_i)\big)\right],$$

where e_i is the ith canonical unit vector. For the second derivatives, we have
(132)
$$\partial_{x_j} \partial_{x_i} V_d(t, x(\beta)) \approx \begin{cases} \frac{1}{4h_i h_j} \Big[V_d(t, x(\beta + e_i + e_j)) - V_d(t, x(\beta + e_i - e_j)) \\ \qquad\quad - V_d(t, x(\beta - e_i + e_j)) + V_d(t, x(\beta - e_i - e_j)) \Big], & i \neq j, \\ \frac{1}{h_i^2} \Big[V_d(t, x(\beta + e_i)) - 2V_d(t, x(\beta)) + V_d(t, x(\beta - e_i)) \Big], & i = j. \end{cases}$$

The discretization of the nonlocal integral term will be addressed separately in the next paragraph.

After applying the finite difference scheme for a sparse grid approximation in space, the resulting system of ordinary differential equations is solved with an appropriate time stepping scheme. Since the differential part of the PIDE is of parabolic type, we choose a discontinuous Galerkin method of order 1 for this purpose. Details on the topic (in particular error estimates) can be found in [33, Chap. 12]. Defining a partition $0 = t_0 < \cdots < t_{n_T} = T$ of $[0, T]$, we calculate a solution in the space

(133) $\qquad S_d^h := \{ v \in \mathrm{L}^2(0, T; U_d^h); \ v\chi_{[t_{m-1}, t_m]} \in \Pi_1(t_{m-1}, t_m; U_d^h), \ m = 1, \ldots, n_T \},$

where $\Pi_1(t_{m-1}, t_m; U_d^h)$ is the space of polynomials of degree at most 1 having values in U_d^h. This method yields a stable algorithm, allowing for large time steps even in the presence of convection terms.

NONLOCAL INTEGRAL TERMS   One of the main difficulties when solving the PIDE is the nonlocal nature of the integral term. In contrast to differential operators, this term involves the solution on the whole of $\mathbb{R}^d$. Since we have already introduced artificial zero boundary conditions, the solution can easily be continued with 0 outside the truncated domain Ω_R. However, numerical quadrature formulas will in general lead to full matrices. This is contradictory to the essential use of sparse matrices even for problems on relatively low-dimensional spaces. An efficient way to reduce the number of nonzero matrix entries is by using wavelet compression schemes. These make use of the fast decline of entries corresponding to wavelet basis functions with larger distance of their supports. Entries close to 0 are then discarded (cf. the references given at the beginning of this section).

Wavelets are the method of choice if the jump distributions in the additive model (7) cover large parts of the domain. However, the examples we are going to examine in Section 5 are of type (2), i.e., Hilbert space-valued jump-diffusions with scalar driving processes. Here, at every time $t \in [0, T]$, jumps are restricted to one-dimensional subspaces, spanned by $\eta_k(t) \in H$. If the number of driving jump processes is low, a different approach is feasible, too. Instead of wavelet compression, direct numerical quadrature (Gauß or Newton–Cotes method) can be applied. We will now have a closer look at this specific case.

Let $0 = t_0 < t_1 < \ldots < t_{N_T} = T$, $N_T \in \mathbb{N}$, be the time discretization grid. It is not necessary to compute the Lévy measure ν_L defined in (18) explicitly. Instead, we

define measures

$$(134) \qquad \nu_L(t_k, t_l; B) := \frac{1}{t_l - t_k} \int_{t_k}^{t_l} \int_H \chi_B(\eta(s, \xi)) \nu(d\xi) \, ds,$$

$$(135) \qquad \nu_L(t; B) := \int_H \chi_B(\eta(t, \xi)) \nu(d\xi)$$

for any $B \subset \mathcal{B}(H)$, $0 \le k < l \le n_T$ and $t \ge 0$. For each time step $[t_k, t_{k+1}]$ in the method of lines, we use integrals with respect to $\nu_L(t_k, t_{k+1}; \cdot)$, i.e., the equivalent Lévy measure for the current time interval. We are looking for an approximation of the integral

$$\int_H \left[V_d(t, x + \mathcal{P}_d \zeta) - V_d(t, x) - \sum_{i=1}^d \langle \zeta, p_i \rangle_H \, \partial_{x_i} V_d(t, x) \right] \nu_L(t_k, t_{k+1}; d\zeta)$$

$$(136) \qquad = \int_H V_d(t, x + \mathcal{P}_d \zeta) \nu_L(t_k, t_{k+1}; d\zeta) - \lambda V_d(t, x)$$

$$- \sum_{i=1}^d \partial_{x_i} V_d(t, x) \int_H \langle \zeta, p_i \rangle_H \, \nu_L(t_k, t_{k+1}; d\zeta).$$

To reduce the computational effort for the method of lines, the time steps Δt are chosen rather large. In order to nevertheless achieve sufficient accuracy for the approximation of the above integrals, additional substeps of each interval $[t_k, t_{k+1}]$ are introduced. Let $n_s \in \mathbb{N}$ be the number of substeps for each major time step. Then we have

$$\int_H V_d(t, x + \mathcal{P}_d \zeta) \nu_L(t_k, t_{k+1}; d\zeta)$$

$$(137) \qquad \approx \frac{1}{n_s} \sum_{j=1}^{n_s} \int_H V_d(t, x + \mathcal{P}_d \zeta) \, \nu_L\left(t_k + j \frac{t_{k+1} - t_k}{n_s}; d\zeta\right)$$

$$= \frac{\lambda}{n_s} \sum_{j=1}^{n_s} \int_H V_d\left(t, x + \mathcal{P}_d\left[\eta\left(t_k + j \frac{t_{k+1} - t_k}{n_s}, y\right)\right]\right) P^Y(dy)$$

for $0 \le k < n_T$. Similarly, we obtain

$$(138) \qquad \int_H \langle \zeta, p_i \rangle_H \, \nu_L(t_k, t_{k+1}; d\zeta) \approx \lambda \frac{1}{n_s} \sum_{j=1}^{n_s} \int_H \left\langle \eta\left(t_k + j \frac{t_{k+1} - t_k}{n_s}, y\right), p_i \right\rangle_H P^Y(dy).$$

The integrals with respect to P^Y can then be approximated with quadrature formulas. This requires interpolation of the function at quadrature points, which are not necessarily identical to grid points. Since we have assumed that jumps occur only along a relatively small subspace of Ω_R, only a small number of grid points is involved, and the corresponding matrices remain sparse. The idea is illustrated in Figure 1 for a single jump process, where P^Y is defined on $\mathbb{R}$ and the jumps are given by $\eta(t)y \in H$.
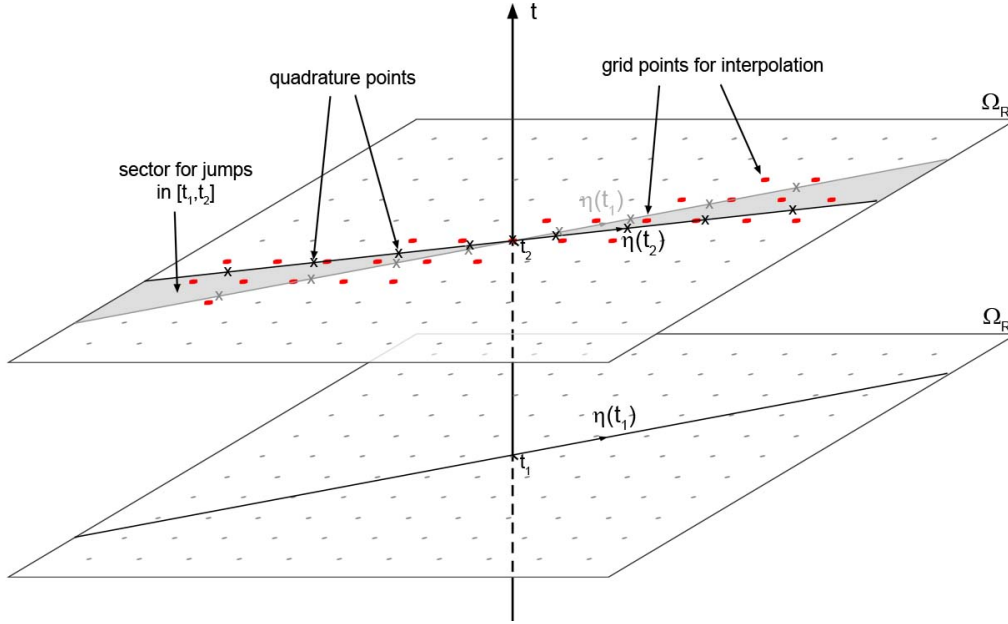
Fig. 1: Grid points (on full grid) involved in the numerical quadrature of jump term integrals in the time interval $[t_1, t_2]$.

SPARSE GRIDS   Besides the nonlocality of the integral term, the exponentially increasing computational complexity with increasing dimension d is the major numerical problem when solving the PIDE. This *curse of dimension* can be broken by using sparse grids. A comprehensive overview of this topic can be found in [8]. Figure 2 shows a sparse grid in two-dimensional space.

In particular, we make use of the combination technique [26]. Thus, we use a standard PIDE solver on a series of full, regular, but anisotropic grids. Instead of using all grids G_α with $\|\alpha\|_\infty \leq M \in \mathbb{N}$, only grids satisfying

$$(139) \qquad M \leq \|\alpha\|_1 \leq M + d - 1$$

for some $M \in \mathbb{N}$ are employed. Denote the approximation of V_d on the grid G_α by $\widetilde{V}_d^\alpha$. An approximation $\widetilde{V}_d^M$, corresponding to a sparse grid solution, can then be obtained by linear combination as follows:

$$(140) \qquad \widetilde{V}_d^M(x) := \sum_{k=1}^{d} (-1)^{k+1} \binom{d-1}{k-1} \sum_{\|\alpha\|_1 = M+d-k} \widetilde{V}_d^\alpha.$$

Since artificial zero boundary conditions are applied, grid points on the boundary $\partial \Omega_R$ do not have to be included.

Because of the anisotropic truncation of the domain introduced in (128), an equal number of refinements in every coordinate would result in finer mesh widths for coordinates which are in fact the least important ones. This mismatch can be remedied
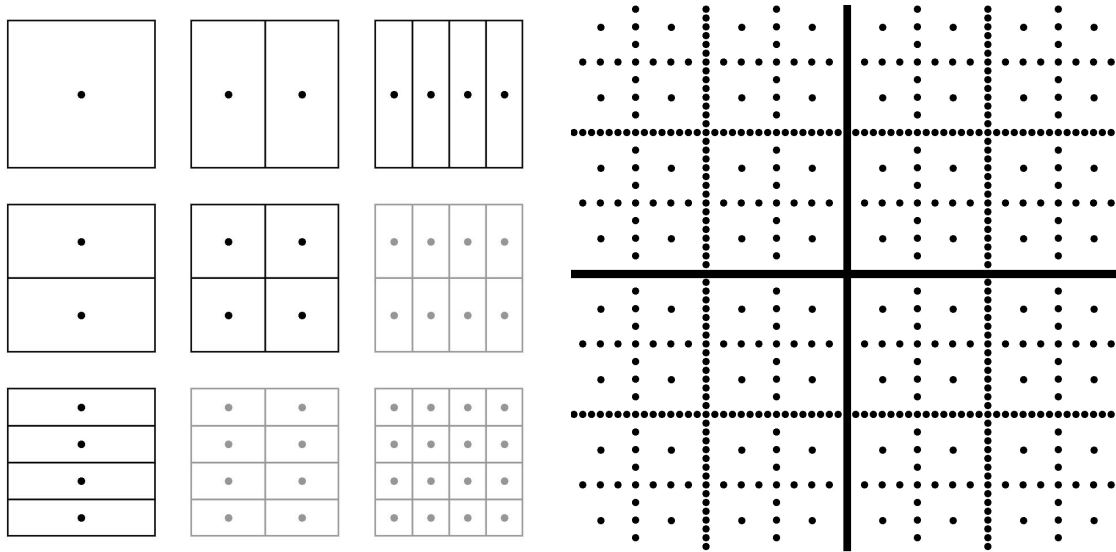
Fig. 2: Construction of sparse tensor product (left) and sparse grid without boundary points (right) in $\mathbb{R}^2$.

by introducing additional constraints on the multiindex of grids used. In order to achieve similar mesh widths, we demand

$$(141) \qquad \alpha_i \le M + \ln\left(\frac{R_i}{R_1}\right) / \ln(2), \quad i = 1, \dots, d,$$

which is equivalent to $2^{\alpha_i} \le 2^M \frac{R_i}{R_1}$. Since this yields a set of grids different from the one obtained by condition (139) alone, the corresponding coefficients in (140) have to be modified. For a detailed presentation of how to choose coefficients in anisotropic sparse grids, see [19].

5 NUMERICAL EXPERIMENTS

In this section, we will examine the performance of the presented option pricing method. To this end, both the dimension reduction and the PIDE solver have been implemented in C++. The program was applied to test problems which are described in detail below.

5.1 TEST PROBLEMS

BASKET OPTION   The first test problem we consider is a basket option on 6 assets. We use a jump-diffusion model similar to [38]. The corresponding Hilbert space is $H = (\mathbb{R}^6, \langle \cdot, \cdot \rangle_2)$, where $\langle \cdot, \cdot \rangle_2$ denotes the Euclidean scalar product. The 6-dimensional

price process $S = (S_1, \ldots, S_6)$ satisfies the dynamics

$$
(142) \quad \frac{dS_i(t)}{S_i(t)} = r\,dt + \sigma_i\,dW_i(t) + \eta_i^0\,d[N_0(t) - \lambda_0 t] + \eta_i^1\,d[N_i(t) - \lambda_i t], \quad i = 1, \ldots, 6,
$$

where $r = 0.05$ is the constant risk-free interest rate. The scalar-valued Brownian motions W_i are supposed to be correlated with correlation matrix

$$
(143) \qquad (\rho_{ij})_{i,j=1}^6 =
\begin{pmatrix}
1 & 0.8 & 0.6 & 0.4 & 0.2 & 0 \\
0.8 & 1 & 0.8 & 0.6 & 0.4 & 0.2 \\
0.6 & 0.8 & 1 & 0.8 & 0.6 & 0.4 \\
0.4 & 0.6 & 0.8 & 1 & 0.8 & 0.6 \\
0.2 & 0.4 & 0.6 & 0.8 & 1 & 0.8 \\
0 & 0.2 & 0.4 & 0.6 & 0.8 & 1
\end{pmatrix}.
$$

We set the volatilities to $\sigma_i = 0.2$, $i = 1, \ldots, 6$. The processes N_0 and N_i are independent Poisson processes with intensities $\lambda_0 = \lambda_i = 1$, describing jumps common for all assets and independent jumps for individual assets, respectively. The discounted value of every asset is thus a martingale under the risk neutral measure. The relative jump heights are set to $\eta_i^0 = -0.2$ and $\eta_i^1 = -0.05$. We can write the value of each asset as the exponential of a Lévy process as follows:

$$
(144) \quad
\begin{aligned}
S_i(t) = S_i(0) \exp &\left\{ \left(r - \frac{1}{2}\sigma_i^2 - \eta_i^0 \lambda_0 - \eta_i^1 \lambda_1 \right) t + \sigma_i W(t) \right. \\
&\left. + \ln\left(1 + \eta_i^0(t)\right) N_0(t) + \ln\left(1 + \eta_i^1(t)\right) N_i(t) \right\}, \quad i = 1, \ldots, 6.
\end{aligned}
$$

We choose the initial value $S_i(0) = \frac{100}{6}$, $i = 1, \ldots, 6$, and strike

$$
(145) \qquad K = 100 \cdot e^{rT} = E\left[\sum_{i=1}^6 S_i(T) \right].
$$

The discounted price of the basket option with maturity $T = 1.0$ at time $t \leq T$ is

$$
(146) \qquad V(t) = e^{-rT} E\left[\left(\sum_{i=1}^6 S_i(T) - K \right)^+ \Big| \mathcal{F}_t \right].
$$

ELECTRICITY SWAPTION   The second problem is an instance of the example described in Section 1, an option on an electricity swap. We use the exponential additive model given in (2). The corresponding Hilbert space is $H = L^2([T_1, T_2], \lambda)$, where $D = [T_1, T_2]$ is the delivery period of the swap and λ is the Lebesgue measure. For the diffusive part of the model, we use two factors, similar to as in [20]. The volatilities are given by

$$
(147) \qquad \sigma_1(s, u) \equiv 0.15, \quad \sigma_2(s, u) = 0.3\, e^{-1.4(u-s)}.
$$

Moreover, we assume additional normally distributed jumps, which yields a *Merton model*. For the jumps, we use a compound Poisson process with intensity $\lambda_1 = 12$ and jump distribution $Y \sim \mathcal{N}(0.1, 0.1)$. The additional factor for dampening the jumps is

$$(148) \qquad \eta_1(s, u) = 0.5 - 0.5 \frac{u - T_1}{T_2 - T_1}.$$

Since an electricity swap requires no payment before the delivery period starts, it is a martingale under the risk neutral measure. Thus, we need the following drift term (compare [11, Chap. 8.4.1]):

$$(149) \qquad \gamma(s, u) = -\frac{1}{2} \sum_{i=1}^{2} \sigma_i^2(s, u) - \lambda_1 \int_{\mathbb{R}} (e^{\eta_1(s,u)y} - 1 - \eta_1(s, u)y) P^Y(dy).$$

The risk-free interest rate is assumed to be constant at $r = 0.05$. We consider a monthly swap maturing in one year, i.e., $T = T_1 = 1$, $T_2 = T_1 + 30/365$. The initial forward curve at time $t = 0$ is $S_0(u) \equiv 50$, $u \in [T_1, T_2]$, and the strike is $K = 50$. It remains to specify the discretization of the delivery period $[T_1, T_2]$. As we have a continuous forward curve model, we may use an arbitrary number of discretization points. On energy markets, monthly swaps on electricity are usually based on daily prices. Thus we will use exactly $n = 30$ components. We will futher assume that there are 8 delivery hours per day. Setting $u_i = T_1 + (i - 1)\frac{T_2 - T_1}{30}$ for $i = 1, \ldots, 30$, we obtain

$$(150) \qquad V(t) = e^{-rT} \cdot 8 \cdot 30 \cdot E\left[\left(\sum_{i=1}^{30} w(u_i; T_1, T_2) S(T, u_i) - K \right)^+ \Bigg| \mathcal{F}_t \right]$$

for the discounted price of the swaption, without making any discretization error.

5.2 Results

Since no analytical solutions are available, a large number of Monte Carlo (MC) simulations were performed for each test problem to obtain a precise solution. All errors were computed using these MC reference values. In order to make use of the easily parallelized methods (MC simulation as well as the sparse grid combination technique), the experiments were run on a Linux workstation with 6 Opteron processors at 2.7 GHz.

RESULTS FOR BASKET OPTION   We first examine the number of POD components needed to obtain a sufficiently good approximation. The eigenvalues μ_i of the covariance operator $\mathcal{C}_X$ defined in (26) are given in Table 1. They exhibit a strong decay, which is not uncommon (also for real market data). The corresponding explained variability, defined by $\sum_{j=1}^{i} \mu_j / \sum_{j=1}^{6} \mu_j$, is also shown in the table.

The computed POD components are displayed in Figure 3. They resemble those known from fixed income markets. In particular, the first three basis vectors represent the typical shift, tilt, and bend. Further components feature higher frequencies.

i	μ_i	μ_i/μ_1	Expl. var.
1	0.4490	1.0000	0.8096
2	0.0623	0.1389	0.9221
3	0.0177	0.0395	0.9541
4	0.0106	0.0237	0.9732
5	0.0079	0.0177	0.9875
6	0.0069	0.0154	1.0000

Table 1: Basket option – Eigenvalues and explained variability.

We now examine the errors of the PIDE method. The "exact" reference solution used here is the result from 10^7 MC simulations with a standard deviation of 0.0063 (estimated from 100 MC series). It turns out that due to the stable discontinuous Galerkin method, the mesh size for the time variable has little influence on the PIDE results. Thus we do all computations with fixed, equidistant time steps of size $\Delta t = \frac{1}{10}$. The spatial grid, on the other hand, has considerable impact on the accuracy. Figure 4 displays the (signed) relative error for different dimensions d of the projected problem. The number $N := 2^M$ is the maximal number of discretization points in one coordinate and is always taken to be a power of 2. Two effects can be observed here. For each fixed value of N, the method converges to a certain limit when the dimension of the problem is increased. These limits, in turn, converge to the exact solution with increasingly fine meshes. Thus, we might accidentally get a very precise result for a low-dimensional computation on a coarse grid if the two errors (from dimension and mesh) happen to cancel each other. In practice, both the dimension and the number of grid points have to be chosen large enough in order to guarantee a precise result. In our case, $d = 4$ and $N = 2^8$ are sufficient to obtain relative errors below 1%. Approximating the basket value with a log-normal distribution (Lévy approximation), on the other hand, yields a relative error of 3.1%.

Finally, we have a look at the computational time needed for the PIDE method. To this end, we fix $N = 2^9$ and plot the error for various dimensions. Figure 5 displays the results; both y-axes (for time and error) have a logarithmic scale. The error decreases approximately exponentially with increasing dimension. The computational time, on the other hand, increases exponentially. Note that the solution of the problem without dimension reduction takes approximately 2 minutes, even though we use the sparse grid combination technique. However, the increase of the projected problem dimension d by one increases the computational effort by only a factor of 4 to 5, despite the fact that we use up to $2^9 = 512$ grid points in each coordinate. A reasonably precise solution (in practice within the bounds of the model error) can be computed within a few seconds. While this is slightly faster than the MC method, it is not an extreme gain. The computational effort might be greatly reduced by using more sophisticated grids, featuring more grid points around the origin and fewer close to the boundary of the domain, which we will not consider here.
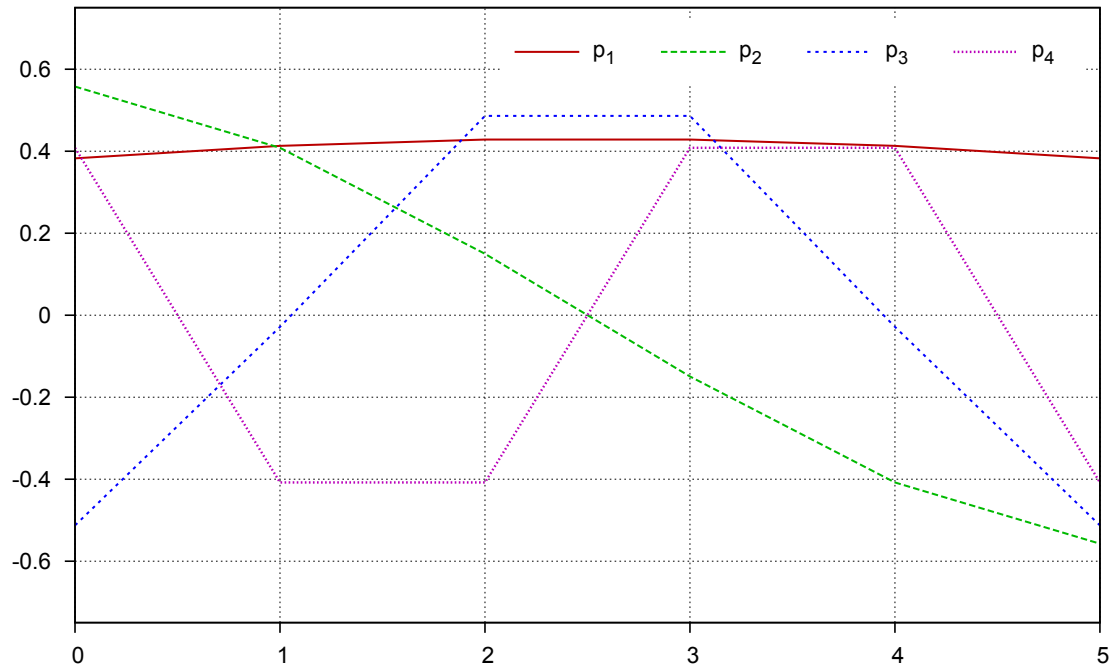
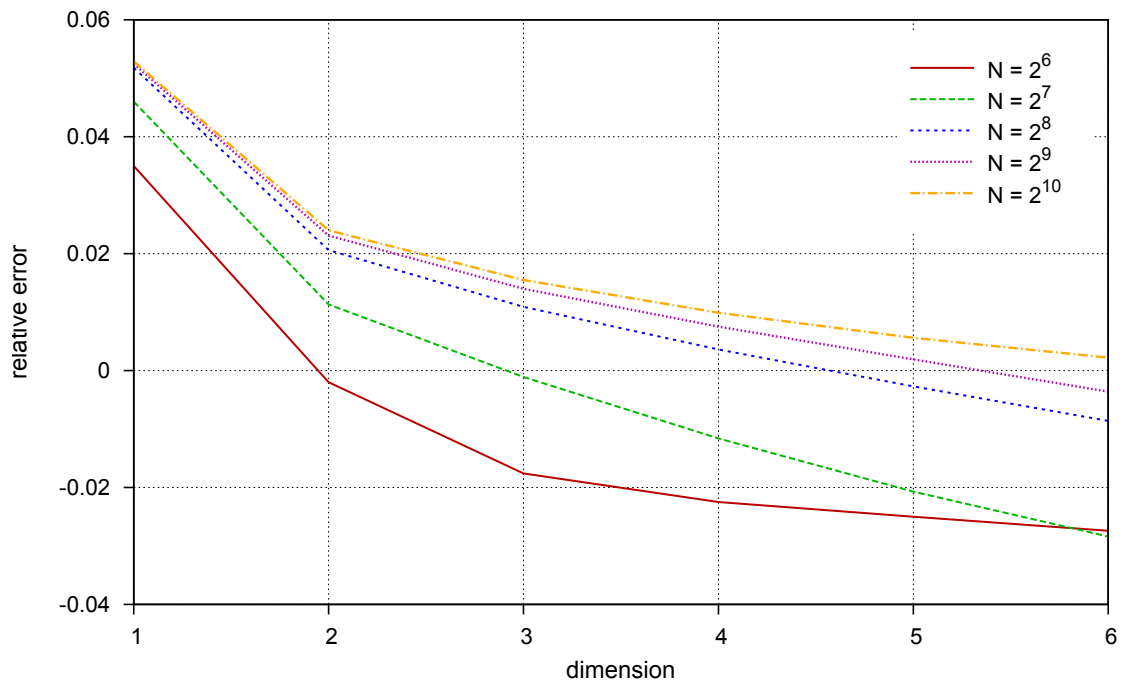Fig. 3: Basket – The first four POD basis components.

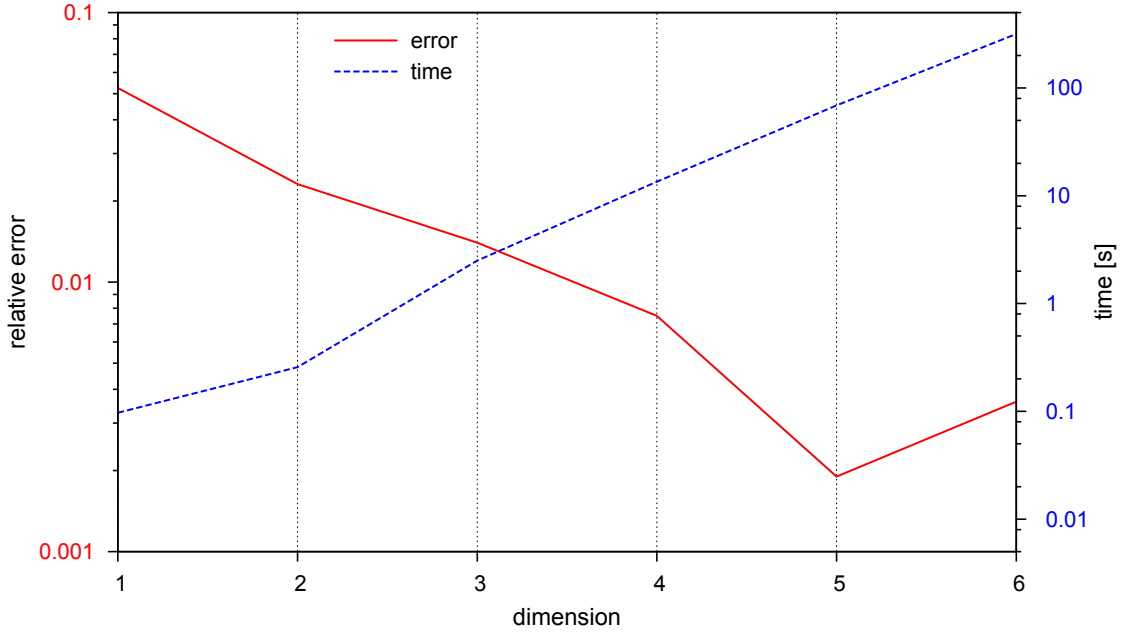Fig. 4: Basket – Relative error of PIDE method with different meshes.

Fig. 5: Basket – Relative error and computational time of PIDE method with at most $N = 2^9$ grid points per coordinate.

RESULTS FOR ELECTRICITY SWAPTION   In the test case for electricity swaptions, two POD basis components are already sufficient to explain almost 100% of the volatility, and the sum of the remaining eigenvalues satisfies $\sum_{i=3}^{\infty} \mu_i \approx 0$. This is of course due to the strong correlation between the price changes for different maturities $u \in [T_1, T_2]$, which makes the dimension reduction technique a particularly well suited method for this type of derivative. The forward curve defined on the delivery interval does not change its shape arbitrarily. The two POD components accurately describe the possible shape changes in our (rather simple) test setting. We are thus able to compute accurate prices by solving a two-dimensional PIDE.

Figure 6 displays the relative error and time of the PIDE method for different mesh widths. As was to be expected, the computational effort increases exponentially in the number of grid refinements (and thus linear in the total number of grid points). The relative error decreases exponentially. However, in contrast to the basket option, a very accurate solution can be computed for the electricity swaption within a fraction of a second. This is made possible by the very efficient dimension reduction from $n = 30$ to $d = 2$.

A comparison of the PIDE method, MC simulation, and the log-normal approximation is displayed in Figure 7. The log-normal approximation yields a relative error of 3.5%, which is by far the largest of all the methods. MC simulation yields very good results for 10^6 runs and above. However, with $n = 30$ it takes considerably longer than in the case of the 6-dimensional basket. The standard deviation after 10^6 simulations
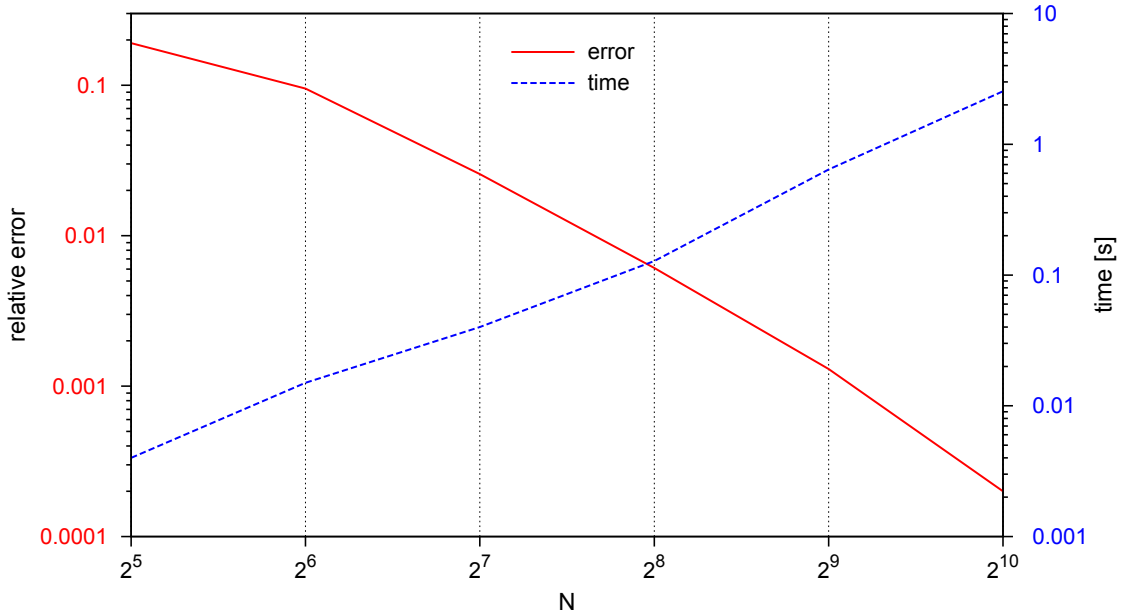
35

Fig. 6: Swaption – Relative error and computational time of PIDE method for dimension $d = 2$.

is 0.4667 (estimated from 100 MC series). The PIDE solver is indeed the fastest and most accurate method for this second test problem.

6 CONCLUSION

In this article, numerical pricing of European options in Hilbert space-valued jump-diffusion models is discussed. We have presented a feasible approach employing PIDEs and a dimension reduction method based on Karhunen–Loève expansion. The PIDE can be projected to an approximating low-dimensional equation by solving an eigenvalue problem. Existence, uniqueness, and convergence of the approximating solutions have been shown based on the variational formulation of the PIDE. The numerical solution of the problem is based on a sparse grid combination method for spatial discretization and a discontinuous Galerkin time stepping method.

   Numerical experiments have been performed for two applications: an electricity swaption and a basket option. The swaption can be priced very efficiently with the presented algorithm, which gives more accurate results in less time than MC simulation. For the basket option, the PIDE method yields comparable performance to MC simulation, depending on the correlation of the assets. Using PIDEs, however, should have considerable advantages when pricing path-dependent options or options driven by infinite activity Lévy models, both topics for future research. In addition, future work will be concerned with adaptive and graded grids, further improving the per-
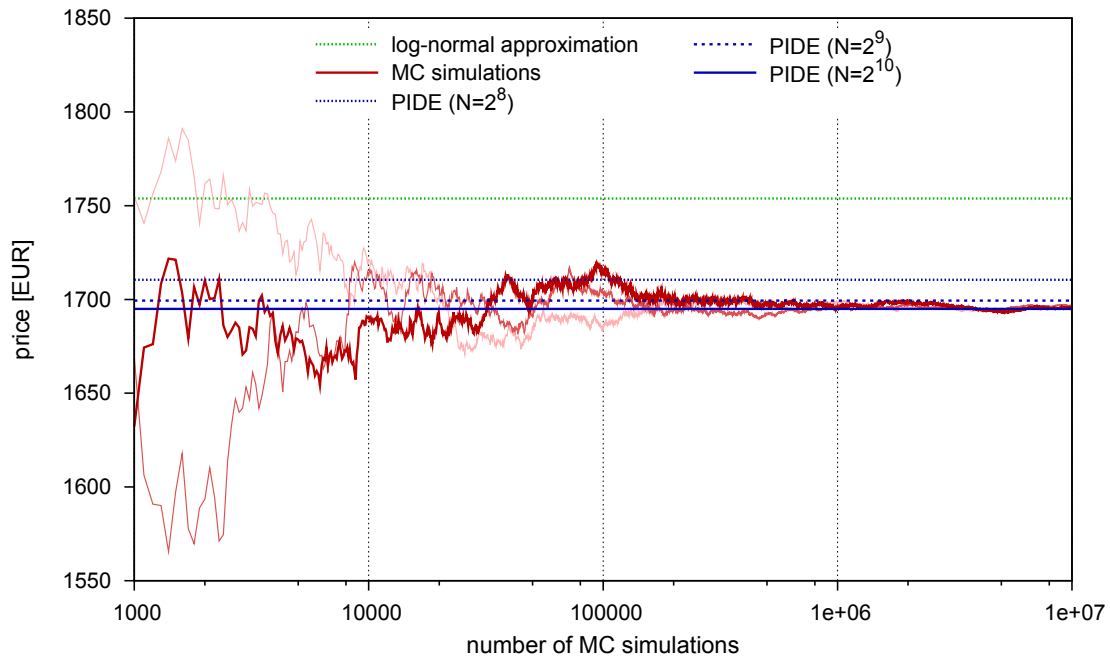
Fig. 7: Swaption – PIDE solutions ($d = 2$), MC simulations and log-normal approximation.

formance of the method.

Acknowledgments

References

[1] N. Audet, P. Heiskanen, J. Keppo, and I. Vehviläinen. Modeling electricity forward curve dynamics in the Nordic market. In D. W. Bunn, editor, *Modelling Prices in Competitive Electricity Markets*, chapter 12. Wiley Series in Financial Economics, Chichester, 2004.

[2] A. Bensoussan and J.-L. Lions. *Impulse Control and Quasi-Variational Inequalities*. Gauthier-Villars, Paris, 1984.

[3] F. E. Benth, Á. Cartea, and R. Kiesel. Pricing forward contracts in power markets by the certainty equivalence principle: explaining the sign of the market risk premium. *Journal of Banking and Finance*, 32(10):2006–2021, 2008.

[4] F. E. Benth and S. Koekebakker. Stochastic modeling of financial electricity contracts. *Energy Economics*, 30(3):1116–1157, 2008.

[5] F. E. Benth, S. Koekebakker, and F. Ollmar. Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variation. *Journal of Derivatives*, 52(15):62–66, 2007.

[6] F. E. Benth and T. Meyer-Brandis. The information premium for non-storable commodities. *Journal of Energy Markets*, 2(3):111–140, 2009.

[7] F. E. Benth, J. Šaltytė Benth, and S. Koekebakker. *Stochastic Modeling of Electricity and Related Markets*. World Scientific, Singapore, 2008.

[8] H.-J. Bungartz and M. Griebel. Sparse grids. *Acta Numer.*, 13:147–269, 2004.

[9] R. Carmona and V. Durrleman. Generalizing the Black-Scholes formula to multivariate contingent claims. *J. Comput. Finance*, 9(2):43–67, 2006.

[10] R. Carmona and M. Tehranchi. *Interest Rate Models: An Infinite-Dimensional Stochastic Analysis Perspective*. Springer, Berlin, 2006.

[11] R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall, Boca Raton, 2004.

[12] R. Cont and E. Voltchkova. A finite difference scheme for option pricing in jump diffusion and exponential Lévy models. *SIAM J. Numer. Anal.*, 43(4):1596–1626, 2005.

[13] R. Cont and E. Voltchkova. Integro-differential equations for option prices in exponential Lévy models. *Finance and Stochastics*, (9):299–325, 2005.

[14] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, Cambridge, 1992.

[15] G. Da Prato and J. Zabczyk. *Second Order Partial Differential Equations in Hilbert Spaces*. Cambridge University Press, Cambridge, 2004.

[16] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, Graduate Studies in Mathematics, Volume 19, Providence, 1998.

[17] E. Hausenblas. Existence, uniqueness and regularity of parabolic SPDEs driven by Poisson random measure. *Electron. J. Probab.*, 10:1496–1546, 2005.

[18] E. Hausenblas. A note on the Itô formula of stochastic integrals in Banach spaces. *Random Oper. Stochastic Equations*, 14(1):45–58, 2006.

[19] M. Hegland, J. Garcke, and V. Challis. The combination technique and some generalisations. *Linear Algebra Appl.*, 420:249–275, 2007.

[20] R. Kiesel, G. Schindlmayr, and R. Börger. A two-factor model for the electricity forward market. *Quant. Finance*, 9(3):279–287, 2009.

[21] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numer. Math.*, 90(1):117–148, 2001.

[22] H. Kunita. Stochastic integrals based on martingales taking values in Hilbert space. *Nagoya Math. J.*, 38:41–52, 1970.

[23] M. Loève. *Probability Theory II*. Springer, New York, fourth edition, 1978.

[24] A.-M. Matache, T. von Petersdorff, and C. Schwab. Fast deterministic pricing of options on Lévy-driven assets. *ESAIM: Mathematical Modelling and Numerical Analysis*, 38(1):37–71, 2004.

[25] S. Peszat and J. Zabczyk. *Stochastic Partial Differential Equations with Lévy Noise: An Evolution Equation Approach*. Cambridge University Press, Cambridge, 2007.

[26] C. Pflaum and A. Zhou. Error analysis of the combination technique. *Numer. Math.*, 84:327–350, 1999.

[27] P. E. Protter. *Stochastic Integration and Differential Equations*. Springer, Berlin, second edition, 2005.

[28] N. Reich. *Wavelet Compression of Anisotropic Integrodifferential Operators on Sparse Tensor Product Spaces*. PhD thesis, ETH Zürich, 2008.

[29] B. Rüdiger and G. Ziglio. Itô formula for stochastic integrals with respect to compensated Poisson random measures on separable Banach spaces. *Stochastics*, (6):377–410, 2006.

[30] E. W. Sachs and A. K. Strauss. Efficient solution of a partial integro-differential equation in finance. *Appl. Numer. Math.*, 58(11):1687–1703, 2008.

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

[32] C. Schwab and R. A. Todor. Karhunen-Loève approximation of random fields by generalized fast multipole methods. *J. Comput. Phys.*, 217:100–122, 2006.

[33] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*. Springer, Berlin, second edition, 2006.

[34] R. A. Todor. Robust eigenvalue computations for smoothing operators. *SIAM J. Numer. Anal.*, 44(2):865–878, 2006.

[35] D. Werner. *Funktionalanalysis*. Springer, Berlin, 5th edition, 2005.

[36] C. Winter. *Wavelet Galerkin Schemes for Option Pricing in Multidimensional Lévy Models*. PhD thesis, ETH Zürich, 2009.

[37] J. Wloka. *Partial Differential Equations*. Cambridge University Press, Cambridge, 1987.

[38] G. Xu and H. Zheng. Approximate basket options valuation for a jump-diffusion model. *Insurance Math. Econom.*, 45(2):188–194, 2009.