Eidgenössische
Technische Hochschule
Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Wavelet finite element method for option pricing in highdimensional diffusion market models

N. Hilber, S. Kehtari, C. Schwab and C. Winter

Research Report No. 2010-01
January 2010

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

Wavelet finite element method for option pricing in highdimensional diffusion market models

N. Hilber*, S. Kehtari*, C. Schwab*, C. Winter*

January 14, 2010

Abstract

We consider the numerical solution of high-dimensional partial differential equations arising in option pricing problems in computational finance. To reduce the complexity in the number of degrees of freedom, sparse tensor product spaces are applied for Galerkin discretization in log-price space. Using this technique we are able to price multi-asset options with up to eight underlying assets for the Black-Scholes framework and stochastic volatility models. Dimensionality reduction by principal component analysis and asymptotic expansion is investigated in order to price options on indices by considering the whole vector process of all of their constituents.

Keywords: Multi-asset options, stochastic volatility, sparse tensor finite elements, wavelets

1 Introduction

Consider a basket of $d \geq 1$ risky assets whose log returns $X_t = (X_t^1, \dots, X_t^d)^\top \in \mathbb{R}^d$ at time $t > 0$ are modeled by a diffusion process $X = \{X_t : t \geq 0\}$ with state space $\mathbb{R}^d$. Arbitrage free prices u of European contingent claims with payoffs $g(\cdot)$ and maturity T are given by the conditional expectation

$$u(t,x) = \mathbb{E}\big[e^{-\gamma(T-t)} g(e^{X_T}) \mid X_t = x\big], \tag{1.1}$$

with γ denoting throughout the risk-free interest rate. Here, the expectation is taken with respect to an a-priori chosen martingale measure equivalent to the historical measure. Sufficiently smooth value functions u in (1.1) can be obtained as solutions of a partial differential equation (PDE)

$$\partial_t u + \mathcal{A}u - \gamma u = 0\,, \tag{1.2}$$

*Seminar for Applied Mathematics, ETH Zurich, 8092 Zurich, Switzerland

where $\mathcal{A}$ is the infinitesimal generator of the process X. To allow low smoothness assumption on the payoff g, we opt for variational solutions which are the basis for variational discretization methods such as finite element discretizations. To convert (1.2) into variational form, we formally integrate against a test function v and obtain (assuming $\gamma = 0$ for convenience)

$$(\partial_t u, v) + a(u, v) = 0, \tag{1.3}$$

where the bilinear form is given by $a(u, v) = \langle \mathcal{A}u, v \rangle$. Note that $\mathcal{A} = \mathcal{A}(x)$ is admissible. For solving problem (1.2) or (1.3) numerically, straightforward application of standard schemes fails due to the so-called 'curse of dimension': the number of degrees of freedom on a mesh of width h in dimension d grows like $\mathcal{O}(h^{-d})$ as $h \to 0$. To avoid this problem, several authors (see [5, 8] and the references therein) use finite differences on sparse grids. In this paper we follow [9] and consider sparse tensor product spaces for the discretization to reduce the complexity in the number of degrees of freedom from $\mathcal{O}(h^{-d})$ to $\mathcal{O}(h^{-1} |\log h|^{d-1})$. In particular, choosing wavelet bases we additionally obtain an efficient preconditioner for the resulting linear equations.

The outline of the paper is as follows. We first introduce diffusion market models where we focus on the multivariate Black-Scholes and stochastic volatility models. In Section 3 we give the abstract variational formulation of the option pricing problem. The discretization is discussed in Section 4 and sensitivities in Section 5. We give numerical examples in Section 6 and formalize an extension to higher dimensions in Section 7.

2 Diffusion market models

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space that satisfies the usual hypotheses, i.e., $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets of $\mathcal{F}$ and $\mathcal{F}_{t^+} := \cap_{h>0}\mathcal{F}_{t+h} = \mathcal{F}_t$, $t \geq 0$. Let $W = \{W_t : t \geq 0\}$ be a $\mathbb{R}^d$-valued standard Brownian motion. We assume that the filtration is generated by W and furthermore that W is independent of $\mathcal{F}_0$.

We consider a process Z to model the dynamics of the underlying stock prices and of the *background* volatility drivers in case of stochastic volatility models. Let g be the payoff in real price, $T > 0$ the maturity, $\gamma \geq 0$ the (time-constant) interest rate and $\mathbb{Q}$ an equivalent martingale measure (EMM) to $\mathbb{P}$, i.e., $\mathbb{Q} \sim \mathbb{P}$ such that the discounted process is a $\mathbb{Q}$-martingale. If Z is Markovian, the fair price of a European style contingent claim with underlying Z is given by

$$u(t, z) = \mathbb{E}^{\mathbb{Q}}\big[e^{-\gamma(T-t)} g(Z_T) \mid Z_t = z\big]. \tag{2.1}$$

We model the market Z by the stochastic differential equation (SDE)

$$\mathrm{d}Z_t = \mu(Z_t)\,\mathrm{d}t + \Sigma(Z_t)\,\mathrm{d}W_t, \quad Z_0 = z. \tag{2.2}$$

Herewith, for $G \subseteq \mathbb{R}^d$, the coefficients $\mu : G \to \mathbb{R}^d$, $\Sigma : G \to \mathbb{R}^{d\times d}$ are assumed to be globally Lipschitz continuous. Thus, for a given random vector z which is $\mathcal{F}_0$-measurable, the SDE (2.2) admits a unique $(\mathcal{F}_t)$-adapted solution $Z = (Z_t)_{t\geq 0}$ such that $Z_0 = z$ a.s.. We consider next two kinds of market dynamics, namely the Black-Scholes model and stochastic volatility models.

2.1 Aggregated Black-Scholes models

2.1.1 Full-rank Black-Scholes

Consider d assets $S = (S^1, \dots, S^d)$ with spot price dynamics $Z^i = S^i$ given by

$$\mathrm{d}S_t^i = \mu_i S_t^i \mathrm{d}t + \sum_{j=1}^{d} \mathbf{\Sigma}_{ij} S_t^i \mathrm{d}W_t^j, \quad i = 1, \dots, d, \tag{2.3}$$

where $W = \{W_t : t \geq 0\}$ is a standard Brownian motion in $\mathbb{R}^d$ and

$$\mu \;:=\; \big(\mu_i\big)_{1\leq i\leq d} \in \mathbb{R}^d, \tag{2.4}$$

$$\Sigma \;:=\; \big(\mathbf{\Sigma}_{ij}\big)_{1\leq i,j\leq d} \in \mathbb{R}^{d\times d}, \tag{2.5}$$

are the constant drift vector and volatility matrix, respectively, with assumption that rank $\mathbf{\Sigma} = d$. The state space domain is given by $G = \mathbb{R}^d$. Under the unique EMM, the log-price dynamics $X^i := \log S^i$ are given by

$$\mathrm{d}X_t^i = \eta_i \mathrm{d}t + \sum_{j=1}^{d} \mathbf{\Sigma}_{ij} \mathrm{d}W_t^j, \quad i = 1, \dots, d, \tag{2.6}$$

where $\eta_i := \big(\gamma - 1/2\mathcal{Q}_{ii}\big)$, $i = 1, \dots, d$ and $\mathcal{Q} := \mathbf{\Sigma}\mathbf{\Sigma}^\top \in \mathbb{R}^{d\times d}_{\text{sym.}}$ denotes the volatility covariance matrix. Since $\mathcal{Q}$ is symmetric positive definite, there exists an orthogonal matrix $\mathbf{U} \in \mathbb{R}^{d\times d}$ such that $\mathbf{U}\mathcal{Q}\mathbf{U}^\top = \mathbf{D}$ with diagonal matrix $\mathbf{D} := \operatorname{diag}(s_1^2, \dots, s_d^2)$, $s_1 \geq \ldots \geq s_d > 0$. Without loss of generality, we rescale time in (2.3) such that $t \to t^* = s_1^2 t$ yielding $\mathbf{D}^* := \operatorname{diag}(s_1^{*2}, \dots, s_d^{*2})$ with normalized $s_1^* = 1$, $s_i^* = s_i/s_1$, $i = 2, \dots, d$. In the remainder, we drop the * and define a process $Y := \{\mathbf{U}X_t : t \geq 0\}$ with dynamics

$$\mathrm{d}Y_t^i = \lambda_i \mathrm{d}t + s_i \mathrm{d}W_t^i, \quad i = 1, \dots, d, \tag{2.7}$$

where $\lambda := \mathbf{U}\eta$. The components $Y^1, \dots, Y^d$ now satisfy the system of d decoupled SDEs (2.7).

2.1.2 Low-rank Black-Scholes

Let $1 \leq r < d$ be a parameter and define $\widehat{\mathbf{D}} := \operatorname{diag}(\hat{s}_1^2, \dots, \hat{s}_d^2) \in \mathbb{R}^{d\times d}$ with

$$\hat{s}_i = \begin{cases} s_i & 1 \leq i \leq r, \\ 0 & r+1 \leq i \leq d, \end{cases} \tag{2.8}$$

and $\widehat{\mathbf{\Sigma}} := \mathbf{U}^\top \widehat{\mathbf{D}}^{\frac{1}{2}}$ with $\mathbf{U}$ and s_i, $i = 1, \ldots, d$, as in Section 2.1.1. Consider the log-price process $\widehat{X} := \{\widehat{X}_t : t \geq 0\}$ with dynamics

$$\mathrm{d}\widehat{X}_t^i = \widehat{\eta}_i \,\mathrm{d}t + \sum_{j=1}^{r} \widehat{\mathbf{\Sigma}}_{ij} \,\mathrm{d}W_t^j, \quad i = 1, \ldots, d, \tag{2.9}$$

where $\widehat{\eta}_i := \big(\gamma - 1/2\widehat{\mathcal{Q}}_{ii}\big)$, $i = 1, \ldots, d$ and $\widehat{\mathcal{Q}} = \widehat{\mathbf{\Sigma}}\widehat{\mathbf{\Sigma}}^\top \in \mathbb{R}^{d\times d}_{\mathrm{sym.}}$ and W_t^j, $j = 1, \ldots, r$, as in (2.3). We designate the process $\widehat{X}$ as *rank-reduced* from d to r since, under the change of basis induced by $\mathbf{U}$, the process $\widehat{Y} := \{\mathbf{U}\widehat{X}_t : t \geq 0\}$ has dynamics

$$\begin{aligned} \mathrm{d}\widehat{Y}_t^i &= \widehat{\lambda}_i \,\mathrm{d}t + \hat{s}_i \,\mathrm{d}W_t^i \\ &= \widehat{\lambda}_i \,\mathrm{d}t + \mathbb{1}_{\{1 \leq i \leq r\}} s_i \,\mathrm{d}W_t^i, \quad i = 1, \ldots, d, \end{aligned} \tag{2.10}$$

where $\widehat{\lambda} := \mathbf{U}\widehat{\eta}$. The components $\widehat{Y}^{r+1}, \ldots, \widehat{Y}^d$ are therefore deterministic. Under such market dynamics, the price of a European contingent claim (2.1) becomes

$$\begin{aligned} u(t, \hat{x}) &= v(t, \hat{y}) \\ &= \widehat{v}(t, \hat{y}_1, \ldots, \hat{y}_r; \hat{y}_{r+1}, \ldots, \hat{y}_d) \\ &= \mathbb{E}\big[e^{-\gamma(T-t)} \widehat{f}(e^{\widehat{Y}_T^1}, \ldots, e^{\widehat{Y}_T^r}; e^{\hat{y}_{r+1}}, \ldots, e^{\hat{y}_d}) \,|\, (\widehat{Y}_t^1, \ldots, \widehat{Y}_t^r) = (\hat{y}_1, \ldots, \hat{y}_r)\big], \end{aligned}$$

with $\hat{y} := \mathbf{U}\hat{x}$ and

$$\begin{aligned} \widehat{f}(e^{\hat{y}}) &= \widehat{f}(e^{\hat{y}_1}, \ldots, e^{\hat{y}_r}; e^{\hat{y}_{r+1}}, \ldots, e^{\hat{y}_d}) \\ &= f(e^{\hat{y}_1}, \ldots, e^{\hat{y}_r}, e^{\hat{y}_{r+1} + \widehat{\lambda}_{r+1}(T-t)}, \ldots, e^{\hat{y}_d + \widehat{\lambda}_d(T-t)}), \end{aligned} \tag{2.11}$$

herewith $f(e^y) := g(e^x)$. The price $u(t, \hat{x}) = \widehat{v}(t, \hat{y}_1, \ldots, \hat{y}_r; \hat{y}_{r+1}, \ldots, \hat{y}_d)$ becomes a solution of an r-dimensional PDE with initial condition depending on parameters $\hat{y}_{r+1}, \ldots, \hat{y}_d$.

2.1.3 ϵ-Aggregation

Suppose a time rescaled d-dimensional Black-Scholes market model with log-price process X as in (2.3) has a volatility covariance matrix $\mathcal{Q}$ of full rank. Given $0 \leq \epsilon \ll 1$, assume that a principal component analysis of $\mathcal{Q}$, i.e., $\mathbf{U}\mathcal{Q}\mathbf{U}^\top = \mathbf{D} = \mathrm{diag}(s_1^2, \ldots, s_d^2)$ with $s_1 = 1 \geq \ldots \geq s_d > 0$ yields

$$s_i^2 \leq \epsilon, \quad i = r+1, \ldots, d,$$

for some $r = r(\epsilon)$. This suggests the d-dimensional dynamics to be mainly driven by $r < d$ ϵ-*aggregated price processes* (see, e.g., [8]). We denote $\widehat{X}^\epsilon$ the ϵ-aggregated rank-r process of X with $\hat{s}_{r+1} = \cdots = \hat{s}_d = 0$ as defined in Section 2.1.2 and define the ϵ-residual market process, i.e., the aggregation remainder, $R^\epsilon := X - \widehat{X}^\epsilon$. From (2.6) and (2.9), we have that

$$\mathrm{d}R_t^{\epsilon,i} = \mathrm{d}(X - \widehat{X}^\epsilon)_t^i = (\eta_i - \widehat{\eta}_i)\,\mathrm{d}t + \sum_{j=1}^{d} (\mathbf{\Sigma}_{ij} - \widehat{\mathbf{\Sigma}}_{ij})\,\mathrm{d}W_t^j \tag{2.12}$$

Under the change of basis induced by $\mathbf{U}$, the process $T^\epsilon := \{\mathbf{U}R^\epsilon_t : t \geq 0\}$, i.e., the fluctuation of X about the ϵ-aggregate $\widehat{X}^\epsilon$, has therefore dynamics

$$\begin{aligned} \mathrm{d}T^{\epsilon,i}_t &= (\mathbf{U}(\eta - \widehat{\eta}))_i \,\mathrm{d}t + \sum_{j=1}^{d} (\mathbf{D} - \widehat{\mathbf{D}})_{ij} \,\mathrm{d}W^j_t \\ &= \begin{cases} (\lambda_i - \widehat{\lambda}_i)\,\mathrm{d}t & 1 \leq i \leq r(\epsilon), \\ (\lambda_i - \widehat{\lambda}_i)\,\mathrm{d}t + s_i \,\mathrm{d}W^t_i & r(\epsilon) + 1 \leq i \leq d. \end{cases} \end{aligned} \tag{2.13}$$

Lemma 2.1. *Given $r = r(\epsilon)$, there holds*

$$|\lambda_i - \widehat{\lambda}_i| \leq \frac{d}{2} \sum_{j=r+1}^{d} s_j^2, \quad i = 1, \ldots, d.$$

Proof. From the definitions of η and λ, we have that

$$\begin{aligned} |\lambda_i - \widehat{\lambda}_i| &= \left| \sum_{k=1}^{d} U_{ik} (\eta_k - \widehat{\eta}_k) \right| \leq \sum_{k=1}^{d} |\eta_k - \widehat{\eta}_k| = \frac{1}{2} \sum_{k=1}^{d} |\mathcal{Q}_{kk} - \widehat{\mathcal{Q}}_{kk}| \\ &= \frac{1}{2} \sum_{k=1}^{d} \left| \sum_{j=1}^{d} U_{jk}^2 (D_{jj} - \widehat{D}_{jj}) \right| \leq \frac{1}{2} \sum_{k=1}^{d} \sum_{j=r+1}^{d} s_j^2 \\ &= \frac{d}{2} \sum_{j=r+1}^{d} s_j^2, \quad i = 1, \ldots, d. \end{aligned}$$

□

Remark 2.2. *From* (2.13) *and Lemma* (2.1)*, we conclude that the fluctuation components $T^{\epsilon,i}$, $i = 1, \ldots, r(\epsilon)$, are pure drifts of order ϵ. Furthermore note that, upon setting*

$$d \to \widehat{d} = d - r(\epsilon), \quad t \to \widehat{t} = s^2_{r(\epsilon)+1} t,$$

the fluctuation components $T^{\epsilon,i}$, $i = r(\epsilon) + 1, \ldots, d$, again define a $\widehat{d}$-dimensional full-rank market of type (2.3)–(2.5) *with timescale $\widehat{t}$, allowing, in principle, for recursive ϵ-rank aggregation. This will be elaborated elsewhere.*

2.1.4 ϵ-Aggregation error bound

We again consider the d-dimensional Black-Scholes market model with log-price process X and its ϵ-aggregate rank r process $\widehat{X}^\epsilon$ of the previous section, and we estimate the error of approximating $u(t,x)$ by $\widehat{u}(t, \hat{x}^\epsilon)$ in (2.1).

Theorem 2.3. *Assume that the payoff g is Lipschitz. Then, there exists a constant $C(x)$ independent of ϵ such that*

$$|u(t,x) - \widehat{u}(t,\hat{x}^\epsilon)| \le C(x) \sum_{i=r(\epsilon)+1}^{d} s_i^2,$$

Proof. We introduce the artificial process $\widetilde{Y}^\epsilon$ with dynamics

$$\mathrm{d}\widetilde{Y}_t^{\epsilon,i} = \lambda_i \mathrm{d}t + \mathbb{1}_{\{1\le i\le r(\epsilon)\}} s_i \mathrm{d}W_t^i, \quad i = 1,\dots,d.$$

Under the change of basis induced by $\mathbf{U}$, we have

$$\begin{aligned}|u(t,x) - \widehat{u}(t,\hat{x}^\epsilon)| &= |v(t,y) - \widehat{v}(t,\hat{y}^\epsilon)| \\ &\le |v(t,y) - \widehat{v}(t,\tilde{y})| + |\widehat{v}(t,\tilde{y}) - \widehat{v}(t,\hat{y}^\epsilon)|. \end{aligned} \tag{2.14}$$

The two terms in (2.14) are estimated separately. Since g is globally Lipschitz, we have for the first term, where $f(e^y) = g(e^x)$ and constants may change between lines,

$$\begin{aligned}|v(t,y) - \widehat{v}(t,\tilde{y})| &= \left|\mathbb{E}\left[f(e^{y+Y_{T-t}}) - f(e^{y+\widetilde{Y}^\epsilon_{T-t}})\right]\right| \le C\sum_{i=1}^{d}\mathbb{E}\left[\left|e^{y_i+Y^i_{T-t}} - e^{y_i+\widetilde{Y}^{\epsilon,i}_{T-t}}\right|\right] \\ &= C\sum_{i=r(\epsilon)+1}^{d} e^{y_i+\lambda_i(T-t)}\,\mathbb{E}\left[\left|e^{s_iW^i_{T-t}} - 1\right|\right] \\ &= C\sum_{i=r(\epsilon)+1}^{d} e^{y_i+\lambda_i(T-t)}\int_{\mathbb{R}}|e^z - 1|\,e^{-z^2/(2s_i^2)}\,\mathrm{d}z \\ &\le C\sum_{i=r(\epsilon)+1}^{d} e^{y_i}s_i^2 \le C(y)\sum_{i=r(\epsilon)+1}^{d} s_i^2.\end{aligned}$$

Similarly, using Lemma (2.1), we have for the second term

$$\begin{aligned}|\widehat{v}(t,\tilde{y}) - \widehat{v}(t,\hat{y}^\epsilon)| &= \left|\mathbb{E}\left[f(e^{y+\widetilde{Y}^\epsilon_{T-t}}) - f(e^{y+\widehat{Y}^\epsilon_{T-t}})\right]\right| \le C\sum_{i=1}^{d}\mathbb{E}\left[\left|e^{y_i+\widetilde{Y}^{\epsilon,i}_{T-t}} - e^{y_i+\widehat{Y}^{\epsilon,i}_{T-t}}\right|\right] \\ &= C\sum_{i=1}^{d} e^{y_i}\mathbb{E}\left[e^{\mathbb{1}_{\{1\le i\le r(\epsilon)\}}s_iW^i_{T-t}}\right]\left|e^{\lambda_i(T-t)} - e^{\widehat{\lambda}_i(T-t)}\right| \\ &\le C(y)\sum_{i=1}^{d} e^{y_i+\frac{1}{2}s_i^2(T-t)}|\lambda_i - \widehat{\lambda}_i| \le C(y)\sum_{i=1}^{d} e^{y_i}(1+s_i^2)\sum_{j=r(\epsilon)+1}^{d} s_j^2 \\ &\le C(y)\sum_{j=r(\epsilon)+1}^{d} s_j^2.\end{aligned}$$

Since $y = \mathbf{U}x$, $C(y) = C'(x)$ which completes the proof. $\square$

2.2 Stochastic volatility models

Similarly to the one-dimensional case, multivariate stochastic volatility models replace the constant volatilities $\mathbf{\Sigma}_{ij}$ in the Black-Scholes model (2.3) by stochastic processes $\Sigma_{ij} = f_{ij}(Y)$, where f_{ij} are non-negative functions and Y is an additional source of randomness, which is modeled by an Itô diffusion in $\mathbb{R}^d$.

We consider the stochastic volatility extension of the Black-Scholes model as described in [1, Chapter 10.6]. We set $Z := (X, Y)$, where X describes again the log-price dynamics of $n > 1$ assets and Y is an $\mathbb{R}^n$-valued Itô diffusion describing the stochastic volatility $\Sigma_{ij} = f_{ij}(Y)$. In particular, we assume that each Y^i evolves according to the SDE

$$\mathrm{d}Y_t^i = c_i(Y_t^i)\,\mathrm{d}t + b_i(Y_t^i)\,\mathrm{d}\widetilde{W}_t^i, \quad Y_0^i = y^i, \quad i = 1, \dots, n.$$

We pose the following assumptions: the state space domain of Y is $G^Y \subseteq \mathbb{R}^n$, and the coefficients $c_k, b_k : G^Y \to \mathbb{R}$ are globally Lipschitz continuous and at most linearly growing. Furthermore, the $\mathbb{R}^n$-valued standard Brownian motion $(\widetilde{W}_t)_{t\geq 0}$ is correlated to the $\mathbb{R}^n$-valued standard Brownian motion $(W_t)_{t\geq 0}$ that drives the process X by $\widetilde{W}^k = \sum_{j=1}^n \rho_{jk} W^j + \rho^* \widehat{W}^k$, where $(W, \widehat{W})$ is a standard Brownian motion in $\mathbb{R}^d$ with $d = 2n$, and $\rho_k^* := (1 - \sum_{j=1}^n \rho_{jk}^2)^{1/2}$.

Denoting by $z := (x, y) = (x_1, \dots, x_n,\ y_1, \dots, y_n)$, the coefficients μ, Σ in (2.2) under a non-unique EMM are given by

$$\mu(z) := \big(\gamma - 1/2 f_{11}^2(y), \dots, \gamma - 1/2 f_{nn}^2(y), c_1(y_1), \dots, c_n(y_n)\big)^\top \in \mathbb{R}^d, \tag{2.15}$$

$$\Sigma(z) := \begin{pmatrix} \Sigma^X(z) & 0 \\ \Sigma^Y(z) & D(z) \end{pmatrix} \in \mathbb{R}^{d\times d}, \tag{2.16}$$

where the matrices $\Sigma^X, \Sigma^Y, D \in \mathbb{R}^{n\times n}$ are

$$\Sigma^X(z) := \big(f_{ij}(y)\big)_{1\leq i,j\leq n}, \quad \Sigma^Y(z) := \big(\rho_{ji} b_i(y_i)\big)_{1\leq i,j\leq n},$$
$$D(z) := \operatorname{diag}\big(\rho_1^* b_1(y_1), \dots, \rho_n^* b_n(y_n)\big).$$

The smooth functions $f_{ij} : G^Y \to \mathbb{R}_+$ are assumed to be bounded from below and above. The state space domain of the pair process $Z = (X, Y)$ is $G = \mathbb{R}^n \times G^Y$.

Example 2.4 (Volatility processes)**.** *In [1, Chapter 10.6], it is assumed that each volatility component Y^k follows a mean-reverting Ornstein-Uhlenbeck process, i.e., $c_k(y_k) = \alpha_k(m_k - y_k)$, $b_k(y_k) = \beta_k$, $1 \leq k \leq n$. Here, $\alpha_k > 0$ is called the rate of mean reversion and $m_k \geq 0$ is the long-run mean level of Y^k. Under an EMM, the drift term c_k becomes $c_k(y) = \alpha_k(m_k - y_k) - \beta_k \Lambda_k(y)$, for some volatility risk premium $\Lambda(y) = (\Lambda_1(y), \cdots, \Lambda_n(y))^\top$. See [1, Chapter 2.5] for a representation of Λ in the one dimensional case $n = 1$.*

3 Option pricing

We change to time-to-maturity $t \to T - t$ and let $J := (0, T]$. Assume that the pricing function $u(t, z)$ in (2.1) satisfies $u \in C^{1,2}(J \times G)$. Then, from the Feynman-Kac Theorem, u solves the parabolic partial differential equation

$$\begin{array}{rcll} \partial_t u + \mathcal{A}u + \gamma u & = & 0 & \text{in} \quad J \times G, \\ u(0, z) & = & g(z) & \text{in} \quad G, \end{array} \tag{3.1}$$

where the infinitesimal generator $\mathcal{A}$ of the semigroup generated by the process Z (2.2) is given by

$$\mathcal{A} := -\frac{1}{2}\mathrm{tr}\big[\mathcal{Q}(z)D^2\big] - \langle \mu(z), D \rangle, \tag{3.2}$$

with $\mathcal{Q} = \mathbf{\Sigma}\mathbf{\Sigma}^\top$, and the differential operators D^2, D given by $D^2 = (\partial_{x_i x_j})_{1 \le i,j \le d}$, $D = (\partial_{x_1}, \ldots, \partial_{x_d})^\top$. Furthermore, $\mathrm{tr} : \mathbb{R}^{d \times d} \to \mathbb{R}$ denotes the trace operator, i.e., $\mathrm{tr}\mathbf{B} = \sum_{i=1}^d \mathbf{B}_{ii}$, $\mathbf{B} \in \mathbb{R}^{d \times d}$, and $\langle \cdot, \cdot \rangle$ denotes the usual Euclidian scalar product in $\mathbb{R}^d$.

3.1 Localization

For numerical implementation we truncate the parabolic PDE (3.1) to a bounded domain G_R and impose boundary conditions on ∂G_R. Typically, G_R is a multidimensional hypercube, i.e., $G_R = \prod_{j=1}^d (a_j, b_j)$ for some $a_j, b_j \in \mathbb{R}$, $b_j > a_j$, $j = 1, \ldots, d$. Therefore, we consider the truncated problem

$$\begin{array}{rcll} \partial_t u_R + \mathcal{A}u_R + \gamma u_R & = & 0 & \text{in} \quad J \times G_R, \\ u_R(t, \cdot) & = & 0 & \text{on} \quad J \times \partial G_R, \\ u_R(0, z) & = & g(z)|_{G_R} & \text{in} \quad G_R. \end{array} \tag{3.3}$$

The truncation to a bounded domain G_R amounts to approximating the solution u of (3.1) by the function u_R pricing a barrier option. The function u_R is given by

$$u_R(t, z) = \mathbb{E}^{\mathbb{Q}}\big[e^{-\gamma(T-t)} g(Z_T) \mathbb{1}_{\{T < \tau_{G_R}\}} \mid Z_t = z\big],$$

where $\tau_{G_R} = \inf\{s \ge 0 \mid Z_s \in G_R^c\}$ is the first hitting time of the complement $G_R^c = G \setminus G_R$ by the process Z. For $R := \max_{1 \le j \le d} |b_j - a_j|$ we assume that u_R converges exponentially fast in R to u on a subset of $G_0 \subset G_R$. This holds for general multidimensional Lévy models as shown in [7, Theorem 4.14] and for certain stochastic volatility models as in [3, Theorem 3.6]. Hence, we neglect the truncation error.

3.2 Variational formulation

The finite element method (FEM) is based on the weak or variational formulation of the pricing equation (3.3). Its functional setting can be described as follows.

Let

$$V \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} V^*,$$

be a Gelfand triplet. We denote by $\|\cdot\|$, $\|\cdot\|_V$ the norms in H, V, by $(\cdot,\cdot)$ the inner product in H and by $\langle\cdot,\cdot\rangle_{V^*\times V}$ the duality pairing between V and its dual V^*. To the infinitesimal generator $\mathcal{A}$ in (3.2) we associate the Dirichlet form $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ via

$$a(u, v) := \langle \mathcal{A}u, v\rangle_{V^*\times V} + \gamma(u, v), \quad u, v \in V, \tag{3.4}$$

and assume: $\exists \alpha_1, \alpha_2 > 0$, $\alpha_3 \geq 0$ such that $\forall u, v \in V$

$$|a(u, v)| \leq \alpha_1 \|u\|_V \|v\|_V, \qquad a(v, v) \geq \alpha_2 \|v\|_V^2 - \alpha_3 \|v\|^2. \tag{3.5}$$

In the Black-Scholes setting property (3.5) holds with the spaces $V = H_0^1(G_R)$ and $H = L^2(G_R)$ as shown in [7, Theorem 4.8]. For stochastic volatility models we obtain (3.5) in weighted Sobolev spaces [2, 3]. The weak formulation of (3.3) reads: Given $g \in H$, find $u_R \in L^2(J; V) \cap H^1(J; V^*)$ such that

$$(\partial_t u_R(t, \cdot), v) + a(u_R(t, \cdot), v) = 0, \quad \forall v \in V. \tag{3.6}$$

Note that the homogeneous Dirichlet boundary conditions (3.3) are imposed on the space V. If (3.5) holds, then $\mathcal{A} + \alpha_3 I \in \mathcal{L}(V, V^*)$ is an isomorphism and the weak formulation (3.6) admits a unique (weak) solution $u_R \in L^2(J; V) \cap H^1(J; V^*)$. By integration by parts and by the homogeneous essential boundary conditions, it follows that the Dirichlet form $a(\cdot,\cdot)$ in (3.4) associated to the operator $\mathcal{A}$ in (3.2) is given by

$$\begin{aligned} a(\varphi, \phi) = \frac{1}{2}\int_{G_R} \nabla\varphi^\top \mathcal{Q}(x)\nabla\phi \,\mathrm{d}x + \int_{G_R} \langle \frac{1}{2}\,\underline{\nabla}\mathcal{Q}(x) + \mu(x), \nabla\varphi\rangle \phi \,\mathrm{d}x \\ + \int_{G_R} \gamma\varphi\,\phi \,\mathrm{d}x, \end{aligned} \tag{3.7}$$

where for a matrix $\mathbf{B} \in [H^1(G_R)]^{d\times d}$, we denote by $\underline{\nabla}\mathbf{B}$ (with a slight abuse of notation) the vector

$$\underline{\nabla}\mathbf{B} := (\nabla \underline{\mathbf{B}}_j)_{j=1}^d, \quad \underline{\mathbf{B}}_j := (B_{1j}, \ldots, B_{dj})^\top \in [H^1(G_R)]^d.$$

4 Discretization

Straightforward application of standard finite element schemes for discretizing (3.6) fails due to the "curse of dimension": the number of degrees of freedom on a tensor product finite element mesh of uniform width h in dimension d grows like $\mathcal{O}(h^{-d})$ as $h \to 0$. Spline wavelets can overcome the problem while still being easy to compute. Choosing wavelet bases has twofold advantages. Firstly, we can break the curse of dimension using sparse tensor products to obtain essentially dimension independent complexity. Secondly, wavelets provide norm equivalences which lead to efficient preconditioning of the resulting linear system.

4.1 Space and time discretization

The abstract finite element semi-discretization in (log) price space of (3.6) reads: given a finite dimensional subspace $V_h \subset V$ with $\dim V_h = N < \infty$, find $u_h \in L^2(J;V_h) \cap H^1(J;V_h^*)$ such that $u_h(0,\cdot) = u_h^0$ and such that

$$(\partial_t u_h(t,\cdot), v) + a(u_h(t,\cdot), v) = 0, \quad \forall v \in V_h, \quad \forall t \in J, \tag{4.1}$$

where u_h^0 is the H-projection of the payoff g onto V_h, i.e.,

$$(u_h^0, v) = (g, v), \quad \forall v \in V_h. \tag{4.2}$$

Choosing a basis $\mathcal{B} := \{\Phi_j\}_{j=1}^N$ of V_h, (4.1) is equivalent to: Given $\underline{u}_h^0 \in \mathbb{R}^N$, find $\underline{u}_h(t) \in \mathbb{R}^N$ such that

$$\mathbf{M}\underline{\dot{u}}_h(t) + \mathbf{A}\underline{u}_h(t) = 0, \tag{4.3}$$

where $\underline{u}_h(t)$ denotes the coefficient vector of $u_h(t,\cdot)$ with respect to the basis $\mathcal{B}$ of V_h. In (4.3), the matrices $\mathbf{M}, \mathbf{A} \in \mathbb{R}^{N\times N}$ are the so-called mass and moment or stiffness matrix, respectively, which are given by

$$\mathbf{M} := \big((\Phi_{i'}, \Phi_i)\big)_{1\le i,i'\le N}, \quad \mathbf{A} := \big(a(\Phi_{i'}, \Phi_i)\big)_{1\le i,i'\le N}. \tag{4.4}$$

In order to discretize in time, we use the θ-scheme. For $M \in \mathbb{N}$, define the time step $k := TM^{-1}$ and the time grid points $t^m := km$, $m = 0,\dots,M$. The fully discrete scheme to (3.6) reads: Given $u_h^0 \in V_h$, for $m = 0,\dots,M-1$ find $u_h^{m+1} \in V_h$ such that

$$\big(k^{-1}(u_h^{m+1} - u_h^m), v\big) + a(u_h^{m+\theta}, v) = 0, \quad \forall v \in V_h. \tag{4.5}$$

Here, $u_h^{m+\theta} := \theta u_h^{m+1} + (1-\theta)u_h^m$, $\theta \in [0,1]$. In matrix form, (4.5) reads: Given $\underline{u}_h^0 \in \mathbb{R}^N$, find $\underline{u}_h^{m+1} \in \mathbb{R}^N$ such that

$$\big(\mathbf{M} + \theta k\mathbf{A}\big)\underline{u}_h^{m+1} = \big(\mathbf{M} + (1-\theta)k\mathbf{A}\big)\underline{u}_h^m, \quad m = 0,1,\dots,M-1, \tag{4.6}$$

Remark 4.1. *It is well known that the convergence of the θ-scheme is of first order if $\theta \in [0,1]\setminus\{1/2\}$, and of second order if $\theta = 1/2$. For the more involved hp-dG time stepping scheme, which converges exponentially under sufficient smoothness conditions, see [9] and the references therein.*

Remark 4.2. *Two issues arise in the choice of the finite element space V_h. First, for classical tensor product Lagrange finite element spaces, the dimension N grows exponentially in the dimension d of the basket, i.e., $N = \mathcal{O}(h^{-d})$ as the mesh width $h \to 0$. Second, the condition number κ of the stiffness matrix $\mathbf{A}$ behaves like $\kappa \sim h^{-2d}$. Clearly, $\mathbf{A}$ is ill-conditioned, which makes preconditioning necessary. To overcome "the curse of dimension", we shall use the aforementioned sparse tensor product spaces and to get efficient and easily implementable preconditioners, these spaces are spanned by spline wavelets which are used as $\mathcal{B}$.*

4.2 Spline wavelets

We describe wavelet finite elements in the interval $G = (0, 1)$. Define the mesh $\mathcal{T}_\ell$ given by the nodes $j2^{-\ell}$, $j = 0, \dots, 2^\ell$, with mesh-width $h_\ell = 2^{-\ell}$. Let $\mathcal{V}_\ell$ be the space of continuous piecewise linear polynomials[1] on the mesh $\mathcal{T}_\ell$ which vanish on ∂G. We write $N_\ell := \dim \mathcal{V}_\ell$, $N_{-1} := 0$ and $M_\ell := N_\ell - N_{\ell-1}$. We use a wavelet basis $\psi_{\ell,k}$, $k = 1, \dots, M_\ell$, $\ell = 0, 1, 2, \dots$ of $\mathcal{V}_\ell$ with the properties,

$$\mathcal{V}_L = \operatorname{span}\{\psi_{\ell,k} \mid 0 \le \ell \le L;\ 1 \le k \le M_\ell\}, \quad \operatorname{diam}(\operatorname{supp}\psi_{\ell,k}) \le C2^{-\ell}. \tag{4.7}$$

Any function $v \in \mathcal{V}_L$ has the representation

$$v = \sum_{\ell=0}^{L} \sum_{k=0}^{M_\ell} v_{\ell,k}\psi_{\ell,k}, \tag{4.8}$$

with the coefficients $v_{\ell,k} = (v, \widetilde{\psi}_{\ell,k})$, where the $\widetilde{\psi}_{\ell,k}$ are the dual wavelets. For $v \in L^2(G)$, one obtains the series

$$v = \sum_{\ell=0}^{\infty} \sum_{k=0}^{M_\ell} v_{\ell,k}\psi_{\ell,k}, \tag{4.9}$$

which converges in $L^2(G)$ and in $H^1_0(G)$. Moreover, for $v \in \widetilde{H}^s(G)$ where $\widetilde{H}^s(G) := [L^2(G), H^1_0(G)]_{s,2}$ there holds the *norm equivalence*

$$c_1\|v\|^2_{\widetilde{H}^s(G)} \le \sum_{\ell=0}^{\infty} \sum_{k=0}^{M_\ell} 2^{2\ell s}|v_{\ell,k}|^2 \le c_2\|v\|^2_{\widetilde{H}^s(G)}, \quad 0 \le s \le 1. \tag{4.10}$$

For $v \in L^2(G)$ we define a bi-orthogonal projection $P_L : L^2(G) \to \mathcal{V}_L$ by truncating (4.9):

$$P_L v := \sum_{\ell=0}^{L} \sum_{k=0}^{M_\ell} v_{\ell,k}\psi_{\ell,k}, \qquad P_{-1} := 0. \tag{4.11}$$

This projection satisfies the approximation property

$$\|u - P_L u\|_{\widetilde{H}^s(G)} \le c2^{-(t-s)L}\|u\|_{H^t(G)}, \quad 0 \le s \le 1,\ s \le t \le p+1. \tag{4.12}$$

The increment or detail spaces W_ℓ are defined by

$$W_\ell := \operatorname{span}\{\psi_{\ell,k} \mid 1 \le k \le M_\ell\},\ \ell = 1, 2, 3, \dots, \quad W_0 := V_0. \tag{4.13}$$

Then

$$\mathcal{V}_\ell = \mathcal{V}_{\ell-1} \oplus W_\ell \text{ for } \ell \ge 1, \quad \text{and} \quad \mathcal{V}_\ell = \bigoplus_{j=0}^{\ell} W_j,\ \ell \ge 0, \tag{4.14}$$

and $Q_\ell := P_\ell - P_{\ell-1}$ is a projection from $L^2(G)$ onto W_ℓ.

[1] Higher order polynomials of degree $p > 1$ are also possible.

Example 4.3. *For $\ell \geq 0$ let $\mathcal{T}^\ell$ be mesh $\mathcal{T}^\ell = \{0 = x_0^\ell < \ldots < x_{2^{\ell+1}}^\ell = 1\}$ with mesh points $x_j^\ell = j2^{-\ell-1}$, $j = 0, \ldots, 2^{\ell+1}$. Note that $N^\ell = 2^{\ell+1} - 1$ and $M_\ell = 2^\ell$. On the coarsest level $\ell = 0$, let $\psi_{0,1}(x) = \gamma_0 \max(0, 1 - 2|x - 1/2|)$. On finer levels $\ell \geq 1$ the left boundary wavelet $\psi_{\ell,1}$ has values $\psi_{\ell,1}(x_1^\ell) = 2\gamma_\ell$, $\psi_{\ell,1}(x_2^\ell) = -\gamma_\ell$ and zero at all other mesh points. The right boundary wavelet ψ_{ℓ,M_ℓ} has values $\psi_{\ell,M_\ell}(x_{N^\ell}^\ell) = 2\gamma_\ell$, $\psi_{\ell,M_\ell}(x_{N^\ell-1}^\ell) = -\gamma_\ell$ and zero at all other mesh points. The interior wavelet $\psi_{\ell,k}$, $1 < k < M_\ell$ has values $\psi_{\ell,k}(x_{2k-2}^\ell) = \psi_{\ell,k}(x_{2k}^\ell) = -\gamma_\ell$ and $\psi_{\ell,k}(x_{2k-1}^\ell) = 2\gamma_\ell$. We choose the constant γ_ℓ such that $\|\psi_{\ell,k}\|_{L^2(G)} = 1$. Thus, $\gamma_0 = \sqrt{3}$ and $\gamma_\ell = \sqrt{3}/2\, 2^{\ell/2}$, $\ell \geq 1$.*

4.3 Sparse tensor product spaces

For $d > 1$, let $G = (0,1)^d$ and define the full tensor product space V_L as the d-fold tensor product of the spaces $\mathcal{V}_L$ as $V_L := \mathcal{V}_L \otimes \cdots \otimes \mathcal{V}_L$ which can be written as

$$V_L = \operatorname{span}\{\psi_{\boldsymbol{\ell},\mathbf{k}} \ : \ 0 \leq \ell_i \leq L,\, 1 \leq k_i \leq M_{\ell_i}, i = 1, \ldots, d\},$$

with basis functions $\psi_{\boldsymbol{\ell},\mathbf{k}} = \psi_{\ell_1,k_1} \otimes \cdots \otimes \psi_{\ell_d,k_d}$ and multi-indices $\boldsymbol{\ell} = (\ell_1, \ldots \ell_d) \in \mathbb{N}_0^d$, $\mathbf{k} = (k_1, \ldots, k_d) \in \mathbb{N}^d$. We define the norms $|\boldsymbol{\ell}|_\infty := \max\{\ell_j \ : \ 1 \leq j \leq d\}$ and $|\boldsymbol{\ell}|_1 := \ell_1 + \cdots + \ell_d$. Using (4.14), the space V_L can be written as

$$V_L = \bigoplus_{|\boldsymbol{\ell}|_\infty \leq L} W_{\ell_1} \otimes \cdots \otimes W_{\ell_d}, \tag{4.15}$$

and the sparse tensor product space $\widehat{V}_L$ at level $L \geq 0$ as

$$\widehat{V}_L := \bigoplus_{|\boldsymbol{\ell}|_1 \leq L} W_{\ell_1} \otimes \cdots \otimes W_{\ell_d}, \tag{4.16}$$

where the increment spaces W_{ℓ_i}, $1 \leq i \leq d$, are as in (4.13).

As $L \to \infty$, we have $N_L := \dim(V_L) = O(2^{dL})$, and $\widehat{N}_L := \dim(\widehat{V}_L) = O(2^L L^{d-1})$, i.e., the spaces $\widehat{V}_L$ have considerably smaller dimensions than V_L. However, both spaces have similar approximation properties, provided the function to approximate is sufficiently smooth. To characterize the extra smoothness requirements, we introduce the spaces $\mathcal{H}^s(G)$, $s \in \mathbb{N}_0$, of all measurable functions $u : G \to \mathbb{R}$ such that the norm

$$\|u\|_{\mathcal{H}^s}^2 := \sum_{\substack{0 \leq \alpha_i \leq s \\ i=1,\ldots,d}} \|\partial_{x_1}^{\alpha_1} \cdots \partial_{x_d}^{\alpha_d} u\|_{L^2(G)}^2, \tag{4.17}$$

is finite. That is, $\mathcal{H}^s([0,1]^d) = \otimes_{j=1}^d H^s([0,1])$. For $s \geq 0$ the space $\mathcal{H}^s$ is defined by interpolation. Due to the underlying tensor product structure, one infers from (4.10) that for

$$v = \sum_{\substack{\ell_i = 0 \\ i=1,\ldots,d}}^{\infty} \sum_{1 \leq k_i \leq M_{\ell_i}} v_{\boldsymbol{\ell},\mathbf{k}} \psi_{\boldsymbol{\ell},\mathbf{k}},$$

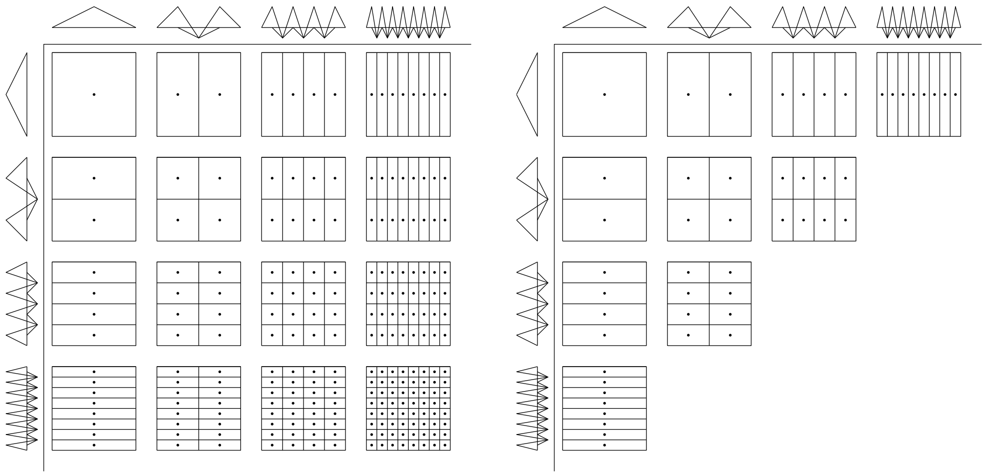

Figure 1: Schematic of full tensor product space V_L (left) and sparse tensor product space $\widehat{V}_L$ (right) for $L = 3$. For both spaces the spline wavelets of Example 4.3 are displayed.

there holds the norm equivalence

$$c_1\|v\|^2_{\mathcal{H}^s} \leq \sum_{\substack{\ell_i=0\\ i=1,\ldots,d}}^{\infty} \sum_{1\leq k_i\leq M_{\ell_i}} 2^{2s|\boldsymbol{\ell}|_1}|v_{\boldsymbol{\ell},\mathbf{k}}|^2 \leq c_2\|v\|^2_{\mathcal{H}^s}, \quad 0 \leq s \leq 1.$$

We now derive a representation of the matrices $\mathbf{M}, \mathbf{A}$ given by (4.4) in the space $V_h := \widehat{V}_L$. To this end, for some function $w : \mathbb{R} \to \mathbb{R}$, define the matrices $\mathbf{M}^w$, $\mathbf{S}^w$ and $\mathbf{C}^w$ with respect to the i-th coordinate direction

$$\mathbf{M}^{w(x_i)} := \Big(\int_{a_i}^{b_i} \psi_{\ell_i,k_i}(x_i)\psi_{\ell_i',k_i'}(x_i)w(x_i)\,\mathrm{d}x_i\Big)_{\substack{0\leq \ell_i',\ell_i\leq L\\ 1\leq k_i'\leq M_{\ell_i'},1\leq k_i\leq M_{\ell_i}}} \tag{4.18}$$

$$\mathbf{S}^{w(x_i)} := \Big(\int_{a_i}^{b_i} \psi'_{\ell_i,k_i}(x_i)\psi'_{\ell_i',k_i'}(x_i)w(x_i)\,\mathrm{d}x_i\Big)_{\substack{0\leq \ell_i',\ell_i\leq L\\ 1\leq k_i'\leq M_{\ell_i'},1\leq k_i\leq M_{\ell_i}}} \tag{4.19}$$

$$\mathbf{C}^{w(x_i)} := \Big(\int_{a_i}^{b_i} \psi'_{\ell_i,k_i}(x_i)\psi_{\ell_i',k_i'}(x_i)w(x_i)\,\mathrm{d}x_i\Big)_{\substack{0\leq \ell_i',\ell_i\leq L\\ 1\leq k_i'\leq M_{\ell_i'},1\leq k_i\leq M_{\ell_i}}} \tag{4.20}$$

Let $\mathbf{X}^i$ be any matrix given by (4.18)–(4.20). We view the matrix $\mathbf{X}^i$ as a collection of block matrices, i.e.,

$$\mathbf{X}^i = (\mathbf{X}^i_{\ell',\ell})_{0\leq\ell',\ell\leq L}, \quad \text{where} \quad \mathbf{X}^i_{\ell',\ell} := (\mathbf{X}^i_{(\ell',k'),(\ell,k)})_{1\leq k'\leq M_{\ell'},1\leq k\leq M_\ell},$$

and define a sparse tensor product $\mathbf{X}^1 \mathbin{\widehat{\otimes}} \mathbf{X}^2 \mathbin{\widehat{\otimes}} \cdots \mathbin{\widehat{\otimes}} \mathbf{X}^d$ by tensor products of block matrices

$$\mathbf{X}^1 \mathbin{\widehat{\otimes}} \cdots \mathbin{\widehat{\otimes}} \mathbf{X}^d := \Big(\mathbf{X}^1_{\ell_1',\ell_1} \otimes \cdots \otimes \mathbf{X}^d_{\ell_d',\ell_d}\Big)_{0\leq|\boldsymbol{\ell}'|_1,|\boldsymbol{\ell}|_1\leq L}$$

for multiindices $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_d)$, $\boldsymbol{\ell}' = (\ell_1', \ldots, \ell_d')$.

Definition 4.4. *For an arbitrary permutation σ,*

$$\sigma : \{1, \dots, d\} \to \{1, \dots, d\}, \quad \{1, \dots, d\} \mapsto \{\sigma(1), \dots, \sigma(d)\},$$

and matrices $\mathbf{X}^{w(x_i)}$, $1 \le i \le d$, $\mathbf{X} \in \{\mathbf{S}, \mathbf{C}, \mathbf{M}\}$, we denote by

$$^{s}\big(\mathbf{X}^{w(x_{\sigma(1)})} \widehat{\otimes} \cdots \widehat{\otimes} \mathbf{X}^{w(x_{\sigma(d)})}\big),$$

the sorted sparse tensor product with factors sorted by increasing indices, i.e.,

$$^{s}\big(\mathbf{X}^{w(x_{\sigma(1)})} \widehat{\otimes} \cdots \widehat{\otimes} \mathbf{X}^{w(x_{\sigma(d)})}\big) := \mathbf{X}^{w(x_1)} \widehat{\otimes} \cdots \widehat{\otimes} \mathbf{X}^{w(x_d)}.$$

Lemma 4.5. *Assume that the coefficients $\mathcal{Q}(x)$ and $\mu(x)$ of the operator $\mathcal{A}$ in (3.2) are given by*

$$\begin{aligned}
\mathcal{Q}(x) = \big(\mathcal{Q}_{ij}(x)\big)_{1 \le i,j \le d} &= \left(\prod_{\ell=1}^{d} q_\ell^{ij}(x_\ell)\right)_{1 \le i,j \le d}, \\
\mu(x) = \big(\mu_i(x)\big)_{1 \le i \le d} &= \left(\prod_{\ell=1}^{d} \mu_\ell^{i}(x_\ell)\right)_{1 \le i \le d},
\end{aligned}$$

for univariate functions $q_\ell^{ij}, \mu_\ell^i \to \mathbb{R}$, $1 \le i, j, \ell \le d$. Then, the stiffness matrix $\mathbf{A}$ in (4.4) of the bilinear form $a(\cdot,\cdot)$ in (3.7) with respect to the sparse tensor product space $\widehat{V}_L$ is given by

$$\begin{aligned}
\mathbf{A} = {} & \frac{1}{2} \sum_{j=1}^{d} {}^{s}\Bigg(\mathbf{S}^{q_j^{jj}(x_j)} \widehat{\otimes} \widehat{\bigotimes_{\substack{1 \le i \le d \\ i \ne j}}} \mathbf{M}^{q_i^{jj}(x_i)}\Bigg) \\
& - \frac{1}{2} \sum_{\substack{j,k=1 \\ j \ne k}}^{d} {}^{s}\Bigg(\mathbf{C}^{q_j^{jk}(x_j)} \widehat{\otimes} \Big(\mathbf{C}^{q_k^{jk}(x_k)} + \mathbf{M}^{\frac{\mathrm{d}}{\mathrm{d}x_k} q_k^{jk}(x_k)}\Big) \widehat{\otimes} \widehat{\bigotimes_{\substack{1 \le i \le d \\ i \notin \{j,k\}}}} \mathbf{M}^{q_i^{jk}(x_i)}\Bigg) \\
& + \frac{1}{2} \sum_{j=1}^{d} {}^{s}\Bigg(\mathbf{C}^{\frac{\mathrm{d}}{\mathrm{d}x_j} q_j^{jj}(x_j)} \widehat{\otimes} \widehat{\bigotimes_{\substack{1 \le i \le d \\ i \ne j}}} \mathbf{M}^{q_i^{jj}(x_i)}\Bigg) \\
& + \frac{1}{2} \sum_{\substack{j,k=1 \\ j \ne k}}^{d} {}^{s}\Bigg(\mathbf{C}^{q_k^{jj}(x_k)} \widehat{\otimes} \widehat{\bigotimes_{\substack{1 \le i \le d \\ i \ne k}}} \mathbf{M}^{\widetilde{q}_i^{jk}(x_i)}\Bigg) \\
& + \sum_{j=1}^{d} {}^{s}\Bigg(\mathbf{C}^{\mu_j^{j}(x_j)} \widehat{\otimes} \widehat{\bigotimes_{\substack{1 \le i \le d \\ i \ne j}}} \mathbf{M}^{\mu_i^{j}(x_i)}\Bigg) + \gamma \widehat{\bigotimes_{1 \le i \le d}} \mathbf{M}
\end{aligned}$$

with weights

$$\widetilde{q}_i^{jk}(x_i) := \begin{cases} q_i^{jk}(x_i) & \text{if } i \ne j \\ \frac{\mathrm{d}}{\mathrm{d}x_i} q_i^{jk}(x_i) & \text{if } i = j \end{cases}.$$

Proof. This follows by elementary, however lengthy, calculations. ∎

4.4 Matrix-vector multiplication

Computing the matrix $\mathbf{A}$ explicitly for $d \gg 1$ requires too much memory. But for solving the ordinary differential equation (4.3) using a time-stepping scheme and an iterative solver, we only need to compute matrix-vector multiplications $\underline{v} = \mathbf{A}\underline{u}$. Using the (sparse) tensor product structure this can be done without computing the matrix $\mathbf{A}$ explicitly.

Let $\mathbf{A} := \mathbf{X}^1 \mathbin{\widehat{\otimes}} \cdots \mathbin{\widehat{\otimes}} \mathbf{X}^d \in \mathbb{R}^{\widehat{N}_L \times \widehat{N}_L}$ and $u_L \in \widehat{V}_L$. We again view the coefficient vector $\underline{u}_L \in \mathbb{R}^{\widehat{N}_L}$ of u_L as a collection of block coefficient vectors,

$$\underline{u}_L = \left(\underline{u}_{\boldsymbol{\ell}}\right)_{0 \le |\boldsymbol{\ell}|_1 \le L}, \quad \text{where} \quad \underline{u}_{\boldsymbol{\ell}} = \left(u_{\boldsymbol{\ell},\mathbf{k}}\right)_{1 \le k_i \le M_{\ell_i}}.$$

The matrix-vector multiplication

$$\underline{v}_L = \mathbf{A}\underline{u}_L = \left(\mathbf{X}^1_{\ell_1',\ell_1} \mathbin{\widehat{\otimes}} \cdots \mathbin{\widehat{\otimes}} \mathbf{X}^d_{\ell_d',\ell_d}\right)_{0 \le |\boldsymbol{\ell}'|_1, |\boldsymbol{\ell}|_1 \le L} \left(\underline{u}_{\boldsymbol{\ell}}\right)_{0 \le |\boldsymbol{\ell}|_1 \le L}$$

is defined by

$$v_{\boldsymbol{\ell}',\mathbf{k}'} = \sum_{|\boldsymbol{\ell}|_1 < L} \sum_{1 \le k_i \le M_{\ell_i}} X^1_{(\ell_1',k_1'),(\ell_1,k_1)} \cdots X^d_{(\ell_d',k_d'),(\ell_d,k_d)} u_{\boldsymbol{\ell},\mathbf{k}}.$$

This multiplication may be computed iteratively as proposed in Algorithm 1.

Algorithm 1 Sparse grid matrix-vector multiplication

Set $v = u$
For $j = 0, 1, \ldots, d$
    For $\left|\boldsymbol{\ell}'\right|_1 = 0, 1, \ldots, L$
        Compute $v_{\boldsymbol{\ell}',\mathbf{k}'} = \sum_{\ell_j,k_j} X^j_{(\ell_j',k_j'),(\ell_j,k_j)} v_{\boldsymbol{\ell},\mathbf{k}}, \quad \forall \mathbf{k}'$,
        with $\ell_i = \ell_i'$, $k_i = k_i'$, $\quad \forall i \neq j$.
    Next $\boldsymbol{\ell}'$
Next j

4.5 Initial condition

Recall that u_h^0 is the H-projection of the payoff g onto V_h (4.2). Thus, $\underline{u}_h^0$ is the unique solution of the system $\mathbf{M}\underline{u}_h^0 = \underline{g}$, with right hand side $\underline{g} := ((g, \Phi_j))_{j=1}^N$. The realization of $\underline{g}$ is non-trivial. Let V_h be given by the sparse tensor product space $\widehat{V}_L$ in (4.16). Then, an arbitrary entry $g_{(\boldsymbol{\ell},\mathbf{k})}$ of $\underline{g}$ is given by

$$g_{(\boldsymbol{\ell},\mathbf{k})} = \int_{G_R} g(x_1, \ldots, x_d) \psi_{\ell_1,k_1}(x_1) \cdots \psi_{\ell_d,k_d}(x_d) \, \mathrm{d}x_1 \cdots \mathrm{d}x_d.$$

If g factorizes, i.e., $g(x_1,\dots,x_d) = \prod_{j=1}^d g_j(x_j)$ for some univariate $g_j : \mathbb{R} \to \mathbb{R}$, $j = 1,\dots,d$, then

$$g_{(\ell,k)} = \prod_{j=1}^d \int_{a_j}^{b_j} g_j(x_j)\psi_{\ell_j,k_j}(x_j)\,\mathrm{d}x_j.$$

However, for most payoffs in option pricing (e.g., basket options), the factorizing property does not hold. While numerical quadrature is applicable, we rather use integration by parts to find, in the sense of distributions,

$$g_{(\boldsymbol{\ell},\mathbf{k})} = \int_{G_R} g^{(-2)}(x_1,\dots,x_d)\psi''_{\ell_1,k_1}(x_1)\cdots\psi''_{\ell_d,k_d}(x_d)\,\mathrm{d}x_1\cdots\mathrm{d}x_d, \tag{4.21}$$

where

$$g^{(-k)}(x) := \int_{[x_0,x]} g^{(-k+1)}(y)\,\mathrm{d}y, \quad x \in \mathbb{R}^d, \quad k \geq 1,$$

for a suitable $x_0 \in \mathbb{R} \cup \{-\infty\}$. Let $\psi_{\ell_i,k_i} \in \widehat{V}_L$ be a continuous, piecewise linear spline inner wavelet and denote its singular support by

$$\operatorname{singsupp}\psi_{\ell_i,k_i} =: \{x^1_{\ell_i,k_i},\dots,x^{n_i}_{\ell_i,k_i}\}.$$

Then, the integral in (4.21) becomes

$$g_{(\boldsymbol{\ell},\mathbf{k})} = \sum_{\substack{1\leq j_i\leq n_i\\ 1\leq i\leq d}} g^{(-2)}\big(x^{j_1}_{\ell_1,k_1},\dots,x^{j_d}_{\ell_d,k_d}\big)\omega^{j_1}_{\ell_1,k_1}\cdots\omega^{j_d}_{\ell_d,k_d},$$

where the weights $\omega^1_{\ell_i,k_i},\dots,\omega^{n_i}_{\ell_i,k_i} \in \mathbb{R}$ depend only on the wavelet ψ_{ℓ_i,k_i}. As an example, consider the L^2-normalized wavelets $\psi_{\ell_i,k_i} : (a_i,b_i) \to \mathbb{R}$ defined on a interval (a_i,b_i) as described in Example 4.3. Then

$$(\omega^1_{\ell_i,k_i},\omega^2_{\ell_i,k_i},\omega^3_{\ell_i,k_i},\omega^4_{\ell_i,k_i},\omega^5_{\ell_i,k_i}) = \sqrt{3}(b_i-a_i)^{-\frac{3}{2}}2^{\frac{3}{2}\ell_i}(-1,4,-6,4,-1),$$

if $\ell_i \geq 1$ and ψ_{ℓ_i,k_i} is an interior wavelet, and

$$\begin{aligned}(\omega^1_{\ell_i,1},\omega^2_{\ell_i,1},\omega^3_{\ell_i,1},\omega^4_{\ell_i,1}) &= \sqrt{3}(b_i-a_i)^{-\frac{3}{2}}2^{\frac{3}{2}\ell_i}(2,-5,4,-1),\\ (\omega^1_{\ell_i,M^{\ell_i}},\omega^2_{\ell_i,M^{\ell_i}},\omega^3_{\ell_i,M^{\ell_i}},\omega^4_{\ell_i,M^{\ell_i}}) &= \sqrt{3}(b_i-a_i)^{-\frac{3}{2}}2^{\frac{3}{2}\ell_i}(-1,4,-5,2),\end{aligned}$$

if $\ell_i \geq 1$ and $\psi_{\ell_i,1},\psi_{\ell_i,M^{\ell_i}}$ is a left or a right boundary wavelet, respectively.

4.6 Multilevel preconditioning

As shown in (4.6), we have to solve M linear systems $\mathbf{B}\underline{u}_h^{m+1} = \underline{f}_h^m$, with $\mathbf{B} := \mathbf{M} + k\theta\mathbf{A}$, $\underline{f}_h^m := (\mathbf{M} + k(1-\theta)\mathbf{A})\underline{u}_h^m$. Due to the norm equivalence (4.10) we can build a simple preconditioner for the ill-conditioned matrix $\mathbf{B}$ using the wavelet basis. For simplicity, we restrict the discussion to the Black-Scholes model.

The norm equivalence (4.10) with $s = 0$ implies for every $u \in \widehat{V}_L$ with coefficient vector $\underline{u} \in \mathbb{R}^{\widehat{N}_L}$

$$C_1 \|u\|^2 \leq \langle \underline{u}, \mathbf{M}\underline{u} \rangle \leq C_2 \|u\|^2,$$

with constants C_1, C_2 independent of L. Denote by $\mathbf{D}_A$ the diagonal matrix with entries $2^{2\ell_1} + \cdots + 2^{2\ell_d}$ for an index corresponding to level $(\ell_1, \ldots, \ell_d)$. Then (3.5) and (4.10) with $s = 1$ imply that

$$C_1 \langle \underline{u}, \mathbf{D}_A \underline{u} \rangle \leq \langle \underline{u}, \mathbf{A}\underline{u} \rangle \leq C_2 \langle \underline{u}, \mathbf{D}_A \underline{u} \rangle.$$

Thus, we have $C_1 \langle \underline{u}, \mathbf{D}\underline{u} \rangle \leq \langle \underline{u}, \mathbf{B}\underline{u} \rangle \leq C_2 \langle \underline{u}, \mathbf{D}\underline{u} \rangle$, with the diagonal matrix $\mathbf{D} := \mathbf{I} + k\theta \mathbf{D}_A$ where $\mathbf{I}$ denotes the indentity matrix. Written in terms of $\underline{\widehat{u}} := \mathbf{D}^{1/2}\underline{u}$, we finally obtain

$$C_1 \|\underline{\widehat{u}}\|^2 \leq \langle \underline{\widehat{u}}, \mathbf{D}^{-1/2}\mathbf{B}\mathbf{D}^{-1/2}\underline{\widehat{u}} \rangle \leq C_2 \|\underline{\widehat{u}}\|^2.$$

The linear system $\widehat{\mathbf{B}}\underline{\widehat{u}} = \underline{\widehat{f}}$ with preconditionend matrix $\widehat{\mathbf{B}} := \mathbf{D}^{-1/2}\mathbf{B}\mathbf{D}^{-1/2}$ and right hand side $\underline{\widehat{f}} = \mathbf{D}^{-1/2}\underline{f}$ can be solved with GMRES in a number of iteration steps which is independent of level index L and thus independent of the size of the linear system.

5 Price sensitivities

Calculating price sensitivities is a central modeling and computational task for risk management and hedging. We distinguish between two classes of sensitivities: the sensitivity of the solution u to variation of a model parameter, like the Greek Vega ($\partial_\sigma u$) and the sensitivity of the solution u to a variation of state spaces such as the Greek Delta ($\partial_x u$). It is shown in [4] that an approximation for the first class can be obtained as a solution of the pricing PDE with a right hand side depending on u. For the second class, a finite difference like differentiation procedure is presented which allows to obtain the sensitivities from the finite element forward price without additional forward solver.

5.1 Sensitivity with respect to model parameters

Suppose the market model, and hence the operator $\mathcal{A}$ in (3.2), depend on some model parameter ϑ. We want to calculate the sensitivity of the solution u of (3.1) with respect to ϑ. To this end, we write $u(\vartheta_0)$ for a fixed realization ϑ_0 of ϑ in order to emphasize the dependence of u on ϑ_0.

Let $\mathcal{C}$ be a Banach space over a domain $G \subset \mathbb{R}^d$. $\mathcal{C}$ is the space of parameters or coefficients in the operator $\mathcal{A}$ and $\mathcal{S}_\vartheta \subseteq \mathcal{C}$ is the set of admissible coefficients. We denote by $u(\vartheta_0)$ the unique solution to (3.1) and introduce the derivative of $u(\vartheta_0)$

with respect to $\vartheta_0 \in \mathcal{S}_\vartheta$ as the mapping $D_{\vartheta_0} u(\vartheta_0) : \mathcal{C} \to V$,

$$\widetilde{u}(\delta\vartheta) := D_{\vartheta_0} u(\vartheta_0)(\delta\vartheta) := \lim_{s\to 0^+} \frac{1}{s}\big(u(\vartheta_0 + s\delta\vartheta) - u(\vartheta_0)\big), \quad \delta\vartheta \in \mathcal{C}.$$

We also introduce the derivative of $\mathcal{A}(\vartheta_0)$ with respect to $\vartheta_0 \in \mathcal{S}_\vartheta$

$$\widetilde{\mathcal{A}}(\delta\vartheta)\varphi := D_{\vartheta_0}\mathcal{A}(\vartheta_0)(\delta\vartheta)\varphi := \lim_{s\to 0^+} \frac{1}{s}\big(\mathcal{A}(\vartheta_0+s\delta\vartheta)\varphi - \mathcal{A}(\vartheta_0)\varphi\big), \quad \varphi \in V, \quad \delta\vartheta \in \mathcal{C}.$$

We assume that $\widetilde{\mathcal{A}}(\delta\vartheta) \in \mathcal{L}(\widetilde{V}, \widetilde{V}^*)$ with $\widetilde{V}$ being a real and separable Hilbert space satisfying

$$\widetilde{V} \subseteq V \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} V^* \subseteq \widetilde{V}^*.$$

We further assume that there exists a real and separable Hilbert space $\overline{V} \subseteq \widetilde{V}$ such that $\widetilde{\mathcal{A}}v \in V^*$, $\forall v \in \overline{V}$. We have the following relation between $D_{\vartheta_0} u(\vartheta_0)(\delta\vartheta)$ and u.

Lemma 5.1. *Let $\widetilde{\mathcal{A}}(\delta\vartheta) \in \mathcal{L}(\widetilde{V}, \widetilde{V}^*)$, $\forall \delta\vartheta \in \mathcal{C}$ and $u(\vartheta_0) : (0,T] \to \overline{V}$, $\vartheta_0 \in \mathcal{S}_\vartheta$ be the unique solution to*

$$\begin{aligned} \partial_t u(\vartheta_0) + \mathcal{A}(\vartheta_0)u(\vartheta_0) &= 0 \quad \textit{in } (0,T) \times \mathbb{R}^d, && (5.1)\\ u(\vartheta_0)(0,\cdot) &= g(x) \quad \textit{in } \mathbb{R}^d. && (5.2) \end{aligned}$$

Then $\widetilde{u}(\delta\vartheta)$ solves

$$\begin{aligned} \partial_t \widetilde{u}(\delta\vartheta) + \mathcal{A}(\vartheta_0)\widetilde{u}(\delta\vartheta) &= -\widetilde{\mathcal{A}}(\delta\vartheta)u(\vartheta_0) \quad \textit{in } (0,T) \times \mathbb{R}^d, && (5.3)\\ \widetilde{u}(\delta\vartheta)(0,\cdot) &= 0 \quad \textit{in } \mathbb{R}^d. && (5.4) \end{aligned}$$

Proof. Since $u(\vartheta_0)(0) = g$ does not depend on ϑ_0 its derivative with respect to ϑ is 0. Now let $\vartheta_s := \vartheta_0 + s\delta\vartheta$, $s > 0$, $\delta\vartheta \in \mathcal{C}$. Subtract from the equation $\partial_t u(\vartheta_s)(t) + \mathcal{A}(\vartheta_s)u(\vartheta_s)(t) = 0$ equation (5.1) and divide by s to obtain

$$\partial_t \frac{u(\vartheta_s)(t) - u(\vartheta_0)(t)}{s} + \frac{\big(\mathcal{A}(\vartheta_s) - \mathcal{A}(\vartheta_0)\big)u(\vartheta_s)(t)}{s} + \frac{\mathcal{A}(\vartheta_0)\big(u(\vartheta_s)(t) - u(\vartheta_0)(t)\big)}{s} = 0.$$

Taking $\lim_{s\to 0+}$ gives equation (5.3). ☐

The PDE for the sensitivity $\widetilde{u}(\delta\vartheta)$ can again be discretized as in Section 4.

5.2 Sensitivity with respect to solution arguments

We also want to calculate the sensitivity of the solution u with respect to a variation of arguments t, x. Let u be the solution of the variational problem (3.1). We discuss the computation of $\mathcal{D}^{\mathbf{n}} u = \frac{\partial^{|\mathbf{n}|}}{\partial_{x_1}^{n_1} \cdots \partial_{x_d}^{n_d}} u$ for arbitrary multiindex $\mathbf{n} \in \mathbb{N}_0^d$. For $\mu \in \mathbb{Z}^d$ and $h \in \mathbb{R}_+$ we define the translation operator $T_h^\mu \varphi(x) = \varphi(x+\mu h)$ and the forward difference quotient $\partial_{h,j}\varphi(x) = h^{-1}(T_h^{e_j}\varphi(x) - \varphi(x))$, where e_j, $j = 1, \dots, d$, denotes the j-th standard basis vector in $\mathbb{R}^d$. For $\mathbf{n} \in \mathbb{N}_0^d$ we denote by $\partial_h^{\mathbf{n}}\varphi = \partial_{h,1}^{n_1} \cdots \partial_{h,d}^{n_d}\varphi$ and by $\mathcal{D}_h^{\mathbf{n}}$ the difference operator of order $n \geq 0$

$$\mathcal{D}_h^{\mathbf{n}}\varphi := \sum_{\mu, |\mathbf{n}|=n} C_{\mu,\mathbf{n}} T_h^{\gamma} \partial_h^{\mathbf{n}}\varphi.$$

Given a basis $\mathcal{B} = \{\Phi_j\}_{j=1}^N$ of V_h, the action of $\mathcal{D}_h^{\mathbf{n}}$ to $v_h \in V_h$ can be realized as matrix-vector multiplication $\underline{v}_h \mapsto \mathbf{D}_h^{\mathbf{n}}\underline{v}_h$, where

$$\mathbf{D}_h^{\mathbf{n}} = \left(\mathcal{D}_h^{\mathbf{n}}\Phi_1, \cdots, \mathcal{D}_h^{\mathbf{n}}\Phi_N\right) \in \mathbb{R}^{N \times N},$$

and $\underline{v}_h$ is the coefficient vector of v_h with respect to basis $\mathcal{B}$, respectively. For more details and numerical examples we refer to [4].

6 Numerical examples

We give numerical examples using the wavelet finite element discretization as described in Example 4.3. We consider a geometric call option for the Black-Scholes model written on up to 8 underlyings and analyze a dimensionally reduced problem from 30 to 5 computational dimensions. We also study a model problem for the stochastic volatility model. All computations are written in FORTRAN and are performed on a 16 $\times$ Quad-Core AMD Opteron(tm) Processor 8356 with 64GB RAM.

6.1 Full-rank d-dimensional Black-Scholes model

We consider the geometric call option with payoff

$$g(x) = \max(0, e^{\sum_{i=1}^d \alpha_i x_i} - K), \quad x \in \mathbb{R}^d, \tag{6.1}$$

with $K = 1$. The antiderivative of $g_{\log}$ for $\alpha_i > 0$, $i = 1, \dots, d$ is given by

$$g^{(-2)}(x) = \prod_{i=1}^d \alpha_i^{-2} \left(e^{\sum_{i=1}^d \alpha_i x_i} - \sum_{k=0}^{2d} \frac{1}{k!} \left(\sum_{i=1}^d \alpha_i x_i - \log K \right)^k \right) \mathbb{1}_{\{\sum_{i=1}^d \alpha_i x_i \geq \log K\}}.$$

We first set $d = 2$ and solve problem (4.6) for various mesh widths $h = 2^{-L}$. Using interest rate $\gamma = 0.01$, covariance $\mathcal{Q} = (\sigma_i\sigma_j\rho_{ij})_{1 \leq i,j \leq 2}$, $\sigma_1 = 0.4$, $\sigma_2 = 0.1$, $\rho_{12} = 0.2$ and weights $\alpha_i = 0.5$, $i = 1, 2$, we plot the convergence rate of the L^2-error

$$\|e_L\| := \|u(T, \cdot) - u_L(T, \cdot)\|_{L^2(G_0)}, \quad G_0 = (K/2, 3/2K)^2,$$

at maturity $T = 1$ in Figure 2.

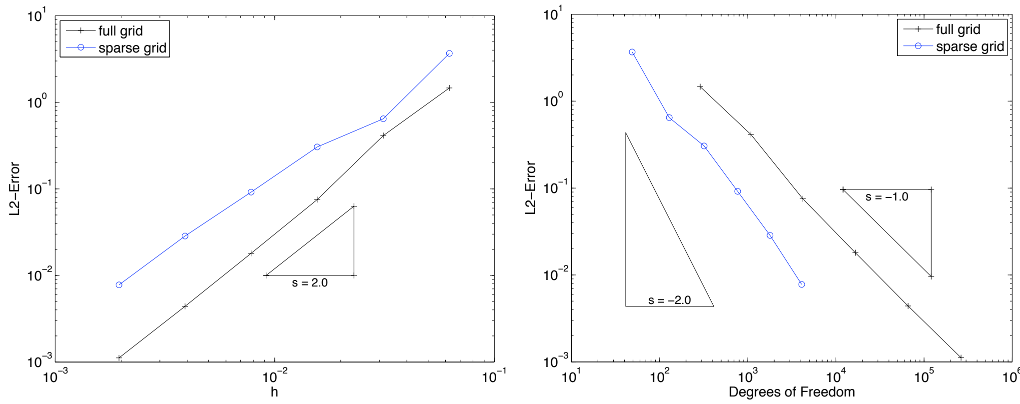

Figure 2: Convergence rates of the 2-dimensional wavelet discretization in terms of the mesh width h (left) and in terms of degrees of freedom (right).

To compare the rates we also solved the problem on full grid. In the left picture the convergence rate of the error on sparse grid can be seen to have (up to a constant) the same rate as on full grid. The superiority of sparse grid can be seen on the right, where the convergence rate is plotted terms of degrees of freedom. From [9], the error in terms of degrees of freedom on the sparse grid behaves like

$$\|e_L\| = \mathcal{O}\big(\widehat{N}_L^{-2}(\log \widehat{N}_L)^{c(d)}\big), \tag{6.2}$$

with $c(d)$ a constant, while on the full grid like

$$\|e_L\| = \mathcal{O}\big(N_L^{-\frac{2}{d}}\big). \tag{6.3}$$

On the sparse grid we have $\widehat{N}_L = \mathcal{O}(L\,2^L)$ and on the full grid $N_L = \mathcal{O}(2^{2L})$. The convergence rate on full grid shows the "curse of dimension", whereas for the sparse grid, we still obtain the optimal second order rate essentially.

For $2 \leq d \leq 8$, we set $\sigma_i = 0.3$, $i = 1, \ldots, d$ and $\rho_{i,j} = 0$, $i \neq j$, $i = 1, \ldots, d$, $j = i, \ldots, d$ and weights $\alpha_1 = 1$, $\alpha_i = 0$, $i = 2, \ldots, d$. The resulting payoff contract hence reduces to a plain vanilla call in the underlying x_1. Accordingly, we plot the convergence rates of the relative L^2-error

$$\|e_L\| := \frac{\|u(T,\cdot) - u_L(T,\cdot)\|_{L^2(G_1)}}{\|u(T,\cdot)\|_{L^2(G_1)}}, \quad G_1 = (K/2, 3/2K) \times \{K\} \times \ldots \times \{K\},$$

at maturity $T = 1$ for $d \in \{2, 3, 4, 6, 8\}$ in terms of $\widehat{N}_L$ in Figure 3.

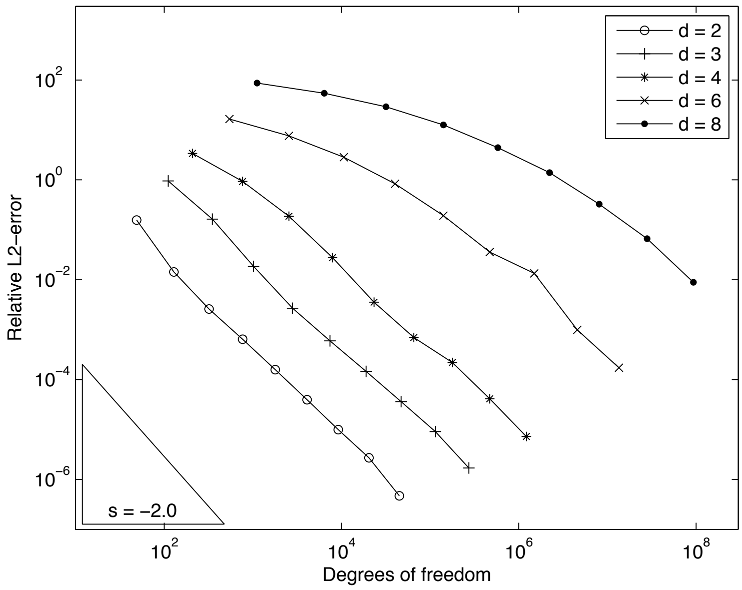

Figure 3: Convergence rates of the Black-Scholes model in different dimensions d.

For low dimensions $d \in \{2, 3, 4\}$, the second order convergence rate on sparse grid can be well observed over all levels. For higher dimensions $d \in \{6, 8\}$, the log terms prevail at low levels, hence the flattened behavior of the convergence curves which then exhibit the expected second order rates at finer discretizations. Despite the smoothness of the solution, we report a steeply increasing constant in the rates as d is raised, which forces us to already set $L = 11$ for $d = 8$ in order to have a relative L^2-error on the order of 10^{-2} which currently prevents us from reasonably increasing d beyond 8. The size of the constant can be traced back to the initial condition u_h^0, the H-projection of the payoff g onto V_h (Section 4.5), showing similar relative L^2-errors. An attempt in order to lower this constant and therefore to alleviate the need for high discretization levels would consist in reformulating problem (3.3) in *excess-to-payoff*[2] v_R by letting

$$v_R(t, z) := u_R(t, z) - g(z), \quad (t, z) \in J \times G_R$$

thus resulting in the following parabolic PDE with homogeneous initial condition

$$\begin{aligned} \partial_t v_R + \mathcal{A} v_R + \gamma v_R &= -\mathcal{A} g - \gamma g && \text{in} \quad J \times G_R, \\ v_R(t, \cdot) &= 0 && \text{on} \quad J \times \partial G_R, \\ v_R(0, z) &= 0 && \text{in} \quad G_R, \end{aligned} \tag{6.4}$$

which is numerically solved along the same lines. This alternative formulation however trades a non-smooth payoff for a non-smooth solution, consequently questioning the improvement of the operation. Under the same settings, we compare

[2] commonly referred to as *premium* in the financial literature.

the convergence rates of the previously defined relative L^2-error for the standard and excess-to-payoff formulations for $d \in \{2,4,6\}$ at maturity $T = 1$ in Figure 4. Except for $d = 2$, the error is reduced at lower levels, as expected from the homo-

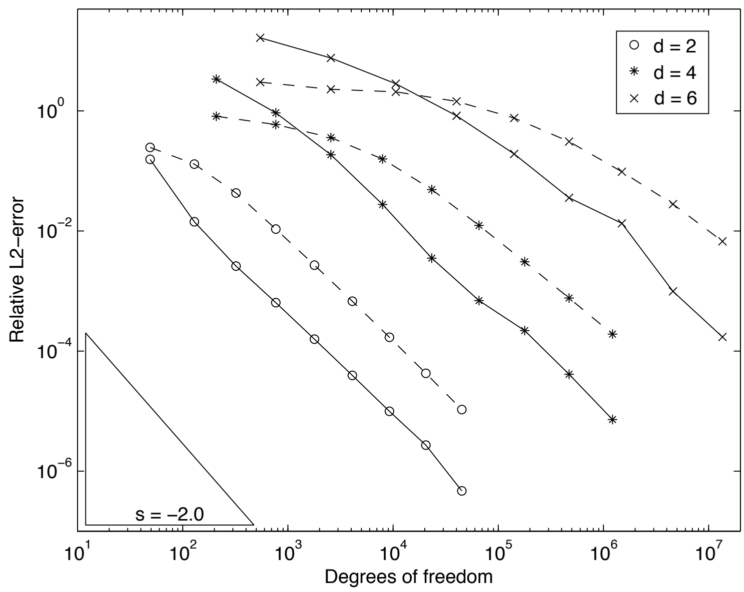

Figure 4: Convergence rates of the Black-Scholes model for $d \in \{2,4,6\}$ in standard (problem (3.3), solid line) and excess-to-payoff formulations (problem (6.4), dashed line).

geneous initial condition, but the optimal convergence rate (6.2) is only reached between higher levels than in the standard formulation case, therefore showing that a non-smooth initial condition is overall preferable. From a computational perspective, the excess-to-payoff formulation moreover gives rise to a right-hand side in (6.4) which requires additional memory requirements on the order of the degrees of freedom. Payoff smoothing thus appears a next sensible step in order to lower the constant resulting from the H-projection of the payoff g onto the approximation space V_h.

6.2 Low-rank r-dimensional Black-Scholes

We consider the geometric call option (6.1) of Section 6.1 written on d underlyings under the Black-Scholes model and we are now interested in the case of larger dimensions, i.e., $d > 8$. This occurs when pricing contingent claims on stock indices considering all d price processes in comparison to handling the index as one single process. Straightforward computations in such high dimensions would currently require too high discretization levels as previously noted. Instead, we rely on the dimensionality reduction by ϵ-aggregation introduced in Section 2.1.3 to identify a rank r ϵ-aggregated process driving a d-dimensional market. In particular, we focus on the Dow Jones industrial index where $d = 30$. We compute the principal components of the volatility covariance matrix[3] $\mathcal{Q} := \mathbf{U}^\top \mathbf{D} \mathbf{U} \in \mathbb{R}^{d\times d}$ of their realized daily log-returns over 252 periods resulting in the spectrum $(s_1^2, \ldots, s_d^2)$, normalized by s_1^2 as in Section 2.1 and shown in Figure 5 (left)[4]. We define the *recovery ratio*

$$\eta_r := \left(\sum_{i=1}^{r} s_i^2\right)\left(\sum_{i=1}^{d} s_i^2\right)^{-1}, \qquad r = 1, \ldots, d,$$

shown in Figure 5 (right), whose values for $r = 1, \ldots, 5$ are reported in Table 1. The eigenvalues are observed to decay exponentially and by virtue of Theorem 2.3, we may therefore expect sufficiently accurate results for $r \leq 5$.

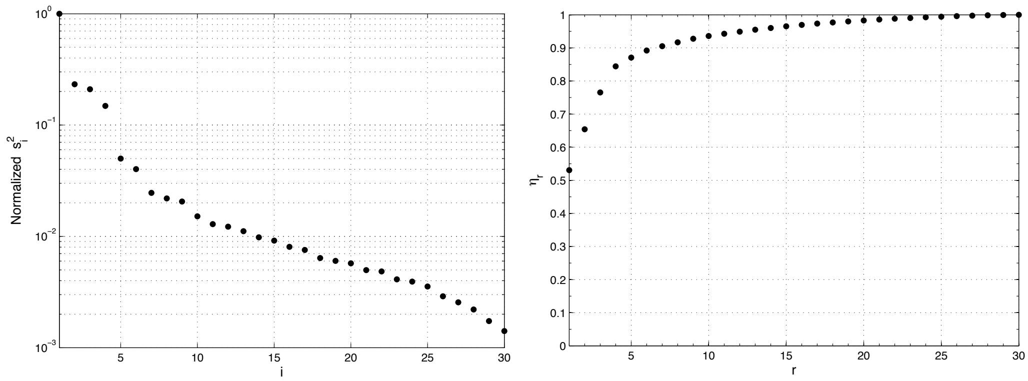

Figure 5: Eigenvalue spectrum (normalized by s_1^2) of the realized volatility covariance matrix of the constituents of the Dow Jones index over a period of 252 days (left) and recovery ratio η_r, $r = 1, \ldots, d$ (right).

r	s_r	s_r^2	η_r
1	0.2052	0.0421	0.5307
2	0.0990	0.0098	0.6542
3	0.0940	0.0088	0.7655
4	0.0791	0.0062	0.8443
5	0.0459	0.0021	0.8708

Table 1: First eigenvalues s_r^2, $r = 1, \ldots, 5$, of the Dow Jones realized volatility covariance matrix $\mathcal{Q}$ and recovery ratio η_r.

As in Section 2.1.2, let $\widehat{\mathbf{D}} := \mathrm{diag}(\hat{s}_1^2, \ldots, \hat{s}_d^2) \in \mathbb{R}^{d\times d}$ with $\hat{s}$ as in (2.8) for some

[3] $\mathcal{Q}$ was computed from the historical time series adjusted for dividends and interest rates of the daily log-returns of the 30 constituents of the Dow Jones industrial average index.

[4] Similar eigenvalue decompositions for the DAX index are presented in [8].

$2 \leq r \leq 5$ which defines the rank-reduced processes $\widehat{X}$ and $\widehat{Y}$ defined by (2.9) and (2.10) respectively. We approximate the exact solution $u(t,x) = v(t,y)$ of the d-dimensional problem (3.3) by a parametrized r-dimensional option price $\widehat{v}(t,\hat{y}) = \widehat{v}(t,\hat{y}_1,\dots,\hat{y}_r;\hat{y}_{r+1},\dots,\hat{y}_d)$ for $\hat{y} = (\hat{y}_1,\dots,\hat{y}_d) \in G := (-R,R)^d$. We therefore consider the option prices $\widehat{v}_R(t,\hat{y}_1,\dots,\hat{y}_r;\hat{y}_{r+1},\dots,\hat{y}_d)$ satisfying

$$\begin{array}{rcll} \partial_t \widehat{v}_R + \widehat{\mathcal{A}}\widehat{v}_R + \gamma \widehat{v}_R & = & 0 & \text{in} \quad J \times \widehat{G}^r_R, \\ \widehat{v}_R(t,\cdot) & = & 0 & \text{on} \quad J \times \partial \widehat{G}^r_R, \\ \widehat{v}_R(0,\hat{y}) & = & \widehat{f}(e^{\hat{y}})|_{\widehat{G}^r_R} & \text{in} \quad \widehat{G}^r_R, \end{array} \tag{6.5}$$

where $\widehat{G}^r_R := (-R,R)^r$. The rank-$r$ operator $\widehat{\mathcal{A}}$ is now given by

$$\widehat{\mathcal{A}} := -\frac{1}{2}\mathrm{tr}\big[\widehat{\mathbf{D}}\widehat{D^2}\big] - \langle \widehat{\lambda}, \widehat{D} \rangle,$$

with $\widehat{\lambda}$ as in (2.10), rank-r differential operators $\widehat{D}$, and $\widehat{D^2}$

$$\widehat{D} := (\partial_{\hat{y}_1},\dots,\partial_{\hat{y}_r},0,\dots,0)^\top_, \quad \widehat{D^2} := \left(\begin{array}{c|c} (\partial^2_{\hat{y}_i\hat{y}_j})_{1\leq i,j\leq r} & 0 \\ \hline 0 & 0 \end{array} \right),$$

respectively, and

$$\begin{array}{rcl} \widehat{f}(e^{\hat{y}}) & = & f(e^{\hat{y}_1},\dots,e^{\hat{y}_r},e^{\hat{y}_{r+1}+\widehat{\lambda}_{r+1}T},\dots,e^{\hat{y}_d+\widehat{\lambda}_d T}) \\ & = & \max(0, e^{\sum_{i=1}^r \alpha_i \hat{y}_i + \sum_{i=r+1}^d \hat{y}_i + \widehat{\lambda}_i T} - K). \end{array}$$

We numerically solve (6.5) for $r = 2,\dots,5$ and various mesh widths $h = 2^{-L}$ with the Dow Jones realized volatility covariance matrix $\mathcal{Q}$ whose principal components are plotted in Figure 5 (left), weights $\alpha_i = 0.3$, $i = 1,\dots,d$, maturity $T = 1$, strike $K = 1$ and interest rate $\gamma = 0.045$, and plot the convergence rate of the relative L^2-error

$$\|e_L\| := \frac{\|v(T,\cdot) - \widehat{v}_L(T,\cdot)\|_{L^2(G_1)}}{\|v(T,\cdot)\|_{L^2(G_1)}}, \quad G_1 = (3/4K, 5/4K) \times \{K\} \times \ldots \times \{K\},$$

at maturity $T = 1$ in Figure 6. The flattening behavior of the convergence rates is explained by further expanding the error into (a) an ϵ-aggregation error made by artificially setting volatilities to zero (Theorem 2.3) and (b) a discretization error [9] as

$$\begin{array}{rcl} \|v(t,y) - \widehat{v}_L(t,\hat{y})\| & = & \|v(t,y) - \widehat{v}(t,\hat{y}) + \widehat{v}(t,\hat{y}) - \widehat{v}_L(t,\hat{y})\| \\ & \leq & \underbrace{\|v(t,y) - \widehat{v}(t,\hat{y})\|}_{(a)} + \underbrace{\|\widehat{v}(t,\hat{y}) - \widehat{v}_L(t,\hat{y})\|}_{(b)}. \end{array}$$

It follows that the lower bounds observed in Figure 6 therefore stem from the ϵ-aggregation error which diminishes as r is increased.

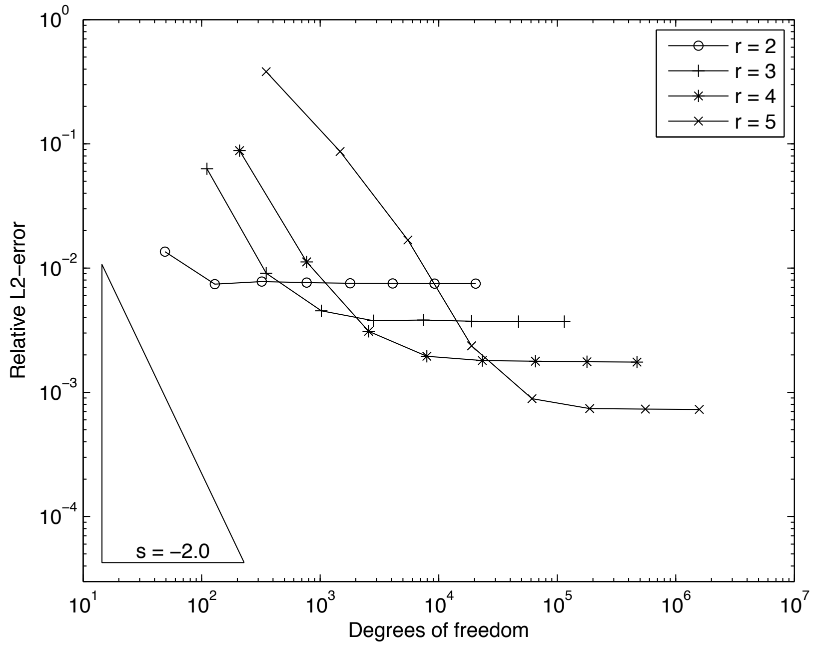

Figure 6: Convergence rates of the approximation of a 30 dimensional option price on the Dow Jones by $r = 2, \dots, 5$ dimensional options in the Black-Scholes model.

6.3 Stochastic volatility models

Since there are no analytically tractable solutions available for the price of a basket with stochastic volatility, we introduce a model problem for which the exact solution is known in closed form. For $z = (x, y) \in G := [0,1]^n \times [0,2]^n$, consider

$$\begin{array}{rcll} \partial_t u + \mathcal{A}u + \gamma u &=& f(t,z) & \text{in} \quad J \times G \\ u(t,\cdot) &=& 0 & \text{on} \quad J \times \partial G \ , \\ u(0,z) &=& g(z) & \text{in} \quad G \end{array}$$

where the operator $\mathcal{A}$ is as in (3.2), with μ, Σ as in (2.15)–(2.16). We let

$$f_{ij}(y) = \begin{cases} |y_i| + |y_{(i+1) \bmod n}|, & \text{if } i = j \\ 0, & \text{else} \end{cases} \quad , \quad 1 \le i, j \le n,$$

and consider functions $c_i(y_i) = \alpha_i(m_i - y_i)$, $b_i(y_i) = \beta_i$, $\Lambda_i = 0$, $1 \le i \le n$, which define the volatility processes as in Example 2.4. Furthermore, we let $\rho_{ji} = 0$ if $i \ne j$. In this setting, the operator $\mathcal{A}$ simplifies to

$$\mathcal{A} = -\frac{1}{2}\sum_{i=1}^{n} \left[f_{ii}^2(y)\partial_{x_i x_i} + \beta_i^2 \partial_{y_i y_i} + 2\beta_i \rho_{ii} f_{ii}(y) \partial_{x_i y_i} + \big(2\gamma - f_{ii}^2(y)\big)\partial_{x_i} + 2c_i(y_i)\partial_{y_i} \right].$$

We set the solution to $u(t,z) = u(t,x,y) := e^{-t}\prod_{i=1}^{n}\sin(\pi x_i)(y_i - 4^{-1}y_i^3)$ and use for $i = 1,\ldots,n$, the values

$$\alpha_i = \frac{1}{10}(1+4\frac{i}{n}),\ \beta_i = \frac{1}{2}(1+\frac{i}{n}),\ m_i = \frac{1}{20}(2-\frac{i}{n}),\ \rho_{ii} = -\frac{1}{10}(1+8\frac{i}{n}),$$

and interest rate $\gamma = 0.05$. We plot the error $\|e_L\| := \|u(T,z) - u_L(T,z)\|_{L^2(G)}$ at $T = 0.5$ against $\widehat{N}_L$ in Figure 7 to obtain the rate of convergence

$$\|e_L\| = \mathcal{O}\big(\widehat{N}_L^{-2}(\log\widehat{N}_L)^{c(n)}\big),$$

with a constant $c(n)$ depending linearly on the dimension of the problem.

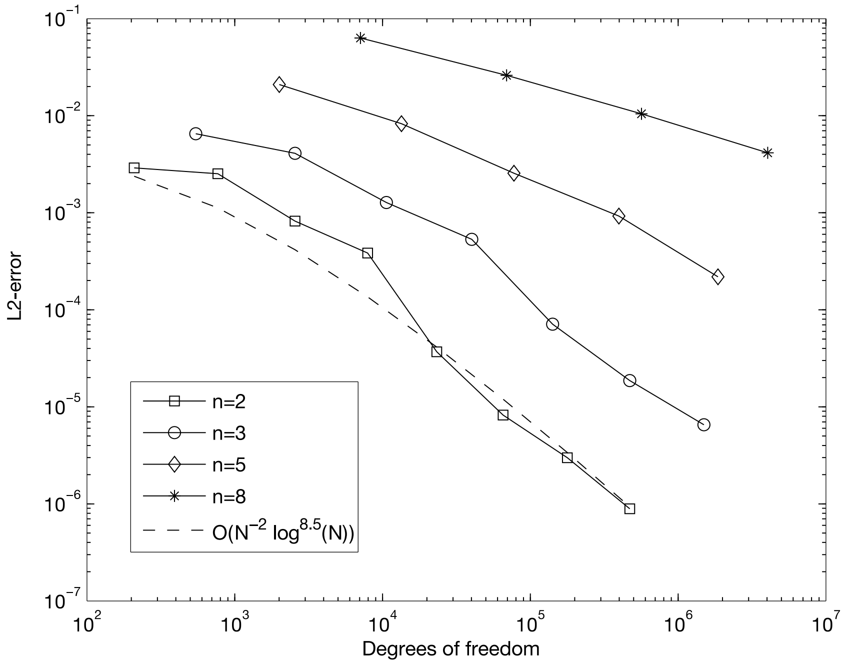

Figure 7: Convergence rates for the stochastic volatility model in different dimensions $d = 2n$.

7 Outlook: Rank-k dimensional corrections

We consider the price u of an option in a full-rank d-dimensional time rescaled Black-Scholes market with principal components $s_1^2 = 1 \geq \ldots \geq s_d^2$, $i = 1,\ldots,d$, of its volatility covariance matrix $\mathcal{Q} = \mathbf{\Sigma}\mathbf{\Sigma}^\top$, $\mathbf{\Sigma} = \mathbf{U}^\top\mathbf{D}^{\frac{1}{2}}$ with $\mathbf{D} = \mathrm{diag}(s_1^2,\ldots,s_d^2)$ and $\mathbf{U}$ as introduced in Section 2.1.1. Without loss of generality, the option value u in (2.1) is also a function of $\mathbf{\Sigma}$ such that we can write

$$u = u(t,x,s) = u(t,x,s_1,\ldots,s_d).$$

For some $1 \le r \le d$, we denote $\hat{s} := (s_1, \dots, s_r, 0, \dots, 0) \in \mathbb{R}^d$ as defined in (2.8), and we expand $u(t,x,s)$ in a Taylor series of order k about the low-rank option price $u(t,x,\hat{s})$

$$\begin{aligned} u(t,x,s) &= u(t,x,\hat{s}) + \sum_{i=r+1}^{d} s_i \frac{\partial u}{\partial s_i}(t,x,\hat{s}) + \frac{1}{2!} \sum_{i,j=r+1}^{d} s_i s_j \frac{\partial^2 u}{\partial s_i \partial s_j}(t,x,\hat{s}) \\ &+ \frac{1}{3!} \sum_{i,j,l=r+1}^{d} s_i s_j s_l \frac{\partial^3 u}{\partial s_i \partial s_j \partial s_l}(t,x,\hat{s}) + \ldots + \mathcal{O}(\|s-\hat{s}\|^{k+1}) \end{aligned} \quad (7.1)$$

The main premise of expansion (7.1) is that the partial derivatives need not be calculated exactly but only up to order $k+1$ which may be achieved with high order finite difference schemes [6]. In particular the case $r=1$, $k=1$ is treated in [8]. We illustrate hereafter the case $k=2$ for any $r \le d$, but extension to arbitrary order k is straightforward. Specifically, (7.1) becomes

$$\begin{aligned} u(t,x,s) &= u(t,x,\hat{s}) + \sum_{i=r+1}^{d} s_i \frac{\partial u}{\partial s_i}(t,x,\hat{s}) + \frac{1}{2} \sum_{i=r+1}^{d} s_i^2 \frac{\partial^2 u}{\partial^2 s_i}(t,x,\hat{s}) \\ &+ \frac{1}{2} \sum_{i=r+1}^{d} \sum_{\substack{j=r+1 \\ j \neq i}}^{d} s_i s_j \frac{\partial^2 u}{\partial s_i \partial s_j}(t,x,\hat{s}) + \mathcal{O}(\|s-\hat{s}\|^3), \end{aligned} \quad (7.2)$$

For $i,j = r+1, \dots, d$, we introduce the rank $r+1$ and $r+2$ vectors $\hat{s}^{(i)}_{(p)}$, $\hat{s}^{(i,j)}_{(p,q)} \in \mathbb{R}^d$ respectively, defined by

$$\left(\hat{s}^{(i)}_{(p)}\right)_m := \begin{cases} s_m & 1 \le m \le r, \\ \frac{1}{2}(p-1)s_i & m = i, \\ 0 & \text{else,} \end{cases} \qquad \left(\hat{s}^{(i,j)}_{(p,q)}\right)_m := \begin{cases} s_m & 1 \le m \le r, \\ \frac{1}{2}(p-1)s_i & m = i, \\ \frac{1}{2}(q-1)s_j & m = j, \\ 0 & \text{else,} \end{cases}$$

With weights[5] $\alpha = (-3, 4, -1)^\top$, $\beta = (4, -8, 4)^\top$, the partial derivatives in (7.2) are approximated by the second order finite differences

$$s_i \partial_{s_i} u(t,x,\hat{s}) = \sum_{p=1}^{3} \alpha_p u(t,x,\hat{s}^{(i)}_{(p)}) + \mathcal{O}(s_i^3), \quad (7.3)$$

$$s_i^2 \partial^2_{s_i} u(t,x,\hat{s}) = \sum_{p=1}^{3} \beta_p u(t,x,\hat{s}^{(i)}_{(p)}) + \mathcal{O}(s_i^3), \quad (7.4)$$

$$s_i s_j \partial^2_{s_i s_j} u(t,x,\hat{s}) = \sum_{p,q=1}^{3} \alpha_p \alpha_q u(t,x,\hat{s}^{(i,j)}_{(p,q)}) + \mathcal{O}(s_i^3 + s_j^3), \quad (7.5)$$

[5] Other choices of weights are possible implying other definitions for vectors $\hat{s}^{(i)}_{(p)}$, $\hat{s}^{(i,j)}_{(p,q)}$.

where $u(t,x,\hat{s}^{(i)}_{(p)})$ and $u(t,x,\hat{s}^{(i,j)}_{(p,q)})$ are *independent* solutions of $r+1$ and $r+2$ dimensional PDEs respectively as shown in Section 2.1.2. Inserting (7.3)–(7.5) into (7.2) and truncating the expansion yields the *rank*-2 corrected *rank*-r option price $u^{(r,2)}(t,x,s)$ approximation of $u(t,x,s)$

$$\begin{aligned} u^{(r,2)}(t,x,s) &:= u(t,x,\hat{s}) \\ &+ \sum_{i=r+1}^{d}\sum_{p=1}^{3}\Big(\alpha_p + \frac{\beta_p}{2}\Big)u(t,x,\hat{s}^{(i)}_{(p)}) \\ &+ \frac{1}{2}\sum_{i=r+1}^{d}\sum_{\substack{j=r+1\\ j\neq i}}^{d}\sum_{p,q=1}^{3}\alpha_p\alpha_q u(t,x,\hat{s}^{(i,j)}_{(p,q)}). \end{aligned} \tag{7.6}$$

Remark 7.1. (7.6) *is a linear combination of one r-dimensional, $2(d-r)$ $r+1$ dimensional and $4(d-r)(d-r+1)$ $r+2$ dimensional independent PDEs which can be solved independently in parallel using the numerical approach described in Section 4.*

Acknowledgments

The authors acknowledge Eckhard Platen from the University of Technology, Sydney, for constructive discussions in the preparation of this article.

References

[1] J.-P. Fouque, G. Papanicolaou, and K.R. Sircar. *Derivatives in financial markets with stochastic volatility.* Cambridge University Press, Cambridge, 2000.

[2] N. Hilber. *Stabilized wavelet methods for option pricing in high dimensional stochastic volatility models.* PhD thesis, ETH Zürich, 2009.

[3] N. Hilber, A.-M. Matache, and C. Schwab. Sparse wavelet methods for option pricing under stochastic volatility. *The Journal of Computational Finance*, 8(4):1–42, 2005.

[4] N. Hilber, C. Schwab, and C. Winter. Variational sensitivity analysis of parametric Markovian market models. In Ł. Stettner, editor, *Advances in Mathematics of Finance*, volume 83, pages 85–106. Banach Center Publ., 2008.

[5] C.C.W. Leentvaar and C.W. Oosterlee. On coordinate transformation and grid stretching for sparse grid pricing of basket options. *J. Comput. Appl. Math.*, 222(1):193–209, 2008.

[6] S.K. Lele. Compact finite difference schemes with spectral-like resolution. *J. Comput. Phys.*, 103, 1992.

[7] N. Reich, C. Schwab, and C. Winter. On Kolmogorov equations for anisotropic multivariate Lévy processes. Research Report 2008-03, Seminar for Applied Mathematics, ETH Zürich, 2008.

[8] C. Reisinger and G. Wittum. Efficient hierarchical approximation of high-dimensional option pricing problems. *SIAM J. Sci. Comput.*, 29(1), 2007.

[9] T. von Petersdorff and C. Schwab. Numerical solution of parabolic equations in high dimensions. *M2AN Math. Model. Numer. Anal.*, 38(1):93–127, 2004.

Research Reports

No.	Authors/Title
10-01	*N. Hilber, S. Kehtari, C. Schwab and C. Winter* Wavelet finite element method for option pricing in highdimensional diffusion market models
09-41	*E. Kokiopoulou, D. Kressner, N. Paragios, P. Frossard* Optimal image alignment with random projections of manifolds: algorithm and geometric analysis
09-40	*P. Benner, P. Ezzatti, D. Kressner, E.S. Quintana-Ortí, A. Remón* A mixed-precision algorithm for the solution of Lyapunov equations on hybrid CPU-GPU platforms
09-39	*V. Wheatley, P. Huguenot, H. Kumar* On the role of Riemann solvers in discontinuous Galerkin methods for magnetohydrodynamics
09-38	*E. Kokiopoulou, D. Kressner, N. Paragios, P. Frossard* Globally optimal volume registration using DC programming
09-37	*F.G. Fuchs, A.D. McMurray, S. Mishra, N.H. Risebrom, K. Waagan* Approximate Riemann solvers and stable high-order finite volume schemes for multi-dimensional ideal MHD
09-36	*Ph. LeFloch, S. Mishra* Kinetic functions in magnetohydrodynamics with resistivity and hall effects
09-35	*U.S. Fjordholm, S. Mishra* Vorticity preserving finite volume schemes for the shallow water equations
09-34	*S. Mishra, E. Tadmor* Potential based constraint preserving genuinely multi-dimensional schemes for systems of conservation laws
09-33	*S. Mishra, E. Tadmor* Constraint preserving schemes using potential-based fluxes. III. Genuinely multi-dimensional central schemes for for MHD equations
09-32	*S. Mishra, E. Tadmor* Constraint preserving schemes using potential-based fluxes. II. Genuinely multi-dimensional central schemes for systems of conservation laws
09-31	*S. Mishra, E. Tadmor* Constraint preserving schemes using potential-based fluxes. I. Multidimensional transport equations