A Sparse Grid Finite Difference method for pricing European options : application to a multi-factor stochastic variance swap model

David Pommier *
INRIA Paris-Rocquencourt, MathFi Project,
Domaine de Voluceau, Rocquencourt,
B.P. 105, 78153 Le Chesnay Cedex, FRANCE

March 3, 2020

Abstract

This paper presents a sparse grid finite difference method to compute an approximation of European option price. In the first part, we will give a short description of our model and the PDE formulation associated to the option pricing problem. Then, we will describe the sparse finite difference method use to solve the PDE. *Keywords: finite difference method, sparse grid, variance swap model.*

Premia 22

Introduction

We derive a numerical scheme to compute approximation price as solution of PDE.

In this paper, we are interested in the PDE formulation and discretization schemes for a multi-factor stochastic volatility model (Varswap-nd for short). We study the pricing problem for a European call option when the volatility of the underlying asset is random and follows the exponential of a sum of Ornstein-Uhlenbeck process. This is a generalization of the Scott model in the one factor case.

*Email: david.pommier@inria.fr

1 Model definition

The random diffusion model proposed is a four-dimensional market process that takes a log-Brownian motion to describe price dynamics and a sum of Ornstein-Uhlenbeck subordinated process describing the randomness of the log-volatility.

$$dS_t = (r - q)S_t dt + \sigma_t S_t dW_t. \tag{1}$$

We suppose $\sigma_t = f(Y_t)$, where f is a positive function and (Y_t) is the driven process. Some examples of driven process and f function are presented in [3],[1] for one factor model.

Definition 1.1 (multi-factor model)**.** *We consider a driven process (Y_t) defined in $\mathbb{R}^n$, so the stochastic volatility (σ_t) depend of a n-dimensional Ornstein-Uhlenbeck process $Y_t = \left(Y_t^1, \dots, Y_t^n\right)^T$. All one dimensional process $Y_t^i, 1 \le i \le n$ follows the stochastic differential equation:*

$$dY_t^i = \lambda_i(m_i - Y_t^i)dt + \beta_i dZ_t, \tag{2}$$

where λ_i and β_i are positive constants. We suppose that W_t and Z_t are two Brownian motion with a correlation factor equals to ρ.

In the following, we consider

$$f(Y_t) = \sigma_0 \exp\left(\frac{1}{2}\sum_{i=1}^{n} Y_t^i\right). \tag{3}$$

With this choice of volatility function, $m_i = 0$ seems to be natural.

Remark 1.2. *Note that n OU process are perfectly correlated. The SDE (2) are driven by only one Brownian motion.*

First, let us note that close formula methods have been not find for this model. This is due to our choice of f, the volatility function. So we propose a Monte Carlo method to deal with convergence of Sparse Grid method.

We will present here some technicals points in problem formulation.

To reduce complexity in implementation and subtract some terms in the PDE formulation, we solve problem on Forward Price in place of Spot Price. So we consider than the discounted price P_t of an European contract depend of $t, F_t, Y_t^1, \dots, Y_t^n$ where F_t is given by :

$$F_t^T = exp((r - q)(T - t))S_t.$$

To give same rules to each Y_t^i, we use a change of variable. With this, all of them have the same variance $(= 1)$ at time $T = \infty$. We define $\widehat{Y}_t^i = \frac{\sqrt{2\lambda_i}}{\beta_i} Y_t^i$, so (2) becomes

$$d\widehat{Y}_t^i = -\lambda_i \widehat{Y}_t^i dt + \sqrt{2\lambda_i} dZ_t, \tag{4}$$

and (3) becomes

$$f\left(\widehat{Y}_t\right) = \sigma_0 \exp\left(\frac{1}{2}\sum_{i=1}^{n} \frac{\beta_i}{\sqrt{2\lambda_i}} \widehat{Y}_t^i\right). \tag{5}$$

We keep the Y_t in place of $\widehat{Y}_t$ in the following.

2 Monte Carlo Method

We use a standard Euler scheme for the Monte-Carlo discretization. We apply standard reduction variance method. This method is consistent with the PDE formulation study in the next part. We will not compute price in our model, but the surprime u, define as the difference between the price in our model and the Black & Scholes price with good choice of volatility.

In other word, we define

$$\widetilde{F}_t^T = exp((r-q)(T-t))\widetilde{S}_t, \quad d\widetilde{F}_t^T = \sigma_0 \widetilde{F}_t^T dW_t.$$

The suprime is obtained by an approximation of:

$$u_t = \mathbb{E}\left[h\left(F_t^T\right) - h\left(\widetilde{F}_t^T\right) \middle| \mathcal{F}_t \right]. \tag{6}$$

To reduce variance on u_t, we consider only Put option in choice of h in (6). To define P_t for a call (resp put) option, we add Black & Scholes price of a call (resp put) at u_t.

We make choice to simulate $x_t = log\left(F_t^T\right)$

3 PDE formulation

With standard assumptions, we can show that

Proposition 3.1. *The price of European option is solution of the following backward PDE, with $x_t = log\left(F_t^T\right)$:*

$$\begin{aligned} \frac{\partial P}{\partial t} + \frac{1}{2} f(y)^2 \frac{\partial^2 P}{\partial x^2} - \frac{1}{2} f(y)^2 \frac{\partial P}{\partial x} + \rho \sum_{i=1}^{n} \beta_i f(y) \frac{\partial^2 P}{\partial x \partial y_i} & \\ + \frac{1}{2} \sum_{i,j=1}^{n} \beta_i \beta_j \frac{\partial^2 P}{\partial y_i \partial y_j} - \sum_{i=1}^{n} \lambda_i y_i \frac{\partial P}{\partial y_i} = & \quad 0, \quad (t,x,y) \in [0,T) \times \mathbb{R} \times \mathbb{R}^n, \\ P(x,y,T) = & \quad h(x), \quad (x,y) \in \mathbb{R} \times \mathbb{R}^n, \end{aligned} \tag{7}$$

where T is the maturity contract and h the payoff function.

We now apply same technique as used in Monte Carlo method.

Proposition 3.2. *The surprice u_t defined by (6) is solution of the following forward PDE:*

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{1}{2} f(y)^2 \frac{\partial^2 u}{\partial x^2} + \frac{1}{2} f(y)^2 \frac{\partial u}{\partial x} - \rho \sum_{i=1}^{n} \beta_i f(y) \frac{\partial^2 u}{\partial x \partial y_i} & \\ - \frac{1}{2} \sum_{i,j=1}^{n} \beta_i \beta_j \frac{\partial^2 u}{\partial y_i \partial y_j} + \sum_{i=1}^{n} \lambda_i y_i \frac{\partial u}{\partial y_i} = \frac{1}{2}\left(f(y)^2 - \sigma_0^2\right) Vega_{BS}(x,t), & \\ u(x,y,0) = 0, \quad (x,y) \in \mathbb{R} \times \mathbb{R}^n, & \end{aligned} \tag{8}$$

with the change of time $t \to T - t$.

4 Short introduction to sparse grid finite difference method.

We will give some notations in the first paragraph. After that, we will give some examples on elliptic problems and the discrete scheme to solve it with sparse grid methods.

Note that consistency results have been prove in [9], more reading prove have been proposed in [5].

4.1 Notations and Preliminary Results

Consider a boundary value problem in the hypercube $\Omega = (0,1)^d$. One can think of a Poisson problem $\Delta u = f$ with the Dirichlet boundary conditions $u = 0$ on $\partial\Omega$. Let $H^1(\Omega)$ be the Hilbert space of the square integrable functions whose first order partial derivatives are square integrable too. The norm in $H^1(\Omega)$ is $\|v\|_{H^1(\Omega)} = \sqrt{\|v\|^2_{L^2(\Omega)} + |v|^2_{H^1(\Omega)}}$ where $|v|^2_{H^1(\Omega)} = \sum_{i=1}^d \|\frac{\partial v}{\partial x_i}\|^2_{L^2(\Omega)}$. Let $H^1_0(\Omega)$ be the completion in $H^1(\Omega)$ of the subspace of smooth functions compactly supported in Ω. The last elliptic problem has a weak or variational formulation in $H^1_0(\Omega)$: find $u \in H^1_0(\Omega)$ such that $\int_\Omega \nabla u \cdot \nabla v = \int_\omega f v$, for all $v \in H^1_0(\Omega)$.
Assume that the solution of the Poisson problem is approximated by a conforming piecewise multi-linear finite element method on a Cartesian mesh, more precisely with piecewise linear functions of total degree $\leq d$. This is the lowest order finite element method on this mesh. Assume that the mesh is uniform and that each element is a cube of size n^{-1}. It is easy to see that the dimension of the approximation space is of the order of n^d: the algorithmic complexity grows exponentially with d, which actually forbids the use of this method for $d > 4$. This too rapid growth in complexity is known as the *curse of dimensionality*.
Yet, quite recent developments have shown that it may be possible to use deterministic Galerkin methods or grid based methods for elliptic or parabolic problems in dimension d, for $4 \leq d \leq 20$: these methods are based either on sparse grids [4] or on sparse tensor product approximation spaces [6].
In this paragraph, we aim at rapidly describing the principle of sparse approximations. This presentation heavily relies on the review article by H.J. Bungartz and M.Griebel [2]. We concentrate on the previously mentioned Dirichlet boundary value problem in Ω. The solution u will be approximated by a Galerkin method, *i.e.*, a variational problem posed in a finite dimensional approximation space V_n instead of $H^1_0(\Omega)$. The goal is to use approximation spaces V_n whose dimensions do not grow too rapidly with d.
The results below are proved in [2].

In this section, bold letters will stand for d-uples: for example, $\boldsymbol{x} = (x_1, \dots, x_d)$ and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d)$. We set $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^d$ and $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^d$. Take a sufficiently smooth function f defined on $[0,1]^d$; if $\boldsymbol{\alpha} \in \mathbb{N}^d$, we call $D^{\boldsymbol{\alpha}} f$ the partial derivative

$$D^{\boldsymbol{\alpha}} f = \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}},$$

where $|\boldsymbol{\alpha}| = \sum_{i=1}^{d} \alpha_i$. For two multi-index $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and a scalar λ, we define

$$\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = \sum_{i=1}^{d} \alpha_i \beta_i, \quad \lambda \boldsymbol{\alpha} = (\lambda \alpha_1, \ldots, \lambda \alpha_d), \quad \mathbf{2}^{\boldsymbol{\alpha}} = (2^{\alpha_1}, \ldots, 2^{\alpha_d}).$$

We say that $\boldsymbol{\alpha} \leq \boldsymbol{\beta}$ if $\alpha_i \leq \beta_i$, $i = 1, \ldots, d$, and that $\boldsymbol{\alpha} < \boldsymbol{\beta}$ if $\boldsymbol{\alpha} \leq \boldsymbol{\beta}$ and $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$.
Let us introduce the functional spaces $X^{q,r}(\Omega)$, for $r \in \mathbb{N}$ and $q \in [1, +\infty]$:

$$X^{q,r}(\Omega) = \{u \in L^q(\Omega),\ \forall \boldsymbol{\alpha} \text{ s.t. } \boldsymbol{\alpha} \leq r\mathbf{1},\ \ D^{\boldsymbol{\alpha}} u \in L^q(\Omega)\}, \tag{9}$$

which are endowed with the semi-norms:

$$|u|_{q,\boldsymbol{\alpha}} = \left(\int_\Omega |D^{\boldsymbol{\alpha}} u|^q \right)^{\frac{1}{q}}, \quad \boldsymbol{\alpha} \leq r\mathbf{1}, \text{ if q} < \infty, \tag{10}$$

$$|u|_{\infty,\boldsymbol{\alpha}} = \|D^{\boldsymbol{\alpha}} u\|_{L^\infty(\Omega)}, \quad \boldsymbol{\alpha} \leq r\mathbf{1}, \text{ if q} = \infty. \tag{11}$$

Note that $X^{q,r}(\Omega)$ is embedded in the more usual Sobolev space $W^{q,r}(\Omega) = \{u \in L^q(\Omega),\ \forall \boldsymbol{\alpha} \text{ s.t. } |\boldsymbol{\alpha}| \leq r,\ \ D^{\boldsymbol{\alpha}} u \in L^q(\Omega)\}$.
For a multi-index $\boldsymbol{\ell}$, consider the Cartesian meshes $\mathcal{T}_{\boldsymbol{\ell}}$ of $\overline{\Omega}$ with mesh steps $\boldsymbol{h}_{\boldsymbol{\ell}} = 2^{-\boldsymbol{\ell}} = (2^{-\ell_1}, \ldots, 2^{-\ell_d})$. The grid nodes of $\mathcal{T}_{\boldsymbol{\ell}}$ are the points $\boldsymbol{x}_{\boldsymbol{\imath}} = \boldsymbol{\imath} \cdot \boldsymbol{h}_{\boldsymbol{\ell}}$, $\mathbf{0} \leq \boldsymbol{\imath} \leq \mathbf{2}^{\boldsymbol{\ell}}$.
We note by ϕ the mother hat function:

$$\phi(x) = \begin{cases} 1 - |x| & \text{if } |x| < 1, \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

and $\phi_{\boldsymbol{\ell},\boldsymbol{\imath}}$ the d-dimensional hat function:

$$\phi_{\boldsymbol{\ell},\boldsymbol{\imath}}(\boldsymbol{x}) = \prod_{k=1}^{d} \phi(2^{\ell_k} x_k - i_k). \tag{12}$$

We call $V_{\boldsymbol{\ell}}$

$$V_{\boldsymbol{\ell}} = \operatorname{span}\left(\phi_{\boldsymbol{\ell},\boldsymbol{\imath}},\ \mathbf{1} \leq \boldsymbol{\imath} \leq \mathbf{2}^{\boldsymbol{\ell}} - \mathbf{1}\right) \tag{13}$$

We also consider the wavelet subspaces:

$$W_{\boldsymbol{k}} = \operatorname{span}\left\{\phi_{\boldsymbol{k},\boldsymbol{\imath}}, \mathbf{1} \leq \boldsymbol{\imath} \leq \mathbf{2}^{\boldsymbol{k}} - \mathbf{1}, i_j \text{ odd }, 1 \leq j \leq d\right\}. \tag{14}$$

We have

$$V_{\boldsymbol{\ell}} = \bigoplus_{\mathbf{1} \leq \boldsymbol{k} \leq \boldsymbol{\ell}} W_{\boldsymbol{k}}.$$

The basis of $V_{\boldsymbol{\ell}}$ obtained by assembling the previously mentioned bases of $W_{\boldsymbol{k}}$ $\mathbf{1} \leq \boldsymbol{k} \leq \boldsymbol{\ell}$ is called the hierarchical basis of $V_{\boldsymbol{\ell}}$. Calling $\boldsymbol{I}_{\boldsymbol{\ell}} = \{\boldsymbol{\imath} \leq \mathbf{2}^{\boldsymbol{\ell}} - \mathbf{1}\ :\ i_j \text{ odd }, 1 \leq j \leq d\}$, the hierarchical basis of $V_{\boldsymbol{\ell}}$ is $\{\phi_{\boldsymbol{k},\boldsymbol{\imath}},\ \boldsymbol{\imath} \in \boldsymbol{I}_{\boldsymbol{k}},\ \boldsymbol{k} \leq \boldsymbol{\ell}\}$. Note that the completion of $\bigoplus_{\mathbf{1} \leq \boldsymbol{k}} W_{\boldsymbol{k}}$ with respect to the $H^1(\Omega)$ norm is exactly $H_0^1(\Omega)$.
Rescaling the $\phi_{\boldsymbol{k},\boldsymbol{\imath}}$ as follows

$$\psi_{\boldsymbol{k},\boldsymbol{\imath}} = -2^{-(\boldsymbol{k}+\mathbf{1})\cdot\mathbf{1}} \phi_{\boldsymbol{k},\boldsymbol{\imath}}, \quad \boldsymbol{\imath} \in \boldsymbol{I}_{\boldsymbol{k}}, \tag{15}$$

we obtain another basis of $W_{\boldsymbol{k}}$.
If a function u is smooth enough, then the coefficients of its expansion in the hierarchical basis are obtained by a simple integral formula:

Lemma 4.1. *If $u \in H_0^1(\Omega) \cap X^{1,2}(\Omega)$, then*

$$u = \sum_{\boldsymbol{k} \geq \boldsymbol{1}} \sum_{\boldsymbol{\imath} \in \boldsymbol{I_k}} u_{\boldsymbol{k},\boldsymbol{\imath}} \phi_{\boldsymbol{k},\boldsymbol{\imath}}, \quad \text{where } u_{\boldsymbol{k},\boldsymbol{\imath}} = \int_\Omega D^{\boldsymbol{2}} u \cdot \psi_{\boldsymbol{k},\boldsymbol{\imath}}. \tag{16}$$

By using Lemma 4.1, one may evaluate the contribution $u_{\boldsymbol{k}}$ of a subspace $W_{\boldsymbol{k}}$ to the hierarchical expansion of u:

Lemma 4.2. *If $u \in H_0^1(\Omega) \cap X^{2,2}(\Omega)$, then the component $u_{\boldsymbol{k}} \in W_{\boldsymbol{k}}$ of the expansion of u in the hierarchical representation is such that*

$$\|u_{\boldsymbol{k}}\|_{L^2(\Omega)} \leq 2^{-2|\boldsymbol{k}|} 3^{-d} |u|_{2,\boldsymbol{2}}, \tag{17}$$

$$|u_{\boldsymbol{k}}|_{H^1(\Omega)} \leq 2^{-2|\boldsymbol{k}|} 3^{-d+\frac{1}{2}} \left(\sum_{j=1}^{d} 2^{2k_j} \right)^{\frac{1}{2}} |u|_{2,\boldsymbol{2}}. \tag{18}$$

4.2 Sparse Methods, approximation results

It is clear that the dimension of $V_{\boldsymbol{\ell}}$ is $\prod_{j=1}^{d} (2^{\ell_j} - 1)$. In particular, $\dim(V_{n\boldsymbol{1}}) = (2^n - 1)^d$. As already mentioned, the full tensor product space $V_{n\boldsymbol{1}}$ is often too large for practical use when $d > 4$.
Let us give an example of a sparse Galerkin method: the discrete space is chosen to be

$$V_n = \bigoplus_{\boldsymbol{1} \leq \boldsymbol{k}, |\boldsymbol{k}| \leq n+d-1} W_{\boldsymbol{k}} \tag{19}$$

instead of the full tensor product space $V_{n\boldsymbol{1}} = \bigoplus_{\boldsymbol{1} \leq \boldsymbol{k} \leq n\boldsymbol{1}} W_{\boldsymbol{k}}$. One may prove that

$$\dim(V_n) = 2^n \left(\frac{n^{d-1}}{(d-1)!} + O(n^{d-2}) \right). \tag{20}$$

Therefore $\dim(V_n)$ is significantly smaller than $\dim(V_{n\boldsymbol{1}})$. It can be seen that a Galerkin method with V_n is feasible for d of the order of 10. On Figure 1, we display the bases of $V_{n\boldsymbol{1}}$ and V_n.
Consider the discretization of the Dirichlet problem in Ω: the discretization error of the Galerkin method with the approximation space V_n (resp. $V_{n\boldsymbol{1}}$) is of the same order as the best fit error when approximating the solution of the continuous problem by a function of V_n (resp. $V_{n\boldsymbol{1}}$). Let us assume that u is smooth. We know that $\inf_{v \in V_{n\boldsymbol{1}}} \|v - u\|_{H^1(\Omega)} \leq C 2^{-n} |u|_{W^{2,2}(\Omega)}$, where $|u|^2_{W^{2,2}(\Omega)} = \sum_{|\boldsymbol{\alpha}|=2} \|D^{\boldsymbol{\alpha}} u\|^2_{L^2(\Omega)}$. Since V_n is much smaller than $V_{n\boldsymbol{1}}$, a similar estimate is not true for $\inf_{v \in V_n} \|v - u\|_{H^1(\Omega)}$. Griebel et al have proved the following theorem:

Theorem 4.3. *If $u \in H_0^1(\Omega) \cap X^{2,2}(\Omega)$, and if $u_n \in V_n$ is the component of the expansion of u in the hierarchical representation,*

$$\|u - u_n\|_{L^2(\Omega)} \leq \left(\frac{2^{-2n+1}}{12^d} \sum_{k=0}^{d-1} \binom{n+d-1}{k} \right) |u|_{2,\boldsymbol{2}} = O(2^{-2n} n^{d-1}) |u|_{2,\boldsymbol{2}}, \tag{21}$$

$$|u - u_n|_{H^1(\Omega)} \leq \left(\frac{2^{-n} d}{\sqrt{3}\, 6^{d-1}} \sum_{k=0}^{d-1} \right) |u|_{2,\boldsymbol{2}} = O(2^{-n}) |u|_{2,\boldsymbol{2}}. \tag{22}$$

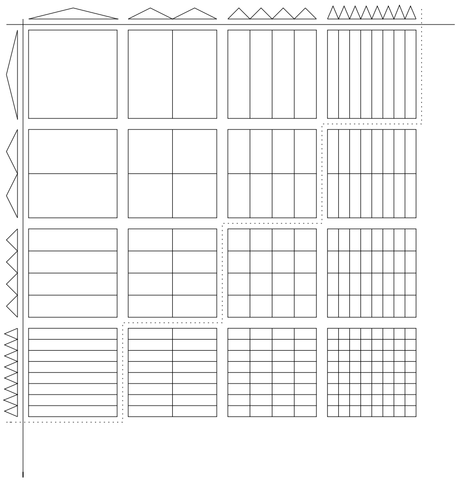

Figure 1: The case $d = 2$: each entry of this array corresponds to a pair of integer $\boldsymbol{k} = (k_1, k_2)$, $1 \leq k_1, k_2 \leq 4$, and contains the grid corresponding to $W_{\boldsymbol{k}}$. Each space $W_{\boldsymbol{k}}$ is the tensor product of two spaces whose bases are plotted on the sides of the array. The full tensor space $V_{n\mathbf{1}}$ is given by $V_{n\mathbf{1}} = \bigoplus_{\mathbf{1} \leq \boldsymbol{k} \leq n\mathbf{1}} W_{\boldsymbol{k}}$ whereas the sparse tensor space V_n is given by $V_n = \bigoplus_{\mathbf{1} \leq \boldsymbol{k}, |\boldsymbol{k}| \leq n+d-1} W_{\boldsymbol{k}}$, (only the spaces $W_{\boldsymbol{k}}$ corresponding to the entries above the diagonal are used to construct V_n)

Theorem 4.3 says that under the assumption that $u \in H_0^1(\Omega) \cap X^{2,2}(\Omega)$ (which is a rather strong regularity assumption, much stronger than the assumption $u \in H_0^1(\Omega) \cap W^{2,2}(\Omega)$ required when the full tensor product space is used), then using the sparse approximation space V_n instead of the full tensor space $V_{n\mathbf{1}}$ does not deteriorate the accuracy, at least with respect to the H^1 semi-norm. There is a moderate deterioration for the L^2 norm of the error.

In our presentation, we have focused on sparse methods based on tensorizing one dimensional hierarchical bases made of hat functions. This technique can be generalized to other classes of bases functions, for example higher order piecewise polynomial functions or wavelets as in Figure 2.

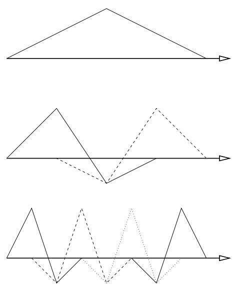

Figure 2: An example of wavelets

4.3 Sparse Grids Finite Difference Method

Before defining finite difference methods on sparse grids, we need to introduce new notations and concepts.

Consider the one variable shape functions: $\phi_{\ell,i}(x) = \phi(2^\ell x - i)$, $\ell \geq 1$, $1 \leq i \leq 2^\ell - 1$, and call V_ℓ the space spanned by $(\phi_{\ell,i})_{1\leq i\leq 2^\ell-1}$. Call W_ℓ the subspace of V_ℓ spanned by $(\phi_{\ell,2i-1})_{1\leq i\leq 2^{\ell-1}}$. We have $V_\ell = W_\ell \oplus V_{\ell-1}$. We have already seen that $V_1 \subset \dots V_\ell \subset V_{\ell+1} \subset \dots$ is a multiresolution analysis of $H_0^1((0,1))$. For a function $u \in \mathcal{C}^0([0,1])$ s.t. $u(0) = u(1) = 0$, we have

$$u = \sum_{\ell=1}^{\infty} \sum_{i=1}^{2^{\ell-1}} u_{\ell,i}\phi_{\ell,2i-1},$$

and the projection of u on V_ℓ is

$$\sum_{i=1}^{2^\ell-1} u^{\ell,i}\phi_{\ell,i} = \sum_{k=1}^{\ell} \sum_{i=1}^{2^{k-1}} u_{k,i}\phi_{k,2i-1}.$$

The change of coordinates $(u^{\ell,i})_{i=1,\dots,2^\ell-1} \mapsto (u_{k,i})_{k=1,\dots\ell,i=1,\dots,2^{k-1}}$ is called T_ℓ, which define transformation from nodal base to wavelet/hierarchical base. We call U_ℓ and U^ℓ the column vectors: $U^\ell = (u^{\ell,1}, \dots, u^{\ell,2^\ell-1}) \in \mathbb{R}^{2^\ell-1}$ and $U_\ell = (u_{\ell,1}, \dots, u_{\ell,2^{\ell-1}}) \in \mathbb{R}^{2^{\ell-1}}$. We have

$$T_\ell U^\ell = \begin{pmatrix} U_1 \\ \vdots \\ U_\ell \end{pmatrix}.$$

We denote by P^ℓ the restriction operator

$$P^\ell : \mathcal{C}^0([0,1]) \to \mathbb{R}^{2^\ell-1}, \quad P^\ell u = U^\ell. \tag{23}$$

Note that T_ℓ^{-1} is the representation of the operator P^ℓ in the wavelet basis, i.e.,

$$P^\ell\left(\sum_{k\leq\ell}\sum_{i=1}^{2^{k-1}} u_{k,i}\phi_{k,2i+1}\right) = T_\ell^{-1}\begin{pmatrix} U_1 \\ \vdots \\ U_\ell \end{pmatrix}.$$

We introduce the interpolation operator I^ℓ:

$$I^\ell : \mathbb{R}^{2^\ell-1} \to \mathcal{C}^0([0,1]), \quad I^\ell U = \sum_{i=1}^{2^\ell-1} u^i\phi_{\ell,i}. \tag{24}$$

We also denote by D^ℓ the finite difference operator for the discretization of $\frac{d^2}{dx^2}$:

$$D^\ell : \mathbb{R}^{2^\ell-1} \to \mathbb{R}^{2^\ell-1}, \tag{25}$$

$$\forall U, V \in \mathbb{R}^{2^\ell-1} \quad (D^\ell U, V) = 2^\ell \int_0^1 (I^\ell U)'(I^\ell V)'. \tag{26}$$

We consider the uniform grids of $(0,1)$: $\omega^\ell = 2^{-\ell}\{1,\dots,2^\ell-1\}$. For $\boldsymbol{\ell}\in\mathbb{N}^d$, $\mathbf{1}\le\boldsymbol{\ell}$, we introduce the Cartesian grid of Ω: $\Omega^{\boldsymbol{\ell}}=\prod_{i=1}^d\omega^{\ell_i}$. A grid function on $\Omega^{\boldsymbol{\ell}}$ is a mapping from $\Omega^{\boldsymbol{\ell}}$ to $\mathbb{R}$. The space of the grid functions on $\Omega^{\boldsymbol{\ell}}$ is exactly $\prod_{i=1}^d\mathbb{R}^{2^{\ell_i}-1}$. The mapping $(u^{\boldsymbol{\imath}})_{\mathbf{1}\le\boldsymbol{\imath}\le\mathbf{2}^{\boldsymbol{\ell}}-\mathbf{1}}\mapsto u=\sum_{\mathbf{1}\le\boldsymbol{\imath}\le\mathbf{2}^{\boldsymbol{\ell}}-\mathbf{1}}u^{\boldsymbol{\imath}}\phi_{\boldsymbol{\ell},\boldsymbol{\imath}}$ is an isomorphism from the space of the grid functions on $\Omega^{\boldsymbol{\ell}}$ onto $V_{\boldsymbol{\ell}}$ defined in (13). Moreover, the function u can be written on the wavelet basis $u=\sum_{\mathbf{1}\le\boldsymbol{k}\le\boldsymbol{\ell}}\sum_{\boldsymbol{\imath}\in\boldsymbol{I}_k}u_{\boldsymbol{k},\boldsymbol{\imath}}\phi_{\boldsymbol{k},\boldsymbol{\imath}}$. Calling $U_{\boldsymbol{k}}$ the vector $(u_{\boldsymbol{k},\boldsymbol{\imath}})_{\boldsymbol{\imath}\in\boldsymbol{I}_k}$, the grid function will be represented by the family $(U_{\boldsymbol{k}})_{\mathbf{1}\le\boldsymbol{k}\le\boldsymbol{\ell}}$.
For a positive integer n, we define the sparse grid Ω^n as follows:

$$\Omega^n=\cup_{\mathbf{1}\le\boldsymbol{\ell},|\boldsymbol{\ell}|\le n+d-1}\Omega^{\boldsymbol{\ell}}\subset\Omega^{n\mathbf{1}}. \tag{27}$$

An example of a sparse grid in dimension $d=2$ is presented in Figure 3.

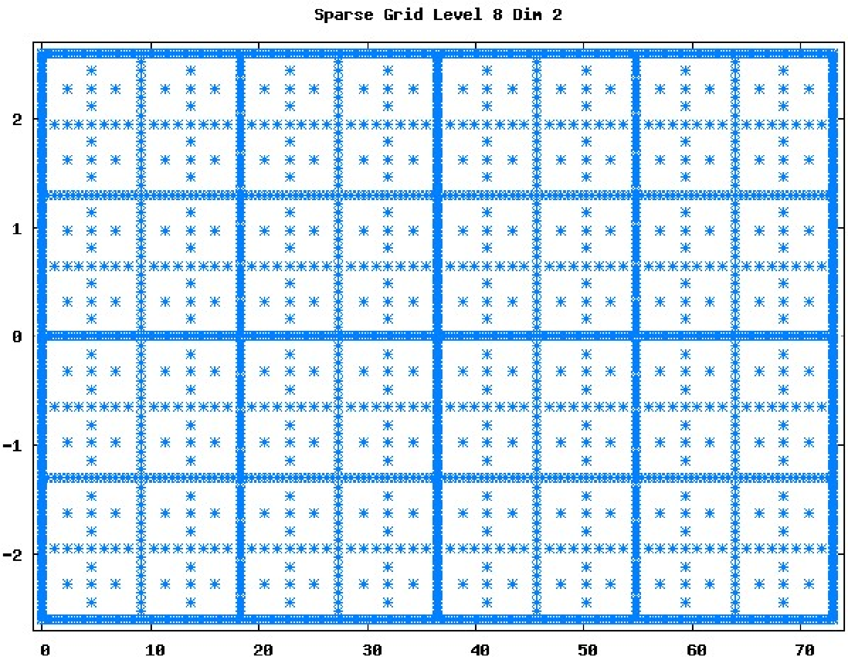

Figure 3: An example of a sparse grid for $d=2$, $n+1=8$

A grid function on Ω^n is a mapping from Ω^n to $\mathbb{R}$. The space of the grid functions on Ω^n is isomorphic to V_n defined in (19). As for the full tensor grid, a grid function on Ω^n can be represented on the wavelet basis by $\sum_{\mathbf{1}\le\boldsymbol{k},|\boldsymbol{k}|\le n}\sum_{\boldsymbol{\imath}\in\boldsymbol{I}_k}u_{\boldsymbol{k},\boldsymbol{\imath}}\phi_{\boldsymbol{k},\boldsymbol{\imath}}$. Calling $U_{\boldsymbol{k}}$ the vector $(u_{\boldsymbol{k},\boldsymbol{\imath}})_{\boldsymbol{\imath}\in\boldsymbol{I}_k}$, the sparse grid function will be represented by the family $(U_{\boldsymbol{k}})_{\mathbf{1}\le\boldsymbol{k},|\boldsymbol{k}|\le n+d-1}$.
We now define the sparse finite difference discretization of $\frac{\partial^2}{\partial x_1^2}$: given the vectors $\check{\boldsymbol{k}}=(k_2,\dots,\dots k_d)\in\mathbb{N}^{d-1}$, $\check{\boldsymbol{\imath}}\in I_{\check{\boldsymbol{k}}}$ and a sparse grid function represented by $(U_{\boldsymbol{k}})_{\mathbf{1}\le\boldsymbol{k},|\boldsymbol{k}|\le n+d-1}$, let $\tilde{k}$ be the positive integer $\tilde{k}=n+d-1-|\check{\boldsymbol{k}}|$; we introduce $U_{\check{\boldsymbol{k}}}$ by

$$U_{\check{\boldsymbol{k}},\check{\boldsymbol{\imath}}}=\begin{pmatrix}U_{(1,\check{\boldsymbol{k}},\check{\boldsymbol{\imath}})}\\ \vdots\\ U_{(\tilde{k},\check{\boldsymbol{k}},\check{\boldsymbol{\imath}})}\end{pmatrix}\qquad\text{where } U_{(j,\check{\boldsymbol{k}},\check{\boldsymbol{\imath}})}=\left(u_{(j,\check{\boldsymbol{k}}),(m,\check{\boldsymbol{\imath}})}\right)^T_{\{m\text{ odd},\,1\le\mathrm{m}\le 2^{\mathrm{j}}-1\}}.$$

Proposition 4.4. *The sparse grid discretization of the operator* $\frac{\partial^2}{\partial x_1^2}$ *is*

$$(U_{\boldsymbol{k}})_{\mathbf{1}\leq \boldsymbol{k},|\boldsymbol{k}|\leq n+d-1} \mapsto (V_{\boldsymbol{k}})_{\mathbf{1}\leq \boldsymbol{k},|\boldsymbol{k}|\leq n+d-1} \tag{28}$$

$$\textit{such that} \quad V_{\check{\boldsymbol{k}},\check{\boldsymbol{\imath}}} = T_{\tilde{k}} D^{\tilde{k}} T_{\tilde{k}}^{-1} U_{\check{\boldsymbol{k}},\check{\boldsymbol{\imath}}}, \quad \forall \boldsymbol{k}, \boldsymbol{\imath} \in I_{\boldsymbol{k}}. \tag{29}$$

The sparse grid discretization of the operators $\frac{\partial^2}{\partial x_j^2}$, $\frac{\partial}{\partial x_j}$, $j = 1, \ldots, d$, can be done in a similar way.

Classical finite difference operator on $(0,1)$ Let we call the linear finite difference schemes. They should be compute as a matrix vector product which include usual scheme of order M, but exclude ENO schemes. All this operator should be use in place of $D^{\tilde{k}}$ in (28).

- For operator $\dfrac{\partial}{\partial x}$:
  - Centered scheme of order $M = 2$:
$$\frac{\partial u}{\partial x}(x = x_i) \approx \left(D_\ell^1 u\right)_i = \frac{u\left(x_i + 2^{-\ell}\right) - u\left(x_i - 2^{-\ell}\right)}{2 \cdot 2^{-\ell}}. \tag{30}$$
  - Left decentered scheme of order $M = 1$:
$$\frac{\partial u}{\partial x}(x = x_i) \approx \left(D_\ell^- u\right)_i = \frac{u(x_i) - u\left(x_i - 2^{-\ell}\right)}{2^{-\ell}}. \tag{31}$$
  - Right decentered scheme of order $M = 1$:
$$\frac{\partial u}{\partial x}(x = x_i) \approx \left(D_\ell^+ u\right)_i = \frac{u\left(x_i + 2^{-\ell}\right) - u(x_i)}{2^{-\ell}}. \tag{32}$$
- for second order operator, centered scheme of order $M = 2$:
$$\frac{\partial^2 u}{\partial x^2}(x = x_i) \approx \left(D_\ell^2 u\right)_i = \frac{u\left(x_i + 2^{-\ell}\right) - 2u(x_i) + u\left(x_i - 2^{-\ell}\right)}{2^{-2\ell}}. \tag{33}$$
- Product by tensor product coefficient : suppose that $c(\boldsymbol{x}) = c_1(x_1)\ldots c_d(x_d)$,
$$\prod_{i=1}^{d} c^i \cdot = \mathrm{T}_{(d)} \circ \mathbf{c}_d \circ \mathrm{T}_{(d)}^{-1} \circ \cdots \circ \mathrm{T}_{(1)} \circ \mathbf{c}_1 \circ \mathrm{T}_{(1)}^{-1}. \tag{34}$$

Remark 4.5. *As see in (34) standard multiplication is could be very expensive in computational time. In fact, we have to do the product on the nodal base, so we have to apply transformation from nodal to hierarchic and the inverse in each direction which c is dependent.*

4.4 Consistency result

It is natural to define the restriction operator $P^{\boldsymbol{\ell}}: u \mapsto u|_{\Omega^{\boldsymbol{\ell}}}$ and the interpolation operator $I^{\boldsymbol{\ell}} = I^{\ell_1} \otimes \cdots \otimes I^{\ell_d} : \prod_{i=1}^{d} \mathbb{R}^{2^{\ell_i}-1} \to \mathcal{C}^0(\Omega)$. The finite difference approximation of $\partial^2_{x_1} u$ on the grid $\Omega^{\boldsymbol{\ell}}$ is $(I^{\boldsymbol{\ell}} \circ (D^{\ell} \otimes Id) \circ P^{\boldsymbol{\ell}})(u)$. It has been proved by Koster, see [5], that the sparse grid approximation of $\partial^2_{x_1} u$ can be written in terms of these finite difference operators:

Theorem 4.6. *For a function $u \in \mathcal{C}^0(\Omega)$ s.t. $u = 0$ on $\partial\Omega$, we note $D_n(u)$ the function of V_n whose expansion in the wavelet basis is given by $(V_{\boldsymbol{k}})_{\boldsymbol{1} \le \boldsymbol{k}, |\boldsymbol{k}| \le n+d-1}$ in (28), where $(U_{\boldsymbol{k}})_{\boldsymbol{1} \le \boldsymbol{k}, |\boldsymbol{k}| \le n+d-1}$ is the expansion on the wavelet basis of the projection of u on V_n. Then*

$$D_n(u) = \left(\sum_{\boldsymbol{1} \le \boldsymbol{k}, |\boldsymbol{k}| \le n+d-1} f(\boldsymbol{k}) I^{\boldsymbol{k}} \circ (D^{k_1} \otimes Id) \circ P^{\boldsymbol{k}} \right)(u), \tag{35}$$

where $f(\boldsymbol{k})$ is recursively defined by

$$\begin{array}{rcll} f(\boldsymbol{k}) & = & 0, & \text{if } |\boldsymbol{k}| > n+d-1 \text{ or } \boldsymbol{k} < \boldsymbol{1}, \\ f(\boldsymbol{k}) & = & 1 - \displaystyle\sum_{\boldsymbol{\ell}: \boldsymbol{k} < \boldsymbol{\ell}} f(\boldsymbol{\ell}), & \text{if } |\boldsymbol{k}| \le n+d-1 \text{ and } \boldsymbol{k} \ge \boldsymbol{1}. \end{array}$$

Before stating a consistency estimate, let us introduce some Hölder spaces: let $\boldsymbol{\alpha}$ belong to $\mathbb{R}^d_+$. Call $[\boldsymbol{\alpha}]$ the vector of $\mathbb{N}^d$ whose i^{th} component is the integer part of α_i. Call $\{\boldsymbol{\alpha}\} = \boldsymbol{\alpha} - [\boldsymbol{\alpha}]$. We note $\mathcal{C}^{\boldsymbol{\alpha}}(\bar{\Omega})$ the space of continuous functions u such that for all $\boldsymbol{\beta} \le [\boldsymbol{\alpha}]$, $D^{\boldsymbol{\beta}} u$ is continuous and

$$\sup \left\{ \frac{|D^{[\boldsymbol{\alpha}]} u(\boldsymbol{x} + \boldsymbol{h}) - D^{[\boldsymbol{\alpha}]}(\boldsymbol{x})|}{|h_1|^{\{\alpha_1\}} \ldots |h_d|^{\{\alpha_d\}}}, \boldsymbol{x}, \boldsymbol{x} + \boldsymbol{h} \in \Omega, |h_i| > 0, i = 1, \ldots, d \right\} < +\infty.$$

The last quantity corresponds to a semi-norm on $\mathcal{C}^{\boldsymbol{\alpha}}(\bar{\Omega})$, which we call $|u|_{\mathcal{C}^{\boldsymbol{\alpha}}(\bar{\Omega})}$.
Theorem 4.6 is the key to the following consistency estimate, obtained in [5]:

Theorem 4.7. *Assume that $u \in \mathcal{C}^{\boldsymbol{\alpha}}(\bar{\Omega})$, where $\alpha_1 > 2$, $\alpha_i > 0$, $i = 2, \ldots, d$, and that $u = 0$ on $\partial\Omega$. Let P^n be the restriction operator on the sparse grid Ω_n: $P^n(u) = u(\Omega_n)$. We have the consistency error estimate*

$$\|P^n(\frac{\partial^2 u}{\partial x_1^2}) - P^n \circ D_n(u)\|_\infty \le C n 2^{-n \min(\alpha_1 - 2, \alpha_2, \ldots, \alpha_d, 2)} |u|_{\mathcal{C}^{\boldsymbol{\alpha}}(\bar{\Omega})}. \tag{36}$$

Similarly, for the sparse discretization of the Laplace operator, the consistency error may be bounded by $Cn2^{-n \min(\alpha_1 - 2, \alpha_2 - 2, \ldots, \alpha_d - 2, 2)} |u|_{\mathcal{C}^{\boldsymbol{\alpha}}(\bar{\Omega})}$ if $u \in \mathcal{C}^{\boldsymbol{\alpha}}(\bar{\Omega})$ with $\alpha_i > 2$, $i = 1, \ldots, d$. We see that the sparse grid discretization of Δ is consistent and that the consistency error is almost of the same order (up to the factor n) as the consistency error obtained with a full tensor grid.
We are left with studying the stability of the sparse grid discretization. As far as we know, there is unfortunately no theoretical stability estimates. There is even no proof that the

matrix D arising in the discrete problem is invertible. Indeed D does not fall into the well studied classes of matrices: in particular, D is neither a symmetric nor a M matrix. No discrete maximum principle is available. Nevertheless, numerical tests were done in [9], indicating that the stability constant , *i.e.*, $\|D^{-1}\|_\infty$ is bounded by Cn^{d-1}.
If such a stability estimate is true, we see that the sparse grid discretization of the Poisson problem is convergent, with an error of the order of $n^d 2^{-n\min(\alpha_1-2,\alpha_2-2,\ldots,\alpha_d-2,2)}$, if $u \in \mathcal{C}^{\alpha}(\bar{\Omega})$ with $\alpha_i > 2$, $i = 1,\ldots,d$.

4.5 Sparse Grid in practice

Consider Poisson problem on $(0,1)^d$ with Dirichlet homogene conditions.

$$\Delta u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} u = f. \tag{37}$$

The discrete approximation of this problem is obtain by (28) :

$$\left(\sum_{i=1}^{n} T_i D^{1,i} T_i^{-1} \right) U = F. \tag{38}$$

In fact, all application T_i,$D^{1,i}$,and T_i^{-1} is associated to matrix with only three non zero entries by row. So Matrix application is never construct to preserve optimal complexity. and the best way to solve (38) without construct matrix, is to use Krylov minimization methods like GMRES or BICGSTAB. We only need to matrix vector application construct by applying step by step all applications in $\left(\sum_{i=1}^{n} T_i D^{1,i} T_i^{-1} \right) U$.

Examples to solve Poisson Problem and Heat equation is given in examples/libpnl repository. It seems to be the best way to study Sparse Grid code.

More details on implementation is given in [7] chaper 9.

5 Application of Sparse Grid Method in Finance

We study in this part, the discretization of (8) by a sparse finite difference scheme.

5.1 Exact PDE formulation

Localization The initial problem (8) is given on a non-bounded domain. To apply numerical method, we have to truncate domain and impose boundary condition.

- on $y_i = y_i^{max}$, flow is out drawing so boundary condition don't impact solution in center of domain.
- on $x = x^{max}$ or $x = x^{min}$, $u(x, y_1, \ldots, y_n, t) \approx 0$.

$$\frac{\partial u}{\partial t} - \frac{1}{2} f(y)^2 \frac{\partial^2 u}{\partial x^2} + \frac{1}{2} f(y)^2 \frac{\partial u}{\partial x} - \rho \sum_{i=1}^{n} \beta_i f(y) \frac{\partial^2 u}{\partial x \partial y_i} \tag{39}$$

$$-\frac{1}{2} \sum_{i,j=1}^{n} \beta_i \beta_j \frac{\partial^2 u}{\partial y_i \partial y_j} + \sum_{i=1}^{n} \lambda_i y_i \frac{\partial u}{\partial y_i} = \frac{1}{2} \left(f(y)^2 - \sigma_0^2 \right) Vega_{BS} (x, t) , \tag{40}$$

with $(x, y, t) \in [x^{min}, x^{max}] \times [-y_i^{max}, y_i^{max}] \times [0, T]$ and

$$\begin{aligned} u(x, y, 0) &= 0 \quad (x, y)\,(x, y, t) \in [x^{min}, x^{max}] \times [-y_i^{max}, y_i^{max}] \\ u(x^{min/max}, y, t) &= 0 \quad (y, t) \in [-y_i^{max}, y_i^{max}] \times [0, T] \\ u(x, y^{min/max}, t) &= 0 \quad (x, t) \in [x^{minx}, x^{max}] \times [0, T] \end{aligned} \tag{41}$$

Technical point To reduce point wise multiplications, operator $\mathcal{L}$ given by

$$\mathcal{L}u = \frac{1}{2} f(y)^2 \frac{\partial^2 u}{\partial x^2} - \frac{1}{2} f(y)^2 \frac{\partial u}{\partial x} + \rho \sum_{i=1}^{n} \beta_i f(y) \frac{\partial^2 u}{\partial x \partial y_i} + \frac{1}{2} \sum_{i,j=1}^{n} \beta_i \beta_j \frac{\partial^2 u}{\partial y_i \partial y_j} - \sum_{i=1}^{n} \lambda_i y_i \frac{\partial u}{\partial y_i}. \tag{42}$$

is written in source code as follow

$$\begin{aligned} \mathcal{L}u = \frac{1}{2} \frac{\partial^2 f(y)^2 u}{\partial x^2} - \frac{1}{2} \frac{\partial f(y)^2 u}{\partial x} + \rho \sum_{i=1}^{n} \beta_i \left(\frac{\partial^2 f(y) u}{\partial x \partial y_i} - \frac{\partial ln f(y)}{\partial y_i} \frac{\partial f(y) u}{\partial x} \right) \\ + \frac{1}{2} \sum_{i,j=1}^{n} \beta_i \beta_j \frac{\partial^2 u}{\partial y_i \partial y_j} - \sum_{i=1}^{n} \lambda_i y_i \frac{\partial u}{\partial y_i}. \end{aligned} \tag{43}$$

5.2 Time discretization

We apply θ-scheme for time discretization. We note $u_n(x, y)$ the approximation of u at time t_n. Equation (39) becomes

$$\frac{u_n - u_{n-1}}{\Delta t} - \mathcal{L} \left(\theta u_n + (1 - \theta) u_{n-1} \right) = \frac{1}{2} \left(f(y)^2 - \sigma_0^2 \right) Vega_{BS} \left(x, t_{n-\frac{1}{2}} \right). \tag{44}$$

We have to solve at each time step, same equation as (37) :

$$\left(1 - \theta \Delta t \mathcal{L} \right) u_n = f, \tag{45}$$

with

$$\left(1 + (1 - \theta) \Delta t \mathcal{L} \right) u_{n-1} + \Delta t \frac{1}{2} \left(f(y)^2 - \sigma_0^2 \right) Vega_{BS} \left(x, t_{n-\frac{1}{2}} \right). \tag{46}$$

6 Conclusion

Lots of details not given here like : use of other boundary conditions, preconditionner construction, code structure, . . . , could be found in [7].

References

[1] Achdou, Y. and Pironneau, O. (2005) .*Computational methods for option pricing*, Frontiers in Applied Mathematics. 2

[2] Bungartz, H. J. and Griebel, M.(2004),*Sparse grids*, Acta Numerica 13. 4

[3] Fouque, Jean-Pierre and Papanicolaou, George and Sircar, K. Ronnie(2000).*Derivatives in financial markets with stochastic volatility*, Cambridge University Press.

[4] Griebel, M.,(1998) *Adaptive Sparse Grid Multilevel Methods for Elliptic PDEs Based on Finite Differences*, Computing 61-2. 2

[5] Koster, F. (2000),*A proof of the consistency of the finite difference technique on sparse grids*, Computing 65-3. 4

[6] Petersdoff, V. T. and Schwab, C. (2004), *Numerical Solutions of Parabolic Equations in High Dimensions*, Mathematical Modelling and Numerical Analysis. 4, 11

[7] Pommier, D. (2008),*Méthodes numériques sur des grilles sparse appliquées à l'évaluation d'options en finance*, Phd thesis, Université Pierre & Marie Curie, Paris. 4

[8] Reisinger, C. (2004),*Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben*, Phd thesis, Universitat Heidelberg. 12, 13

[9] Schiekofer, T. (1998),*Die methode der finiten differenzen auf dünnen gittern zur lösung elliptischer und paraboliкscher*, Phd thesis, Universität Bonn.

4, 12