

---

# Asymptotic Expansion Around Principal Components and the Complexity of Dimension Adaptive Algorithms

Christoph Reisinger

Mathematical Institute, University of Oxford,  
christoph.reisinger@maths.ox.ac.uk

**Summary.** In this short article, we describe how the correlation of typical diffusion processes arising e.g. in financial modelling can be exploited – by means of asymptotic analysis of principal components – to make Feynman-Kac PDEs of high dimension computationally tractable. We explore the links to dimension adaptive sparse grids [GG03], anchored ANOVA decompositions and dimension-wise integration [GH10], and the embedding in infinite-dimensional weighted spaces [SW98]. The approach is shown to give sufficient accuracy for the valuation of index options in practice. These numerical findings are backed up by a complexity analysis that explains the independence of the computational effort of the dimension in relevant parameter regimes.

## 1 Introduction

The motivation for this research comes from financial engineering. Often, the value of financial derivatives is conveniently modelled by partial differential equations. This can lead to highly efficient numerical finite difference and finite element schemes in low to moderate dimensions, however it presents extreme numerical challenges if the dimension is high. There are arguably two main origins: In the first class of applications, the high-dimensionality arises because the financial derivative depends on the whole path of a stock (see e.g. [CMO97]); in the second class, the derivative depends on the value of multiple stocks at a fixed future time. From the PDE perspective, these cases are somewhat different and we focus on the latter category here.

Consider stocks  $S_1, \dots, S_d$ , which are modelled in a standard Black-Scholes setting as geometric Brownian motions with covariance matrix  $\Sigma = (\rho_{ij}\sigma_i\sigma_j)_{1 \leq i, j \leq d}$ , where  $\sigma_i$  are the volatilities and  $\rho_{ij}$  the correlations.

As a running example, consider further a basket  $\sum_{i=1}^d \mu_i S_i$  with weights  $\mu_i \geq 0$ , and a European put option on this basket with strike  $K$  and expiry  $T$ . The value of this option satisfies the Black-Scholes equation

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sum_{i,j=1}^d \rho_{ij} \sigma_i \sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} + r \sum_{i=1}^d S_i \frac{\partial V}{\partial S_i} - rV = 0, \quad (1)$$

for all  $S_i > 0$ ,  $t \in [0, T]$ , supplemented with a terminal condition

$$V(S, T) = P(S) := \max \left( K - \sum_{i=1}^d \mu_i S_i, 0 \right), \quad (2)$$

where  $S = (S_1, \dots, S_d) \in \mathbb{R}^d$  and  $P$  is the payoff.

The Black-Scholes equation (1) has the special property that it can be transformed to the (forward) heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i=1}^d \lambda_i \frac{\partial^2 u}{\partial x_i^2}, \quad (3)$$

for  $x \in \mathbb{R}^d$ , with appropriate initial data  $u(x, 0) = \tilde{P}(x)$ . This is a direct consequence of the assumed log-normality of  $S_i$ . Here,  $\lambda_i$  are the eigenvalues of the covariance matrix of the underlying Brownian drivers and are found to decay rapidly in typical applications (see Section 3). This opens the possibility of a perturbation analysis.

The computational aspects of such an approach are discussed in [RW07]; meanwhile, this has been developed further by [HKS07]. Here, we focus on the underlying expansion itself, and discuss its position within recent work on fundamentally similar ideas. Almost all of this work has been done in the context of high-dimensional cubature problems, and we can relate these approaches to the present context by noting that the solution to (3) can be written as

$$u(x, t) = \frac{1}{t^{d/2} \pi^{d/2} \prod_{i=1}^d \lambda_i^{1/2}} \int_{\mathbb{R}^d} \exp \left( - \sum_{i=1}^d (x_i - x'_i)^2 / (2\lambda_i) \right) \tilde{P}(x') dx'. \quad (4)$$

Principal component analysis provides a natural ordering of dimensions. Such an ordering can also be found for options whose payoff depends on a Brownian path, through a hierarchical (e.g. Brownian bridge) construction. In both of these settings, the feasibility of the high-dimensional integration problem depends on the speed of decay of the high-dimensional contributions, as well as their regularity. A natural way to quantify this is via weighted mixed Sobolev norms [SW98]. We will see later that the  $\lambda_i$  can be directly mapped to those weights.

The upshot is that many applications in finance have a relatively low superposition dimension [WS05], i.e. their solution can be accurately represented by the sum of solutions to lower-dimensional problems. For integration problems, this can be exploited via quasi-Monte Carlo methods [WS05, CMO97], dimension adaptive and generalised sparse grids [GG03, HGC07], dimension-wise integration based on anchored ANOVA decomposition [GH10], or related ideas.

The rest of the article is structured as follows. In Section 2, we present results of a simple dimension adaptive sparse grid strategy applied to PDEs of type (1) and (3) to illustrate the dependence of the convergence speed on the choice of coordinates. Section 3 analyses common properties of typical covariance matrices in equity and interest rate markets, and derives an asymptotic expansion around a principal component, which ultimately leads to an ANOVA-type decomposition. Section 4 shows that for exponentially decaying weights of the higher order PCA components, as observed for index options, the complexity can become independent of the nominal dimension. Section 5 critiques the findings and outlines directions for further research.

## 2 Dimension adaptive sparse grids

In this section, we investigate by numerical experiments the role of the underlying coordinate systems in computations.

### 2.1 Construction

We explain the construction for the two-dimensional case. It is based on a heuristic algorithm for the approximation of a numerical solution in optimal complexity from computations on suitably chosen Cartesian grids. The candidate set of Cartesian grids is obtained by bisection from a coarse grid, say the unit square  $[0, 1]^2$ , such that the mesh widths in the two directions are  $(h_i, h_j) = (2^{-i}, 2^{-j})$ ,  $(i, j) \in \mathbb{N}_0^2$ . We denote the corresponding numerical solution by  $U(i, j)$ , and have in mind pointwise evaluations of a finite difference solution or other functionals thereof.

To assess the contribution of a particular  $U(i, j)$  to the overall solution, we study the *hierarchical surplus* (see e.g. [BG04])

$$\begin{aligned} \delta U(i, j) &= U(i, j) - U(i-1, j) - U(i, j-1) + U(i-1, j-1) \\ &= \delta_1^- \delta_2^- U(i, j), \end{aligned}$$

where  $\delta_1^- U(i, j) = U(i, j) - U(i-1, j)$  and  $\delta_2^- U(i, j) = U(i, j) - U(i, j-1)$  for  $i, j > 0$  are backward difference operators, and  $\delta_1^- U(i, j) = U(i, j)$  if  $i = 0$ ,  $\delta_2^- U(i, j) = U(i, j)$  if  $j = 0$ .

A numerical approximation on level  $n$  is defined by the choice of grids described by an index set  $\mathcal{M}_n \subset \mathbb{N}_0^2$  via

$$u_n := \sum_{(i,j) \in \mathcal{M}_n} \delta U(i, j).$$

We require that  $\mathcal{M}_n$  is convex in the sense

$$\mathcal{M}_n = C(\mathcal{M}_n) := \mathcal{M}_n \cup \{(i, j) \in \mathbb{N}_0^2 : (i, j+1) \in \mathcal{M}_n \vee (i+1, j) \in \mathcal{M}_n\}.$$

For a convergent discretisation, we expect

$$u_n \rightarrow u \quad \text{for } \mathcal{M}_n \uparrow \mathbb{N}_0^2,$$

where the last expression is to be understood in the sense  $\mathcal{M}_n \subset \mathcal{M}_{n+1}$  and  $\forall (i, j) \in \mathbb{N}_0^2 \exists n : (i, j) \in \mathcal{M}_n$ . An error bound will be given by

$$|u - u_n| \leq \sum_{(i,j) \notin \mathcal{M}_n} |\delta U(i, j)|.$$

The following construction of  $\mathcal{M}_n$  is *dimension adaptive* (see [GG03] for a similar strategy). Start with  $\mathcal{M}_0 = \{(0, 0)\}$ . For given  $\mathcal{M}_n$ ,  $n \geq 0$ , consider candidate nodes

$$\mathcal{C}_n = \{(i, j) \in \mathbb{N}_0^2 : (i, j) \notin \mathcal{M}_n, (i-1, j) \in \mathcal{M}_n \vee (i, j-1) \in \mathcal{M}_n\}.$$

Then refine

$$\mathcal{M}_{n+1} = C(\mathcal{M}_n \cup \{(i, j) \in \mathcal{C}_n : |\delta U(i, j)| \geq \gamma c_n\}),$$

where

$$c_n = \max\{|\delta U(i, j)| : (i, j) \in \mathcal{C}_n\}.$$

The idea is that we refine only in those directions where there is a significant contribution from the refined grid. We pick  $\gamma$  slightly smaller than the asymptotic ratio of the surpluses between two refinement levels.

This draws on the theoretical properties of the used discretisation. We use below second order finite difference and finite element schemes. From the analysis of the sparse grid combination technique, see e.g. [PZ99] or [R11], we expect

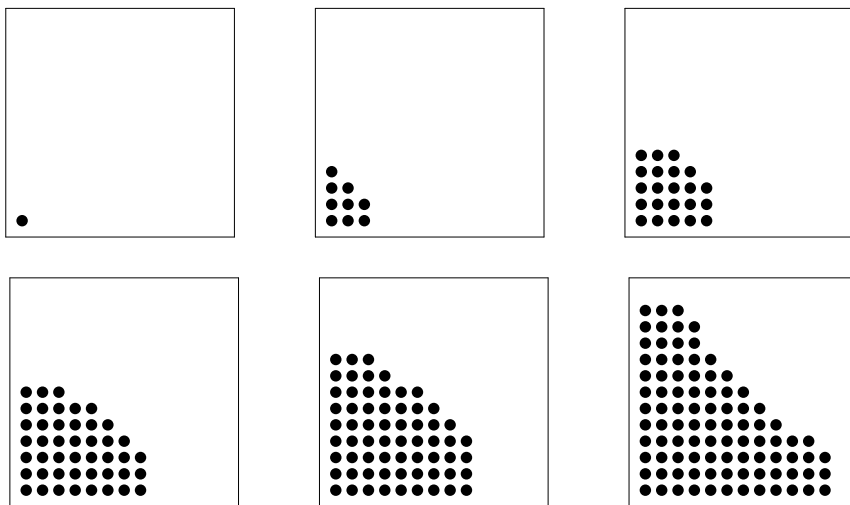
$$\delta U(i, j) \sim c h_i^2 h_j^2. \quad (5)$$

Refining in one direction,  $i \rightarrow i+1$  or  $j \rightarrow j+1$ , we expect the surplus to decrease by a factor of 1/4. Hence we used  $\gamma = 0.2$  in the numerical computations below.

## 2.2 An example

In the following subsections, we illustrate the behaviour of the above strategy on the example of the Black-Scholes model for a put option on an equity basket consisting of BMW and Daimler. This amounts to solving the PDE (1) for  $d = 2$ , for parameters  $\sigma_1 = 0.438$ ,  $\sigma_2 = 0.616$ ,  $\rho = 0.89$ ,  $r = 0.05$ . The terminal condition at  $t = T = 1$  is given by (2) with  $\mu_1 = 0.384$ ,  $\mu_2 = 0.616$ ,  $K = 0.25$ . We are interested in the numerical solution evaluated at  $(S_1, S_2) = (K, K)$  at time  $t = 0$ .

We use a standard central difference scheme with fractional-step  $\theta$ -time-stepping. The domain is truncated at sufficiently large values of  $S_1$  and  $S_2$

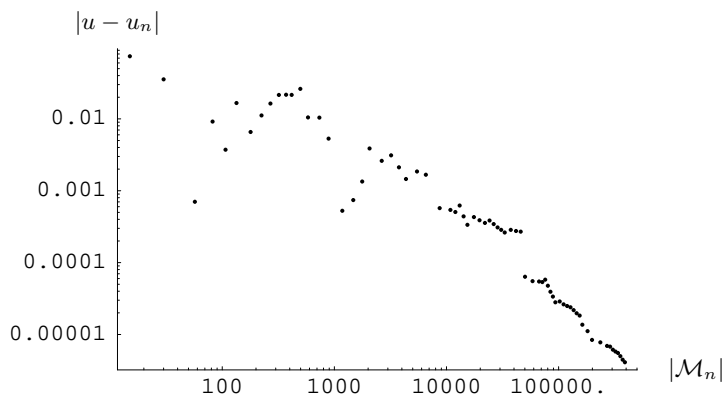


**Fig. 1.** The black dots correspond to indices  $(i, j)$  included in  $\mathcal{M}_n$  at various refinement step  $n$  from  $n = 0$ , i.e.  $i = j = 0$ , top left box. For this example where non of the two directions is dominant, the refinement is largely symmetric.

and then transformed to the unit square, where appropriate asymptotically exact boundary conditions are set.

Fig. 1 illustrates the resulting adaptive refinement strategy. It shows that the grid construction is very similar to that of a standard sparse grid, where  $\mathcal{M}_n = \{(i, j) \in \mathbb{N}_0^2 : i + j \leq n\}$ , confirming experimentally the optimality of the standard sparse grid [Z91, BG04].

Fig. 2 gives an indication of the efficiency of the method. Ideally, grids



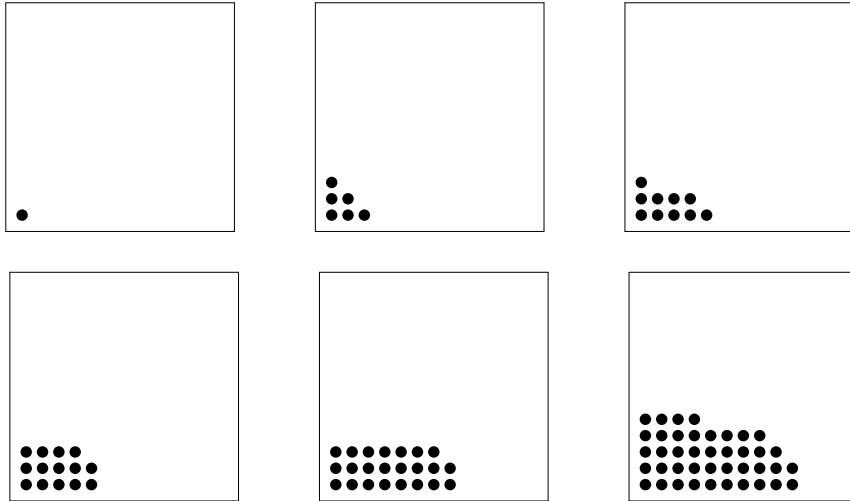
**Fig. 2.** Denoting by  $|\mathcal{M}_n|$  the number of all grid points of all grids in  $\mathcal{M}_n$ , shown is the approximation error for an increasing number of unknowns. The slope for large  $n$  is consistent with asymptotic second order convergence.

would be included in decreasing order of importance. It is seen that for large enough  $n$ , where the grid sizes are small enough for (5) to be a valid approximation, this is largely the case.

### 2.3 Principal components

Sparse grid solutions, and indeed all approximations based on tensor product spaces, are by construction dependent on the choice of an underlying coordinate system, if the underlying problem is anisotropic. We now seek to exploit the fact that our dimension adaptive method can use this to its advantage. To this end, transform (1) to (3); then, in the above example, one finds  $\lambda_1 = 0.431$ ,  $\lambda_2 = 0.024$ .

Repeating the above numerical test in transformed coordinates, leads to the refinement strategy in Fig. 3. Reassuringly, the adaptive algorithm detects



**Fig. 3.** Similar to Fig. 1, but now in eigenvector coordinates. Refinement is stronger in the direction of the principal component.

the anisotropy of the operator and refines mostly in the direction where the variation is largest.

Fig. 4 provides a different angle on this behaviour. It shows the (logarithm of the) hierarchical surplus for different grids, indexed by  $(i, j)$ . Revisiting the derivation of the hierarchical surplus for a central finite difference scheme in [R11], explicitly taking into account the dependence of the truncation error on the eigenvalues, leads to

$$\delta_1^- U(i, j) \sim c_1 \lambda_1 h_i^2, \quad (6)$$

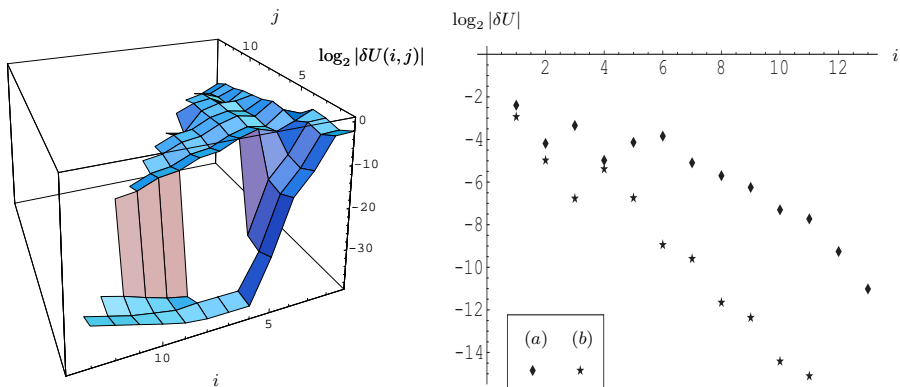
$$\delta_2^- U(i, j) \sim c_2 \lambda_2 h_j^2, \quad (7)$$

$$\delta U(i, j) \sim c_0 \lambda_1 \lambda_2 h_i^2 h_j^2. \quad (8)$$

This is based on the simplifying assumption that the smoothness of the solution in the two directions does not depend on  $\lambda_1$  and  $\lambda_2$ , which will be a reasonable assumption for moderate timescales.

The left plot appears to confirm that the surplus is asymptotically a function of  $i + j$ , as suggested by (8), as long as  $i, j > 2$ , i.e. excluding cases where the solution is solely or predominantly determined by the boundary values.

The plot on the right, which takes cross-section in the first and second directions, reveals the anisotropy of the problem. Close to the ‘edges’ of the grid table, i.e. where  $j$  or  $i$  are small, the surplus will be similar to (6) or (7), respectively. This determines the different ‘constant’ factors in the surplus in these directions, and can explain the offset of the two curves in Fig. 4 (right). As we will discuss later, the ratio of  $\lambda_1/\lambda_2$ , here approximately 20, is not the only relevant factor, but also the initial data to the PDE.



**Fig. 4.** Left: The surplus  $U(i, j)$  for different grids  $(i, j)$ . Right: Cross-sections along the edges (a)  $\log_2 |\delta U(i, 3)|$ , (b)  $\log_2 |\delta U(2, i)|$ .

An intriguing new phenomenon is observed when we go to three dimensions. When we add Volkswagen to the basket from before, the eigenvalues are estimated as  $\lambda_1 = 0.653$ ,  $\lambda_2 = 0.069$ ,  $\lambda_3 = 0.023$ , and we solve the corresponding PDE (3) in  $d = 3$  dimensions.

At the finest computed level, 2197 grids are involved. This compares in complexity roughly to a regular three-dimensional sparse grid on level 20, which consists of 2023 grids. Out of these 2197 grids, 469 grids of  $\mathcal{M}$  are at the ‘boundary’ of the grid table described by

$$\mathcal{B} = \{(i, j, k) \in \mathbb{N}_0^3 : i \cdot j \cdot k = 0\},$$

i.e. for which at least one direction is unrefined, while 1728 grids are in the ‘interior’  $\mathcal{M} \setminus \mathcal{B}$ . However, the contribution from the boundary elements to the solution is

$$\sum_{(i,j,k) \in \mathcal{B} \cap \mathcal{M}} |\delta U(i,j,k)| = 0.4369,$$

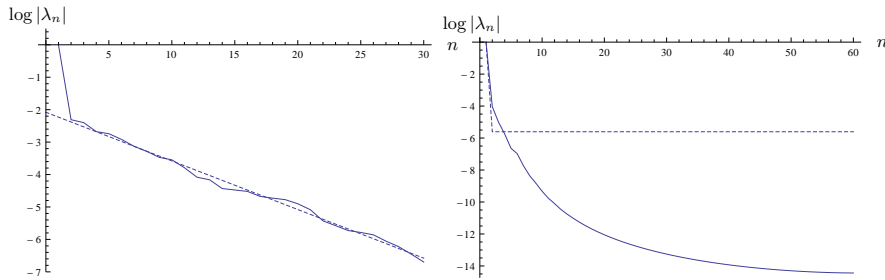
while

$$\sum_{(i,j,k) \in \mathcal{M} \setminus \mathcal{B}} |\delta U(i,j,k)| = 0.0035.$$

This suggests that the problem can be well approximated by sums of two-dimensional approximations. We investigate this for general dimensions in the following section.

### 3 PCA and asymptotic expansion

This section explores communalities of the eigensystem of important covariance matrices, and their use for problem-adapted high-dimensional approximation.



**Fig. 5.** Shown are the (logarithms of) normalised eigenvalues of three representative examples of covariance matrices from financial modelling. Left: Estimated covariances  $\sigma_i \sigma_j \rho_{ij}$  from historical time series of the DAX30, i.e.  $d = 30$ , and a fitted curve (which will be discussed in Section 4). Right: Two stylized forward rate models, both for  $d = 60$ : (a) constant correlation  $\rho = 0.8$  for all tenors, constant volatility; (b) humped-shaped volatility  $\sigma_i = (A + B i) \exp(-C i) + D$  with  $A = 0.1$ ,  $B = 0.1$ ,  $C = 1$ ,  $D = 0.1$ , and exponentially decaying correlation  $\rho_{ij} = \exp(-\alpha|i - j|)$ ,  $\alpha = 0.025$ . Eigenvalues are numbered in decreasing order.

A common feature to the examples in Fig. 5 is the presence of a dominant eigenvalue, which is by a factor of 10 or more larger than the rest of the spectrum. This motivates expanding the solution to (3) in these small parameters to obtain approximations to the high-dimensional solution by solving low-dimensional problems. Consider therefore  $u$  as a function of  $\lambda = (\lambda_i)_{1 < i \leq d} \in \mathbb{R}^{d-1}$ , where



$$\lambda_i \ll \lambda_1, \quad i > 1.$$

Assuming the dependence of  $u$  on  $\lambda$  is sufficiently smooth, first order Taylor expansion gives

$$u(\lambda) = u(0) + \sum_{i=2}^d \lambda_i \frac{\partial u}{\partial \lambda_i}(0) + O(\|\lambda\|_2^2). \quad (9)$$

For special cases, the eigenvalue sensitivities can be calculated analytically. An example is the value of a put option in the Black-Scholes model, where differentiation of the integral formula (4) gives easily computable closed-form expressions similar to the standard one-factor Black-Scholes formula [R04]. The correction terms also satisfy simple PDEs, i.e. for

$$u^{\{i\}} := \frac{\partial u}{\partial \lambda_i}(0), \quad (10)$$

$$\frac{\partial}{\partial t} u^{\{i\}} = \frac{1}{2} \lambda_1 \frac{\partial^2}{\partial x_1^2} u^{\{i\}} + \frac{1}{2} \frac{\partial^2 u}{\partial x_i^2}(0), \quad (11)$$

where the last equation follows from application of (3) to (10).

More generally, the exact sensitivity can be replaced by a finite difference. Writing, by slight abuse of notation,  $u(\lambda_i)$  for the solution to the problem where all coefficients except  $\lambda_1$  and  $\lambda_i$  are zero,

$$\frac{\partial u}{\partial \lambda_i}(0) = \frac{u(\lambda_i) - u(0)}{\lambda_i} + O(\lambda_i).$$

This coincides with the financial industry practice of parameter ‘bumping’. Inserting in (9),

$$u(\lambda) = u(0) + \sum_{i=2}^d (u(\lambda_i) - u(0)) + O(\|\lambda\|_2^2) \quad (12)$$

is a second order approximation in the small parameters  $\lambda_i \leq \lambda_2 \ll \lambda_1$ ,  $i = 2, \dots, d$ .

In a sense, this formula is even ‘better’ than the one using the exact derivatives, as it captures functions of superposition dimension one *exactly*. We discuss this in more detail below.

Moreover, the lead term  $u(0)$  requires the solution of only a one-dimensional PDE in  $x_1$ ,

$$\frac{\partial u}{\partial t}(0) = \frac{1}{2} \lambda_1 \frac{\partial^2 u}{\partial x_1^2}(0),$$

and the correction terms  $u(\lambda_i)$  the solution of  $d - 1$  two-dimensional PDEs in  $x_1$  and  $x_i$ ,

$$\frac{\partial u}{\partial t}(\lambda_i) = \frac{1}{2} \lambda_1 \frac{\partial^2 u}{\partial x_1^2}(\lambda_i) + \frac{1}{2} \lambda_i \frac{\partial^2 u}{\partial x_i^2}(\lambda_i).$$

For the test example of the DAX given above, [RW07] report an error against a Monte Carlo simulation benchmark of  $< 0.06\%$  of the option value, with an absolute error of 0.000073 at the strike  $K = 1$ , which is less than 1 basis point. Results for other (and smaller) baskets are comparable, such that the approximation seems sufficient for practical applications.

What is more, it is not difficult to extend this expansion to higher order, e.g. to order two and using an approximation to the cross-derivatives in  $\lambda_i$  and  $\lambda_j$ ,

$$u(\lambda) = u(0) + \sum_{i=2}^d (u(\lambda_i) - u(0)) + \sum_{i \neq j} (u(\lambda_i, \lambda_j) - u(\lambda_i) - u(\lambda_j) + u(0)) + O(\|\lambda\|_2^3),$$

where  $u(\lambda_i, \lambda_j)$  is the solution where all coefficients except  $\lambda_1$ ,  $\lambda_i$  and  $\lambda_j$  are set zero. Hence, the additional correction terms are the solution of  $(d-1)(d-2)/2$  three-dimensional PDEs.

Proceeding in this way to order  $d$ ,

$$u(\lambda) = u(0) + \sum_{i=2}^d \Delta_i u + \sum_{i \neq j} \Delta_i \Delta_j u + \dots + \Delta_2 \dots \Delta_d u,$$

where  $\Delta_i u = u(\lambda_i) - u(0)$  etc. This is equivalent to an anchored ANOVA decomposition [GH10]. The functions  $u(\lambda_i)$  have the interpretation

$$u(\lambda_i) = \mathbb{E}(\tilde{P}(X_T)), \quad (13)$$

$$dX_t^k = \begin{cases} \lambda_k dW_t^k & k \in \{1, i\}, \\ 0 & \text{else,} \end{cases} \quad (14)$$

$$X_0^k = x_k, \quad (15)$$

where  $W$  is a standard Brownian motion. Similarly,  $u(\lambda_i, \lambda_j)$  corresponds to a process as per (14), but with diffusion in directions  $\{1, i, j\}$ , with the obvious extension to higher order. Writing these in integral form,

$$u(\lambda_i) = \frac{1}{t\pi\sqrt{\lambda_1\lambda_i}} \int_{\mathbb{R}^2} \exp(-(x_1 - x'_1)^2/\lambda_1) \exp(-(x_i - x'_i)^2/\lambda_i) \tilde{P}(x') dx'_1 dx'_i,$$

and similar for higher order terms, leads to dimension-wise quadrature as in [GH10].

It should be noted that this choice of coordinates does not take into account the initial data of the PDE. It is conceivable to construct a pathological example where the approximation is arbitrarily bad, e.g. where the principal component is parallel to a level curve of the initial data, and orthogonal to a direction of high curvature.

Basket options are a fortunate case in this respect. The direction of steepest change of the basket value, which determines the payoff, is roughly aligned with the principal component of the covariance matrix. This results from positive basket weights, and the fact that the principal eigenvector of a positive matrix has positive entries (guaranteed by the Perron-Frobenius Theorem).

## 4 A simple and some more complex complexity results

Formula (12) suggests that the complexity is linear in the dimension  $d$  of the problem, assuming the quadratic remainder term can be neglected, and that the effort for the eigenvalue decomposition and transformation is negligible, which is the case in practice.

We argue in this section that it is not necessary to include all  $d$  terms, by exploiting the decay in  $\lambda_i$  further. The empirical evidence which underpins this analysis is already seen in Fig. 5, where the ordered and normalised ( $\lambda_1 = 1$ ) eigenvalues of the covariance matrix are plotted together with a regression of the set excluding the first one, which gives  $\lambda_n \approx 0.108 \cdot 0.861^{n-2}$  for  $n \geq 2$ . There is clear evidence of exponential decay of the eigenvalues, following a big jump of about 90% of the first eigenvalue.

We therefore assume in the following that  $\lambda_i \leq \delta \cdot \lambda^i$  for some  $\delta \ll 1$  and  $\lambda < 1$ , but with  $\lambda \approx 1$ . In this parameter setting, it is justified to neglect the  $\delta^2$  term (we point to the fact that the residual error was negligible in the numerical case study), but a potentially large number of first order terms gives significant contributions if  $d$  is large. As in [SW98], and more recently [NHMR11, G11], we embed the problem in an infinite-dimensional space as  $d \rightarrow \infty$ .

All of this, combined with (9), motivates the following limiting case as basis for a simple complexity analysis.

**Assumption 4.1** *Assume that*

$$u(\lambda, \delta) = u_0 + \delta \sum_{i=1}^{\infty} \lambda^i u_i, \quad (16)$$

*and there is an algorithm which finds  $u_i$  with accuracy  $\epsilon$  in complexity*

$$C \leq \epsilon^{-1/p},$$

*i.e.  $p$  is the order of the method.*

We do not specify for this abstract result what the  $u_i$  are, but clearly we have in mind first order ANOVA terms. Then, in the above setting, we require the numerical solution of two-dimensional heat equations, so if we use standard linear finite elements and second order time-stepping,  $p = 2/3$ .

**Proposition 1.** *Under Assumption 4.1, there exists an algorithm, which finds  $u$  with accuracy  $\epsilon$  in complexity*

$$C \leq c \cdot (\epsilon/\delta)^{-1/p} \cdot (1 - \lambda^{1/(1+p)})^{-(1+p)/p},$$

*where  $c$  does not depend on  $\epsilon$ ,  $\lambda$ , or  $\delta$ .*

*Proof.* The proof is constructive. Let  $\epsilon_i$  the accuracy of  $u_i$ , then we minimise the sum of the costs  $\epsilon_i^{-1/p}$  under the constraint that the total error is  $\delta \sum_{i=0}^{\infty} \lambda^i \epsilon_i \leq \epsilon$ . A direct application of Lagrangian multipliers gives the result.

This extends straightforwardly to higher order in  $\delta$ , i.e. if

$$u(\lambda, \delta) = u_0 + \delta \sum_{i=1}^{\infty} \lambda^i u_i + \dots + \delta^q \sum_{i_1 \neq \dots \neq i_q} \lambda^{i_1} \dots \lambda^{i_q} u_{i_1, \dots, i_q}, \quad (17)$$

and the complexity for solving for  $u_{i_1, \dots, i_q}$  is  $\epsilon^{-q/p}$ , then the total complexity is of order  $(\epsilon/\delta^q)^{-q/p}$ , although the constant factors become less explicitly computable.

Ultimately, one would want to adaptively pick  $q$  to ensure the overall error is below some desired  $\epsilon$ , see e.g. [SW98, GH10, NHMR11, G11]. An appropriate measure for the overall complexity, in the PDE setting as for cubature, are weighted Sobolev norms.

The size and numerical approximation of  $u_{i_1, \dots, i_q}$  depends on the regularity of the original problem, and, as [GKS10] show, the lower-order ANOVA terms may have higher regularity. We finish this section by outlining how such a result may be derived in the setting of this article, where the correction terms in (17) are essentially determined by the parameter sensitivities

$$u^{\{i_1, \dots, i_m\}} := \frac{\partial^m u}{\partial \lambda_{i_1} \dots \partial \lambda_{i_m}}.$$

For the example of the basket option, these are explicitly computable from the integral formula (4), which is carried out for the first order terms in [R04].

We mention an alternative route, which is more generally applicable. Repeated differentiation of (11) gives

$$\frac{\partial}{\partial t} u^{\{i_1, \dots, i_m\}} = \frac{1}{2} \sum_j \lambda_j \frac{\partial^2}{\partial x_j^2} u^{\{i_1, \dots, i_m\}} + \frac{1}{2} \sum_{k=1}^m \frac{\partial^2}{\partial x_{i_k}^2} u^{\{i_1, \dots, i_m\} \setminus \{i_k\}},$$

which allows the recursive application of standard regularity results for parabolic PDEs. This will give bounds on the sensitivities (ANOVA terms) and corresponding weights for the complexity analysis, and will be the subject of future research.

## 5 Conclusions and extensions

The principal component analysis of multivariate diffusion problems not only gives an ordering of the dimensions, but allows a formal Taylor expansion in terms of the spectrum of the covariance matrix, which leads to a decomposition equivalent to an anchored ANOVA decomposition of the solution. Model problems from financial engineering considered in the literature show that already a first order approximation, which corresponds to an approximation with superposition dimension one, is sufficiently accurate for practical use.

The question arises if this strategy works equally well for models beyond the normal (or log-normal) case. In models with non-constant volatilities and

drift, “freezing” the coefficients at the initial value of the process reduces the problem to the present case. The additional approximation error will be acceptable for not too long time horizons. Jump models will be the subject of future research, and the methodology should be transferable in spirit. An arguably more challenging extension is to non-linear problems, e.g. free boundary problems arising in American option valuation, but again the author does not foresee any fundamental difficulties.

## References

- [BG04] Bungartz, H.-J., Griebel, M.: Sparse grids, *Acta Numer.*, **13**, 11–23 (2004)
- [CMO97] Caffisch, R.E., Morokoff, W., Owen, A.: Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension. *J. Comput. Finance*, **1**, 27–46 (1997)
- [GG03] Gerstner, T., Griebel, M.: Dimension-adaptive tensor-product quadrature. *Computing*, **71**, 65–87 (2003)
- [GH10] Griebel, M., Holtz, M.: Dimension-wise integration of high-dimensional functions with applications to finance. *J. Complexity*, **26**, 455–489 (2010)
- [GKS10] Griebel, M., Kuo, F. Y., Sloan, I.H.: The smoothing effect of the ANOVA decomposition. *J. Complexity*, **26**, 523–551 (2010)
- [G11] Gnewuch, M.: Infinite-dimensional integration on weighted Hilbert spaces, Preprint cucs-016-10, Department of Computer Science, Columbia University (To appear in *Math. Comp.*) (2011)
- [H03] Hegland, M.: Adaptive sparse grids, *ANZIAM J.*, **44**, 335–353 (2003)
- [HGC07] Hegland, M., Garcke, J., Challis, V.: The combination technique and some generalisations. *Lin. Algebra Appl.*, **420**, 249–275 (2007)
- [HKS10] Hilber, N., Kehtari, S., Schwab, C., Winter, C.: Wavelet finite element method for option pricing in highdimensional diffusion market models. SAM Research Report 2010-01, ETH Zürich (2010)
- [NHMR11] Niu, B., Hickernell, F.J., Müller-Gronbach, T., Ritter, K.: Deterministic multi-level algorithms for infinite-dimensional integration on  $\mathbb{R}^N$ . *J. Complexity*, **27**, 331–351 (2011)
- [PZ99] Pflaum, C., Zhou, A.: Error analysis of the combination technique. *Numer. Math.*, **84**, 327–350 (1999)
- [R04] Reisinger, C.: Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben. PhD Thesis, Universität Heidelberg (2004)
- [RW07] Reisinger, C., Wittum, G.: Efficient hierarchical approximation of high-dimensional option pricing problems. *SIAM J. Sci. Comput.*, **29**, 440–458 (2007)
- [R11] Reisinger, C.: Analysis of linear difference schemes in the sparse grid combination technique. Preprint, [arXiv:0710.0491](https://arxiv.org/abs/0710.0491), University of Oxford (2011)
- [SW98] Sloan, I.H., Woźniakowski, H.: When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals? *J. Complexity*, **14**, 1–33 (1998)
- [WS05] Wang, X., Sloan, I.H.: Why are high-dimensional finance problems often of low effective dimension? *SIAM J. Sci. Comput.*, **27**, 159–183 (2005)

- [Z91] Zenger, C.: Sparse grids. In: Hackbusch, W. (ed) Parallel Algorithms for Partial Differential Equations, Notes on Numerical Fluid Mechanics, **31**, Vieweg, Braunschweig/Wiesbaden (1991)