Computing 71, 43–63 (2003)
Digital Object Identifier (DOI) 10.1007/s00607-003-0024-4

Computing
Printed in Austria

Sparse Finite Elements for Stochastic Elliptic Problems – Higher Order Moments*

Ch. Schwab and **R.A. Todor**, Zürich

Received August 6, 2002; revised June 12, 2003
Published online: September 1, 2003

Abstract

We define the higher order moments associated to the stochastic solution of an elliptic BVP in $D \subset \mathbb{R}^d$ with stochastic input data. We prove that the k-th moment solves a deterministic problem in $D^k \subset \mathbb{R}^{dk}$, for which we discuss well-posedness and regularity. We discretize the deterministic k-th moment problem using sparse grids and, exploiting a spline wavelet basis, we propose an efficient algorithm, of logarithmic-linear complexity, for solving the resulting system.

AMS Subject Classification: 65N30.

Keywords: Stochastic pde, sparse grids, finite elements, wavelets.

1 Introduction

We consider an elliptic boundary value problem with stochastic input data in a domain D. Namely, let (Ω, Σ, P) be a σ-finite probability space and $D \subset \mathbb{R}^d$ a bounded open set with Lipschitz boundary ∂D. Consider also a deterministic and uniformly positive on D diffusion coefficient $A \in L^\infty(D, \mathbb{R}^{d\times d}_{sym})$. We define a random field on a submanifold M of $\mathbb{R}^d$ (it will always be D or some part of its boundary) as a jointly measurable function from $M \times \Omega$ to $\mathbb{R}$. Suppose $\partial D = \Gamma_0 \cup \Gamma_1$ (disjoint union), where Γ_0 has positive surface measure, and let f, g and h be random fields on D, Γ_0 and Γ_1 respectively. We consider the following model problem

$$\left.\begin{array}{r} L(\partial_x)u \\ \gamma_0(u) \\ \gamma_n(u) \end{array}\right\} = \left\{\begin{array}{r} -\mathrm{div}(A(x)\nabla u(x,\omega)) \\ u(x,\omega)\ |_{\Gamma_0} \\ n^\top A(x)\nabla u(x,\omega)\ |_{\Gamma_1} \end{array}\right\} = \left\{\begin{array}{ll} f(x,\omega) & \text{in } D \\ g(x,\omega) & \text{on } \Gamma_0 \\ h(x,\omega) & \text{on } \Gamma_1 \end{array}\right. , \qquad (1)$$

where the operators involved in the boundary conditions should be thought of as stochastic counterparts of the classical trace on Γ_0, Γ_1 and distributional conormal derivative operators, γ_0, γ_1 and γ_n respectively. Note that if Ω reduces to

* Supported in part under the IHP network *Breaking Complexity* of the EC (contract number HPRN-CT-2002-00286) with support by the Swiss Federal Office for Science and Education under grant No. BBW 02.0418.

only one point of mass one, the dependence of (1) on ω can be dropped, the stochastic character dissapears, and we are left with a classical mixed BVP, which will be refered to in the following as 'deterministic case'. Since for a stochastic problem the data is uncertain and, moreover, knowing all joint probability densities is in practice hardly the case, reasonable assumptions can be made only on some 'statistics' associated to the data. Here we assume that the k-th order moment, sometimes called k-point correlation of the random data $f(x,\omega)$ in (1) and given by

$$\mathcal{M}^k(f)(x_1,\dots,x_k) := \int_\Omega f(x_1,\omega)\cdot f(x_2,\omega)\cdot\ldots\cdot f(x_k,\omega)dP(\omega),$$

$x_j \in D$, $j = 1,2,\dots,k$, whenever such an integral exists, is available. Correspondingly one is often interested in the higher moments of the stochastic solution. We devoted [9] to the theoretical and numerical study of the expectation (that is, the mean field or first order moment) and two-point correlation of the solution. Both these 'statistics' have been shown to satisfy deterministic elliptic problems which are numerically solvable at essentially the same cost (number of operations, memory requirements for a prescribed relative accuracy) as the deterministic mean field problem,

$$\left.\begin{array}{r} L(\partial_x)E_u \\ \gamma_0(E_u) \\ \gamma_n(E_u) \end{array}\right\} = \left\{\begin{array}{r} -\mathrm{div}(A(x)\nabla E_u(x)) \\ E_u(x)\mid_{\Gamma_0} \\ n^\top A(x)\nabla E_u(x)\mid_{\Gamma_1} \end{array}\right\} = \left\{\begin{array}{ll} E_f(x) & \text{in } D \\ E_g(x) & \text{on } \Gamma_0 \\ E_h(x) & \text{on } \Gamma_1 \end{array}\right. . \tag{2}$$

Here the mean field, or expectation, E_u associated to u, solution of (1), is given by

$$E_u(x) := \mathcal{M}^1(u)(x) = \int_\Omega u(x,\omega)dP(\omega), \quad x \in D.$$

We shall study in the present paper existence, regularity, discretization and complexity issues for the k-point correlation of u, the stochastic solution to (1). Our main goal will be to derive and analyze an algorithm that makes these high order statistics available at a computational cost which exhibits only a mild dependence on k.

2 Problem Formulation

Let $k \geq 1$ be an integer, (Ω,Σ,P) a σ-finite probability space and H a separable Hilbert space. We define the Banach space of L^k, H-valued functions on Ω (see [11]) by

$$L^k(\Omega;H) := \left\{ f:\Omega\to H \,|\, f \text{measurable}, \int_\Omega \|f(\omega)\|_H^k dP(\omega) < \infty \right\} / \sim$$

$$\|f\|^k_{L^k(\Omega;H)} := \int_\Omega \|f(\omega)\|_H^k dP(\omega),$$

where we use the same notation for a P-a.e. equivalence (denoted by $\sim$) class and one of its members. Bochner's Theorem (see [11]) asserts that $f \in L^k(\Omega; H)$ if and only if there exists a sequence of H-valued step functions $(f_j)_{j\in\mathbb{N}}$ such that

$$f_j \to f \ P\text{-a.e. on } \Omega \quad \text{and} \quad \int_\Omega \|f_j - f\|_H^k \to 0, \text{ as } j \to \infty. \tag{3}$$

For each $f \in L^1(\Omega; H)$ one can then define the vector-valued integral

$$\int_\Omega f(\omega) dP(\omega) \quad \in H \tag{4}$$

by means of a sequence of H-valued step functions $(f_j)_{j\in\mathbb{N}}$ satisfying (3) for $k = 1$. Namely,

$$\int_\Omega f(\omega) dP(\omega) := \lim_{j\to\infty} \int_\Omega f_j(\omega) dP(\omega), \quad \text{in } H. \tag{5}$$

We shall consider data for (1) satisfying the regularity assumption with $k \geq 2$,

$$\begin{aligned}
&f \in L^k(\Omega; H^{-1}(D)) \subset L^2(\Omega; H^{-1}(D)) \simeq H^{-1}(D) \otimes L^2(\Omega),\\
&g \in L^k(\Omega; H^{1/2}(\Gamma_0)) \subset L^2(\Omega; H^{1/2}(\Gamma_0)) \simeq H^{1/2}(\Gamma_0) \otimes L^2(\Omega),\\
&h \in L^k(\Omega; H^{-1/2}(\Gamma_1)) \subset L^2(\Omega; H^{-1/2}(\Gamma_1)) \simeq H^{-1/2}(\Gamma_1) \otimes L^2(\Omega).
\end{aligned} \tag{6}$$

For any Sobolev space H we denote by $\mathscr{H}$ its stochastic counterpart, that is, the Hilbert space $H \otimes L^2(\Omega)$ (we refer the reader again to [11] for tensor products of Hilbert spaces). We shall use for instance $\mathscr{L}^2(D) := L^2(D) \otimes L^2(\Omega)$, $\mathscr{H}^1_{(0)}(D) := H^1_{(0)}(D) \otimes L^2(\Omega)$ $\mathscr{H}^{1/2}(\Gamma_1) := H^{1/2}(\Gamma_1) \otimes L^2(\Omega)$, etc. We consider also a deterministic diffusion coefficient $A \in L^\infty(D, \mathbb{R}^{d\times d}_{sym})$, uniformly positive on D, i.e.

$$\exists \alpha, \beta > 0 \text{ s.t. } \alpha\|\xi\|^2 \leq \xi^\top A(x)\xi \leq \beta\|\xi\|^2 \quad \forall \xi \in \mathbb{R}^{\mathrm{d}} \text{ and a.e. } x \in D. \tag{7}$$

With this setup one can prove (see [9]) that (1) has a rigorous variational formulation and a unique random solution, as follows. Note that Id stands for the identity operator in $L^2(\Omega)$.

Theorem 2.1. *Assume that f, g, h satisfy* (6). *Then there exists a unique random solution $u \in \mathscr{H}^1(D)$ such that $(\gamma_0 \otimes \mathrm{Id})u = g$ and*

$$\begin{aligned}
\langle (A \otimes \mathrm{Id})(\nabla \otimes \mathrm{Id})u, (\nabla \otimes \mathrm{Id})v \rangle_{\mathscr{L}^2(D)^d} = &\langle f, v \rangle_{\mathscr{H}^{-1}(D), \mathscr{H}^1_{(0)}(D)}\\
&+ \langle h, (\gamma_1 \otimes \mathrm{Id})v \rangle_{\mathscr{H}^{-1/2}(\Gamma_1), \mathscr{H}^{1/2}(\Gamma_1)}
\end{aligned} \tag{8}$$

for all $v \in \mathscr{H}^1_{(0)}(D)$.

Proof. Since $H^1(D)/H^1_{(0)}(D) \simeq H^{1/2}(\Gamma_0)$ as topological spaces, there exists $u_1 \in \mathscr{H}^1(D)$ such that $(\gamma_0 \otimes \mathrm{Id})(u_1) = g$, so that the problem reduces to the existence and uniqueness of $u_0 \in \mathscr{H}^1_{(0)}(D)$ satisfying

$$\begin{aligned}\mathscr{A}(u_0, v) &:= \langle (A \otimes \mathrm{Id})(\nabla \otimes \mathrm{Id})u_0, (\nabla \otimes \mathrm{Id})v\rangle_{\mathscr{L}^2(D)^d} \\ &= -\langle (A \otimes \mathrm{Id})(\nabla \otimes \mathrm{Id})u_1, (\nabla \otimes \mathrm{Id})v\rangle_{\mathscr{L}^2(D)^d} + \langle f, v\rangle_{\mathscr{H}^{-1}(D),\mathscr{H}^1_{(0)}(D)} \\ &\quad + \langle h, (\gamma_1 \otimes \mathrm{Id})v\rangle_{\mathscr{H}^{-1/2}(\Gamma_1),\mathscr{H}^{1/2}(\Gamma_1)} \end{aligned} \tag{9}$$

for all $v \in \mathscr{H}^1_{(0)}(D)$. And this is a simple consequence of Lax-Milgram Lemma in $\mathscr{H}^1_{(0)}(D)$, as soon as we note that, on account of (7), the bilinear form $\mathscr{A}$ defined by the l.h.s. of (9) is bounded and coercive on $\mathscr{H}^1_{(0)}(D)$ ($\|(\nabla \otimes \mathrm{Id}) \cdot \|_{\mathscr{L}^2(D)^d}$ defines a norm on $\mathscr{H}^1_{(0)}(D)$, equivalent to the usual one), while the r.h.s. is a continuous linear functional on the same space. □

Remark 2.2. *Let* $(e)_{i\geq 1}$ *be an ONB in* $L^2(\Omega)$ *and expand* $f = \sum_i f_i \otimes e_i$ *with* $\sum_i \|f_i\|^2_{L^2(D)} \leq \infty$, (*similarly for* g *and* h). *Then the solution* (*in the sense given by Theorem 2.1*) u *to* (8) *is given by* $u = \sum_i u_i \otimes e_i$ *where the series converges absolutely in* $\mathscr{H}^1(D)$ *and the coefficient function* u_i *solves the deterministic mixed BVP*

$$\left.\begin{matrix} L(\partial_x)u_i \\ \gamma_0(u_i) \\ \gamma_n(u_i) \end{matrix}\right\} = \begin{cases} f_i & \text{in } D \\ g_i & \text{on } \Gamma_0\,. \\ h_i & \text{on } \Gamma_1 \end{cases}$$

This can be seen by choosing the test function in (8) of the form $v = w \otimes e_i$, with $w \in H^1_{(0)}(D)$. Note that the deterministic character of A is essential for this decomposition of (1).

Well-posedness of (1) (in the sense given by (8)) being established, we now investigate the existence and the deterministic computation of the k-th order moment of u solution to (1), for $k \geq 2$.

3 Existence and Regularity of Higher Order Moments $\mathscr{M}^{\mathbf{k}}(\mathbf{u})$

We use here the setup and notations of the previous section and assume for simplicity $g = 0$. We deduce next the existence of the higher order moments associated to the pair (f, h). For $\alpha = (\alpha_j)_{1\leq j\leq k} \in \{0, 1\}^k$ and $s > 0$, we define first the deterministic Hilbert spaces $X^{s,\alpha}_{\pm} := \otimes^k_{j=1} X^{s,\alpha_j}_{\pm}$, where $X^{s,1}_{\pm} := H^{s\pm 1}(D), X^{s,0}_{\pm} := H^{s\pm 1/2}(\Gamma_1)$. Consider also the mapping

$$FH : \Omega \to X^{0,\alpha}_{-}, \quad FH(\omega) := \bigotimes_{j=1}^{k} (\alpha_j f + (1 - \alpha_j)h)(\omega). \tag{10}$$

The strong measurability of FH can be deduced by tensorizing sequences of step functions approximating f and h, while the norm integrability is a consequence of

(6) and the Hölder inequality for the pair of functions $\|f(\cdot)\|_{H^{-1}(D)}^{|\alpha|} \in L^{k/|\alpha|}(\Omega)$, $\|h(\cdot)\|_{H^{-1/2}(\Gamma_1)}^{k-|\alpha|} \in L^{k/(k-|\alpha|)}(\Omega)$. This means, in view of (3), $FH \in L^1(\Omega; X_-^{0,\alpha})$. Consequently, $\mathscr{M}^\alpha(f,h)$, the α-moment of the pair (f,h) can be defined according to (4), by

$$\mathscr{M}^\alpha(f,h) := \int_\Omega FHdP(\omega) \in X_-^{0,\alpha}. \tag{11}$$

Note that if $\alpha = (1,1,\dots,1)$, the moment defined by (11) is actually associated to f and not to the pair (f,h), so that from now on it will be denoted by $\mathscr{M}^k(f)$. Similarly, $\alpha = (0,0,\dots,0)$ leads to $\mathscr{M}^k(h)$.

The problem we address next is the existence of the k-th order moment of u. To state the result we use the notations $H^{\mathbf{v}}(D^k) := \otimes_{j=1}^k H^{v_j}(D)$, $H^{\mathbf{v}}_{(0)}(D^k) := \otimes_{j=1}^k H^{v_j}_{(0)}(D)$ for a multi-index $\mathbf{v} \in (\mathbb{R}_+)^k$ and $\mathbf{s} := (s,s,\dots,s) \in (\mathbb{R}_+)^k$ for $s \in \mathbb{R}_+$.

Theorem 3.1. *Under the regularity assumption* (6), *the k-th order moment of u, solution to* (1) *exists and is an element of $H^{\mathbf{1}}(D^k)$.*

Proof. (1) means that, P-a.e. on Ω, $u(\omega)$ solves a deterministic mixed boundary value problem in D, if we view $u \in H^1_{(0)}(D) \otimes L^2(\Omega)$ as a measurable, $H^1_{(0)}(D)$-valued, square norm integrable function on Ω. More precisely, from (8) in Theorem 2.1 we deduce

$$\langle A\nabla u(\omega), \nabla w\rangle_{L^2(D)} = \langle f(\omega), w\rangle_{H^{-1}(D),H^1_{(0)}(D)} + \langle h(\omega), \mathrm{Tr}_1 w\rangle_{H^{-1/2}(\Gamma_1),H^{1/2}(\Gamma_1)},$$

P-a.e. $\omega \in \Omega$ and for all $w \in H^1_{(0)}(D)$. From the well-posedness of the deterministic problem in D it follows that

$$\|u(\omega)\|_{H^1(D)} \le C \cdot \left(\|f(\omega)\|_{H^{-1}(D)} + \|h(\omega)\|_{H^{-1/2}(\Gamma_1)}\right) \quad P\text{-a.e.}\omega \in \Omega, \tag{12}$$

where the constant C depends only on the coefficient A.

Taking into account the measurability of $u : \Omega \to H^1_{(0)}(D)$, which follows from $u \in \mathscr{H}^1_{(0)}(D)$, (12) implies, in view of (6) and the definition of L^k spaces, the assertion. □

To derive a deterministic equation for $\mathscr{M}^k(u)$, we introduce the following operators:

$$A^{\otimes,k} := \bigotimes_{j=1}^k A \in \mathscr{B}(\otimes_{j=1}^k L^2(D)^d)$$

$$\nabla^{\otimes,k} := \bigotimes_{j=1}^k \nabla \in \mathscr{B}(H^{\mathbf{1}}(D^k), \otimes_{j=1}^k L^2(D)^d)$$

$$\gamma_1^{\otimes,\alpha} := \bigotimes_{j=1}^k (\alpha_j \mathrm{Id}_{H^1(D)} + (1-\alpha_j)\gamma_1) \in \mathscr{B}(H^{\mathbf{1}}(D^k), X_+^{0,\alpha}),$$

where we denote by $\mathscr{B}(X,Y)$ the space of bounded linear operators between the Hilbert spaces X and Y, with $\mathscr{B}(X) := \mathscr{B}(X,X)$.

Theorem 3.2. *$\mathscr{M}^k(u)$ is the unique solution in $H^{\mathbf{1}}_{(0)}(D^k)$ of the variational problem*

$$\langle A^{\otimes,k}\nabla^{\otimes,k}\mathscr{M}^k(u), \nabla^{\otimes,k}\mathscr{M}\rangle_{L^2(D)^{dk}} = \sum_{\alpha\in\{0,1\}^k} \langle \mathscr{M}^\alpha(f,h), \gamma_1^{\otimes,\alpha}\mathscr{M}\rangle_{X_-^{0,\alpha},X_+^{0,\alpha}}, \quad \forall \mathscr{M} \in H^{\mathbf{1}}_{(0)}(D^k). \tag{13}$$

Proof. The existence and uniqueness of a solution to (13) are easily proved using the Lax-Milgram Lemma in appropriate spaces, as soon as we note that tensor products of bounded positive homeomorphisms between Hilbert spaces induce corresponding homeomorphisms between tensor products of these spaces.

Now, since $f \in L^k(\Omega, H^{-1}(D)), h \in L^k(\Omega, H^{-1/2}(\Gamma_1))$, there exist sequences $(f_n)_{n\in\mathbb{N}}, (h_n)_{n\in\mathbb{N}}$ of H-valued step functions on Ω satisfying (3) with $H := H^{-1}(D)$ and $H := H^{-1/2}(\Gamma_1)$, respectively. Let us write $f_n = \sum_{q\in J_n} f_n^q 1_{\Omega_{q,n}}$ and $h_n = \sum_{q\in J_n} h_n^q 1_{\Omega_{q,n}}$, where $1_{\Omega_{q,n}}$ stands for the indicator function of the measurable set $\Omega_{q,n}$, $f_n^q \in H^{-1}(D)$, $h_n^q \in H^{-1/2}(\Gamma_1), \forall q, n$, and for each n, the family $(\Omega_{q,n})_{q\in J_n}$ is a partition of Ω. The above mentioned properties of $(f_n)_{n\in\mathbb{N}}, (h_n)_{n\in\mathbb{N}}$ are also sufficient to ensure, via the Hölder inequality, dominated convergence and (5),

$$\lim_{n\to\infty} \mathscr{M}^\alpha(f_n, h_n) = \mathscr{M}^\alpha(f,h) \quad \text{in } X_-^{0,\alpha}. \tag{14}$$

To the deterministic data (f_n^q, h_n^q) we associate the solution $u_n^q \in H^1_{(0)}(D)$ of the corresponding mixed BVP,

$$\langle A\nabla u_n^q, \nabla v\rangle_{L^2(D)^d} = \langle f_n^q, v\rangle_{H^{-1}(D),H^1_{(0)}(D)} + \langle h_n^q, \gamma_1 v\rangle_{H^{-1/2}(\Gamma_1),H^{1/2}(\Gamma_1)} \tag{15}$$

$\forall v \in H^1_{(0)}(D)$, and set $u_n := \sum_{q\in J_n} u_n^q 1_{\Omega_{q,n}}$. The continuous dependence (12) of the solution of a mixed BVP on the data and (3) for f and h imply

$$\lim_{n\to\infty} u_n \to u \quad P\text{-a.e. on } \Omega, \quad \lim_{n\to\infty} \int_\Omega \|u(\omega) - u_n(\omega)\|^k_{H^1_{(0)}(D)} dP(\omega) = 0. \tag{16}$$

Recalling definition (11) of the k-th order moment, we deduce from (16) and (5) that

$$\lim_{n\to\infty} \mathscr{M}^k(u_n) = \mathscr{M}^k(u) \quad \text{in } H^{\mathbf{1}}(D^k). \tag{17}$$

Choosing in (15) k different deterministic test functions $v_1, v_2, \dots, v_k$, taking the product of the resulting k equalities and summing over q with weights $P(\Omega_{q,n})$, we obtain that $\mathscr{M}^k(u_n)$ solves the deterministic problem

$$\langle A^{\otimes,k}\nabla^{\otimes,k}\mathscr{M}^k(u_n), \nabla^{\otimes,k}\mathscr{M}\rangle_{L^2(D)^{dk}} = \sum_{\alpha\in\{0,1\}^k} \langle \mathscr{M}^\alpha(f_n,h_n), \gamma_1^{\otimes,\alpha}\mathscr{M}\rangle_{X_-^{0,\alpha},X_+^{0,\alpha}} \quad \forall \mathscr{M} \in H^{\mathbf{1}}_{(0)}(D^k) \tag{18}$$

(use here that tensor products of total sets in Hilbert spaces are total in product spaces).

The desired equation for $\mathscr{M}^k(u)$ follows then from (14) and (17) if we let $n \to \infty$ in (18). □

The regularity of $\mathscr{M}^k(u)$ follows naturally from that of the data $\mathscr{M}^\alpha(f,h)$, $\forall \alpha \in \{0,1\}^k$ and the result, as well as its proof, is analogous to the one in [9] for $k = 2$. We only state it, as follows. Recall first that the mean field problem (2) is said to satisfy the shift theorem at order $s > 0$ if $E_f \in H^{-1+s}(D)$ implies $E_u \in H^{1+s}(D)$.

Theorem 3.3. *Suppose that the deterministic boundary value problem on D with the diffusion coefficient A satisfies the shift theorem at order s. Then also for* (13) *holds a shift theorem at order s, in the sense that if* $\mathscr{M}^\alpha(f,h) \in X_-^{s,\alpha}, \forall \alpha \in \{0,1\}^k$, *then* $\mathscr{M}^k(u) \in X_+^{s,\mathbf{1}} = \otimes_{j=1}^k H^{s+1}(D)$.

Remark 3.4. *In the case of a polygon or polyhedron D, a shift theorem at order* $s \geq 0$ *holds in weighted spaces* $H_\beta^{1+s,2}(D)$ (*see* [1]). *The proof of Theorem 3.3 can be correspondingly adapted to deduce then a shift theorem for the correlation equation* (13) *in an anisotropic weighted Sobolev scale in* D^k.

4 FE Discretization

We shall now investigate the numerical approximation of $\mathscr{M}^k(u)$, using the Finite Element Method for the deterministic elliptic equation (13). We assume, for simplicity, $\Gamma_1 = \emptyset$ and we start by defining hierarchical FE spaces in D. Let $V_0 \subset V_1 \subset \ldots \subset V_L \subset \ldots \subset H_0^1(D)$ be a dense hierarchical sequence of finite dimensional subspaces of $H_{(0)}^1(D)$, with $N_L := \dim(V_L) < \infty$ for all L. Suppose that the following *approximation property* holds:

$$\min_{v \in V_L} \|u - v\|_{H_0^1(D)} \leq \Phi(N_L, s)\|u\|_{H^{s+1}(D)}, \quad \forall u \in H^{s+1}(D) \cap H_{(0)}^1(D), \tag{19}$$

where $\Phi(N,s) \to 0$ for $s > 0$ as $N \to \infty$ is the convergence rate. For regular solutions the usual FE spaces based on quasiuniform, shape regular meshes are suitable.

Example 4.1. *If* $\{\mathscr{T}^L\}_{L \in \mathbb{N}}$ *is a nested sequence of regular triangulations of D of meshwidth* $h_L = h_{L-1}/2$, *we choose* V_L *to be the space of all continuous piecewise polynomials of degree p on* $\mathscr{T}^L$ *vanishing on* ∂D. *Then* $N_L = O(2^{d \cdot L})$ *and the functional* Φ *on the r.h.s. of* (19) *reads* $\Phi(N,s) = O(N^{-\delta})$, *with* $\delta := \min\{p,s\}/d$.

Since the k-th order moment $\mathscr{M}^k(u)$ of u solves the elliptic problem (13) on D^k, we shall construct FE spaces in D^k, starting from the hierarchical FE spaces $\{V_L\}_{L \geq 0}$ in D. *Full tensor product* spaces present themselves as natural candidates.

However, due to efficiency reasons, we shall use the *sparse tensor product* spaces that are defined by (see [12], [2])

$$\hat{V}_L := \text{Span}\left\{ \bigotimes_{j=1}^{k} V_{i_j} \mid 0 \le i_1 + i_2 + \cdots + i_k \le L \right\}.$$

Since this description of the sparse tensor space does not help identifying bases, we introduce next at each level $L \ge 0$ a *hierarchic excess* W_L of the scale $\{V_L\}_{L\ge 0}$ to be an arbitrary algebraic summand of V_{L-1} in V_L (here we set $V_{-1} := \{0\}$). As V_L can be obviously decomposed as a direct sum $V_L = \bigoplus_{0\le i\le L} W_i$, one can easily check that $\hat{V}_L$ admits the direct (not necessarily orthogonal!) decomposition

$$\hat{V}_L := \bigoplus_{0\le i_1+i_2+\ldots+i_k\le L} \bigotimes_{j=1}^{k} W_{i_j} \subset \bigoplus_{0\le i_1,i_2,\ldots,i_k\le L} \bigotimes_{j=1}^{k} W_{i_j} = \bigotimes_{j=1}^{k} V_L. \tag{20}$$

The discretized version of (13) using the FE space $\hat{V}_L$ then reads

$$\langle A^{\otimes,k}\nabla^{\otimes,k}\mathcal{M}_L^k(u), \nabla^{\otimes,k}\mathcal{M}_L\rangle_{L^2(D)^{dk}} = \langle \mathcal{M}^k(f), \mathcal{M}_L\rangle_{X_-^{0,\mathbf{1}},X_+^{0,\mathbf{1}}}, \tag{21}$$

$\forall \mathcal{M}_L \in \hat{V}_L$, where we denoted by $\mathcal{M}_L^k(u) \in \hat{V}_L$ the discrete solution of (13). The approximation property (19) allows us to estimate the discretization error in terms of the functional Φ, as follows.

Proposition 4.2. *If* $\mathcal{M}_L^k(u)$ *is the solution to* (21), $L \ge k-1$, *and the approximation property* (19) *holds*, *then*

$$\|\mathcal{M}^k(u) - \mathcal{M}_L^k(u)\|^2_{H^{\mathbf{1}}(D^k)} \le C \cdot \sum_{j=1}^{k} c(j,\Phi) \cdot \sum_{\substack{J\subset\{1,\ldots,k\}\\ \text{Card}(J)=j}} \|\mathcal{M}^k(u)\|^2_{H^{s\cdot e_J+\mathbf{1}}(D^k)} \tag{22}$$

where $e_J \in \{0,1\}^k, e_J(j) = 1$ *iff* $j \in J$ *and*

$$c(j,\Phi) = \sum_{m=1}^{j-1} \sum_{l_1+\cdots+l_m=L-m+1} (\Phi(N_{l_1},s)\cdot\Phi(N_{l_2},s)\cdots\Phi(N_{l_m},s)\cdot\Phi(N_0,s))^2 + \sum_{l_1+\cdots+l_j=L-j+1} \Phi(N_{l_1},s)^2\cdot\Phi(N_{l_2},s)^2\cdots\Phi(N_{l_j},s)^2 \tag{23}$$

Note that the constant C *depends only on the coefficient* A.

Proof. As in [9], the result follows using the quasioptimality of the FE solution, the approximation property (19) and the description (20) of the sparse tensor space with W_L defined as the orthogonal complement of V_{L-1} in V_L w.r.t the usual scalar product $\langle\cdot,\cdot\rangle$ in $H_0^1(D)$. Namely, we employ the following orthogonal decomposition in $H_0^{\mathbf{1}}(D^k)$ equipped with the Hilbert structure induced by the tensor product $\bigotimes_{i=1}^k\langle\cdot,\cdot\rangle$. For the rest of the proof, orthogonality in $H_0^{\mathbf{1}}(D^k)$ is to be understood w.r.t. this natural Hilbert structure.

$$\mathscr{M}^k(u) - P_{\hat{\mathbf{S}}_L}(\mathscr{M}^k(u)) = \sum_{\substack{\alpha_1+\alpha_2+\cdots+\alpha_k \geq L+1 \\ \alpha_i \geq 0, 1 \leq i \leq k}} \bigotimes_{i=1}^{k} P^i_{\alpha_i} \mathscr{M}^k(u) \tag{24}$$

where P^i_α denotes the orthogonal projection on W_α w.r.t. $\langle \cdot, \cdot \rangle$, acting in the i-th dimension of D^k. As the notation suggests, $P_{\hat{V}_L}$ denotes the $H^1_0(D^k)$-orthogonal projection on $\hat{V}_L$, while in the following we shall use also Q^i_α, the projection on V_α acting in the i-th direction of D^k. We note that the sum in the r.h.s. of (24) is $H^1_0(D^k)$-orthogonal, since the excesses $W_\alpha, \alpha \in \mathbb{N}$ are pairwise $H^1_0(D)$-orthogonal. We rewrite the r.h.s. of (24), pointing out those directions $j \in \{1, 2, \ldots, k\}$ for which $\alpha_j = 0$ (coarsest approximation). This decomposition does not coincide with the one in [2], but leads to the same qualitative result.

$$\sum_{\substack{\sum_{i=1}^k \alpha_i \geq L+1 \\ \alpha_i \geq 0}} \left(\bigotimes_{i=1}^k P^i_{\alpha_i} \right) \mathscr{M}^k(u) = \sum_{p=1}^{k} \sum_{\substack{J \subset \{1,2,\ldots,k\} \\ \mathrm{Card} J = p}} \sum_{\substack{\sum_{j \in J} \alpha_j \geq L+1 \\ \alpha_j \geq 1}} \left(\bigotimes_{j \in J} P^j_{\alpha_j} \bigotimes_{j \notin J} P^j_0 \right) \mathscr{M}^k(u) \tag{25}$$

and we cast the first inner sum of projections above for $J = \{j_1, j_2, \ldots, j_p\}$ in the form

$$\begin{aligned}
\sum_{\substack{\sum_{n=1}^p \alpha_n \geq L+1 \\ \alpha_n \geq 1}} \left(\bigotimes_{n=1}^{p} P^{j_n}_{\alpha_n} \bigotimes_{j \notin J} P^j_0 \right) &= (\mathrm{Id} - Q^{j_1}_L) \bigotimes_{n \geq 2} \left(\mathrm{Id} - P^{j_n}_0 \right) \bigotimes_{j \notin J} P^j_0 \\
&+ \sum_{\substack{\alpha_1 \leq L \\ \alpha_1 \geq 1}} P^{j_1}_{\alpha_1} \bigotimes (\mathrm{Id} - Q^{j_2}_{L-\alpha_1}) \bigotimes_{n \geq 3} \left(\mathrm{Id} - P^{j_n}_0 \right) \bigotimes_{j \notin J} P^j_0 \\
&+ \sum_{\substack{\alpha_1+\alpha_2 \leq L \\ \alpha_1, \alpha_2 \geq 1}} \bigotimes_{n=1}^{2} P^{j_n}_{\alpha_n} \bigotimes (\mathrm{Id} - Q^{j_3}_{L-\alpha_1-\alpha_2}) \bigotimes_{n \geq 4} \left(\mathrm{Id} - P^{j_n}_0 \right) \bigotimes_{j \notin J} P^j_0 \\
&+ \cdots + \sum_{\substack{\sum_{n=1}^{p-1} \alpha_n \leq L \\ \alpha_n \geq 1}} \bigotimes_{n=1}^{p-1} P^{j_n}_{\alpha_n} \bigotimes \left(\mathrm{Id} - Q^{j_p}_{L - \sum_{n=1}^{p-1} \alpha_n} \right) \bigotimes_{j \notin J} P^j_0.
\end{aligned} \tag{26}$$

We note that the l-th sum in the r.h.s of (26) consists of those terms in the l.h.s. corresponding to indices $\alpha_1, \alpha_2, \ldots \alpha_p \geq 1$ with $\sum_{n=1}^p \alpha_n \geq L+1$ for which $l \leq p$ is the smallest integer with the property $\sum_{n=1}^l \alpha_n \geq L+1$. Using (26) in (25) and the trivial estimate $\|P^i_\alpha\| \leq \|\mathrm{Id} - Q^i_{\alpha-1}\|$ (operator norm in $H^1_0(D^k)$) we easily get, via (19), the desired inequality (22). □

We specialize Proposition (4.2) by choosing the FE spaces as in Example (4.1), to obtain

Corollary 4.3. *For the sparse tensor product based on the FE spaces in Example 4.1 the following asymptotic estimates hold as* $L \to \infty$,

$$\|\mathscr{M}^k(u) - \mathscr{M}^k_L(u)\|_{H^1(D^k)} \le C \cdot (\log N_L)^{(k-1)/2} N_L^{-\delta} \cdot \|\mathscr{M}^k(u)\|_{H^{\mathbf{s}+\mathbf{1}}(D^k)}$$
$$= O((\log N_L)^{(k-1)/2} N_L^{-\delta}), \tag{27}$$

and

$$\dim \hat{V}_L = O((\log N_L)^{k-1} N_L), \tag{28}$$

where $\mathbf{s} = (s, s, \ldots, s)$ *and* $\delta = \min\{p, s\}/d$.

The full tensor space would require $O(N_L^k)$ degrees of freedom for a relative tolerance $O(N_L^{-\delta})$.

Remark 4.4. *The factor* $(\log N_L)^{(k-1)/2}$ *in* (27) *can not be removed. This would be possible, as shown in* [2], *if we were interested in a* $H^1(D^k)$ (*instead of* $H^{\mathbf{1}}(D^k)$) *approximation of the solution. However, since* $H^{\mathbf{1}}(D^k)$ *is the energy space for the k-point correlation problem, an* $H^1(D^k)$*-approximation is in this case irrelevant.*

Remark 4.5. *The proof of the approximation property of the sparse tensor space, on which Proposition 4.2 is based, carries over to a heterogeneous sparse tensor product, in which the factor spaces are possibly different and satisfy each an approximation property of type* (19) (*see* [10]).

5 Iterative Solution and Complexity

We have seen that sparse FE spaces allow to reduce the number of degrees of freedom needed to compute a discrete solution approximating the exact solution up to a prescribed accuracy. To study the complexity of the discrete problem, we recall that (21) amounts to solving a linear system

$$\hat{S}^L \underline{\mathscr{M}}^k(u) = \underline{\mathscr{M}}^k(f), \tag{29}$$

where $\hat{S}^L$ denotes the stiffness matrix of (13) with respect to some basis of the sparse tensor product space $\hat{V}_L \subset H^{\mathbf{1}}_{(0)}(D^k)$. To solve (29) efficiently, we use the conjugate gradient (CG) method, which is suitable once the matrix $\hat{S}^L$ is well-conditioned and sparse. The first property will be ensured by a wavelet preconditioning procedure, while the second, (which does not hold, actually!) can be replaced by a proper use of the anistropic structure of the problem. Here and in what follows, $\mathscr{F}$ denotes a family of double indices running in $\mathbb{N}^d \times \mathbb{N}^d$.

Assumption 5.1. *There exist a family* $(\psi_{j,i})_{(j,i)\in\mathscr{F}} \subset H_0^1(D)$ *and constants* $C_1, C_2 > 0$ *such that each* $u \in H_0^1(D)$ *can be expanded as a convergent series in* $H_0^1(D)$, $u = \sum_{(j,i)\in\mathscr{F}} c_{j,i}\psi_{j,i}$ *and the following 'stability condition' is fulfilled*

$$C_1 \sum_{(j,i)\in\mathscr{F}} |c_{j,i}|^2 \leq \| \sum_{(j,i)\in\mathscr{F}} c_{j,i}\psi_{j,i}\|^2_{H^1_0(D)} \leq C_2 \sum_{(j,i)\in\mathscr{F}} |c_{j,i}|^2. \tag{30}$$

We present some examples of families satisfying Assumption 5.1 for $D = (0,1)$ or $D = (0,1)^d$, but mention that such constructions are available also for polygonal domains (see [6]).

Example 5.2. *For $D = (0,1)$, let us consider ϕ the hat function on $\mathbb{R}$, piecewise linear, taking values $0, 1, 0$ at $0, 1/2, 0$ and vanishing outside $(0,1)$. We set $\mathscr{F} := \{(j,i)|0 \leq j, 1 \leq i \leq 2^j\}$ and $\psi_{j,i}(x) := 2^{-j/2}\phi(2^j x - i + 1), x \in (0,1)$. The family $(\psi_{j,i})_{(j,i)\in\mathscr{F}}$ satisfies then Assumption 5.1.*

Example 5.3. *With D, $\mathscr{F}$ and ϕ as above, we define on $\mathbb{R}$ the function ψ, piecewise linear, taking values $(1, -6, 10, -6, 1)$ at $(1/2, 1, 3/2, 2, 5/2)$ and vanishing outside $(0,3)$. Similarly, ψ^l take $(9, -6, 1)$ at $(1/2, 1, 3/2)$ and ψ^r assumes values $(1, -6, 9)$ at $(1/2, 1, 3/2)$. Further, we define $\psi_{0,1} := \phi$ (scaling function) and $\psi_{j,1}(x) := 2^{-j/2}\psi^l(2^j x)$, $\psi_{j,2^j} := 2^{-j/2}\psi^r(2^j x - 2^j + 1)$, $x \in (0,1)$, for $j \geq 1$ (boundary wavelets). Analogously, $\psi_{j,i}(x) := 2^{-j/2}\psi(2^j x - i + 2), x \in (0,1)$ for $2 \leq i \leq 2^j - 1$ and $j \geq 2$ (interior wavelets). The family $(\psi_{j,i})_{(j,i)\in\mathscr{F}}$ constructed in this way satisfies Assumption 5.1.*

For further examples see [4] and references therein.

Example 5.4. *If $D = (0,1)^d$, we choose $\mathscr{F} := \{(j,i) \in \mathbb{N}^d \times \mathbb{N}^d | 0 \leq j, 1 \leq i \leq 2^j\}$ (inequalities involving multi-indices should be understood componentwise). Then, starting from the family in Example 5.3, we put $\psi_{j,i}(x) = \prod_{q=1}^d \psi_{(j(q),i(q))}(x_q)$ $\forall x = (x_q)_{1\leq q\leq d} \in D$ to obtain (after rescaling) a family $(\psi_{j,i})_{(j,i)\in\mathscr{F}}$ which still satisfies Assumption 5.1 (see [8]).*

Formally, an increasing FE space sequence in $D \subset \mathbb{R}^d$ can be defined in terms of the family $(\psi_{j,i})_{(j,i)\in\mathscr{F}}$ in Assumption 5.1 by

$$V_L := \mathrm{Span}\{\psi_{j,i} | 0 \leq |j|_\infty \leq L\} \tag{31}$$

(j may be a vector, as in the example above, and $|j|_\infty := \max_{1\leq q\leq d} j_q$). We define further an algebraic complement W_L of V_{L-1} in V_L by

$$W_L := \mathrm{Span}\{\psi_{j,i} \| j|_\infty = L\}. \tag{32}$$

We then obtain, via (20), the following explicit description of the sparse tensor space $\hat{V}_L$ through a basis,

$$\hat{V}_L = \mathrm{Span}\left\{ \psi_{\mathbf{j},\mathbf{i}} := \bigotimes_{\nu=1}^k \psi_{\mathbf{j}(\nu),\mathbf{i}(\nu)} \middle| \sum_{\nu=1}^k |\mathbf{j}(\nu)|_\infty \leq L \right\}, \tag{33}$$

where $\mathbf{j}(\nu)$ is the ν-th line of the $k \times d$ matrix $\mathbf{j}$ and similarly for $\mathbf{i}$.

The algebraic excess $\hat{W}_L$ of the sparse tensor scale $(\hat{V}_L)_{L\geq 0}$ is then given by

$$\hat{W}_L = \mathrm{Span}\left\{ \psi_{\mathbf{j},\mathbf{i}} := \bigotimes_{\nu=1}^{k} \psi_{\mathbf{j}(\nu),\mathbf{i}(\nu)} \,\middle|\, \sum_{\nu=1}^{k} |\mathbf{j}(\nu)|_\infty = L \right\}, \tag{34}$$

and can be further decomposed as

$$\hat{W}_L = \bigoplus_{\substack{\underline{l}\in\mathbb{N}^k \\ |\underline{l}|=L}} W_{\underline{l}} \ \text{ with } \ W_{\underline{l}} = \mathrm{Span}\{\psi_{\mathbf{j},\mathbf{i}} \,|\, |\mathbf{j}(\nu)|_\infty = l_\nu\}, \tag{35}$$

where

$$|\underline{l}| := l_1 + l_2 + \cdots + l_k, \ \ \forall \underline{l} \in \mathbb{N}^k.$$

For further reference, let us collect, for $L \geq 0$, in a vector denoted $\mathbf{\Psi}_L$, the basis functions in the definition (32) of W_L. Similarly, for $\underline{l} \in \mathbb{N}^k$ let $\mathbf{\Psi}_{\underline{l}}$ be the vector containing the basis functions of $W_{\underline{l}}$, as defined in (35).

Concerning the properties of the stiffness matrix $\hat{S}^L$ that are of interest for solving (29), namely well-conditioning and sparsity, it holds

Proposition 5.5. i) *The matrix $\hat{S}^L$ has uniformly bounded condition number, as $L \to \infty$.*

ii) *For examples above as well as for similar wavelet constructions, the matrix $\hat{S}^L$ is not sparse, in the sense that* $\mathrm{nnz}(\hat{S}^L) \geq O(N_L^2)$ (*compare* (28)).

Proof. i) (30) can be rephrased by saying that the basis $(\psi_{j,i})_{(j,i)\in\mathscr{F}}$ gives a homeomorphism of Hilbert spaces between ℓ^2 and $H_0^1(D)$, or that

$$u = \sum_{(j,i)\in\mathscr{F}} c_{j,i}\psi_{j,i} \longrightarrow |u|_w^2 := \sum_{(j,i)\in\mathscr{F}} |c_{j,i}|^2 \tag{36}$$

defines an equivalent norm on $H_0^1(D)$. The same holds then for the basis $\psi_{\mathbf{j},\mathbf{i}}$ introduced in (33). It follows that for $\underline{\mathscr{M}} := (\mathscr{M}_{\mathbf{j},\mathbf{i}})_{\mathbf{j},\mathbf{i}} \in \mathbb{R}^{\hat{N}_L}$ with $\hat{N}_L := \dim \hat{V}_L$, $\mathscr{M} := \sum_{\mathbf{j},\mathbf{i}} \mathscr{M}_{\mathbf{j},\mathbf{i}}\psi_{\mathbf{j},\mathbf{i}}$ is an element of $\hat{V}_L$ and

$$\begin{aligned} \langle \hat{S}^L \underline{\mathscr{M}}, \underline{\mathscr{M}} \rangle_{\mathbb{R}^{\hat{N}_L}} &= \langle A^{\otimes,k}\nabla^{\otimes,k}\mathscr{M}, \nabla^{\otimes,k}\mathscr{M}\rangle_{L^2(D)^{dk}} \\ &\sim \|\mathscr{M}\|^2_{H_0^1(D^k)} \sim \sum_{\mathbf{j},\mathbf{i}} |\mathscr{M}_{\mathbf{j},\mathbf{i}}|^2 = \|\underline{\mathscr{M}}\|^2_{\mathbb{R}^{\hat{N}_L}}. \end{aligned}$$

As for ii), one can easily see that the entries of $\hat{S}^L$ corresponding to the indices $\mathbf{i}$, $\mathbf{j}$, $\mathbf{i}'$, $\mathbf{j}'$ with $\mathbf{j}(1) = \mathbf{j}'(2) = (L, L, \ldots, L)$ are in general nonzero, implying the desired lower bound. □

The nonsparsity makes the storage and use of $\hat{S}^L$ rather costly. However, the alternative, that is a full tensor product FE space in D^k, proves already inefficient, for $k \geq 3$, due to its huge dimension N_L^k. A further improvement in the efficiency of solving (29) on a sparse tensor FE space can be achieved (see [9] for the case $k = 2$) by taking into account the special structure of the discrete operator (or, equivalently, of $\hat{S}^L$), which inherits the tensor product structure of the continuous operator (see (13)). More precisely, we shall see that one should store only the matrix S^L corresponding to the case $k = 1$ and relate $\hat{S}^L$ to S^L to perform one step of the CG-algorithm. Of course, storage of the load vector is necessary too, but, due to (28), this requires only a log-linear (in N_L) amount of memory. The Algorithm 6.13 in [9] will be then shown to be applicable to this higher order case to achieve the log-linear complexity of the matrix-vector multiplication needed to perform one step of the CG-algorithm.

We shall derive next the relation between $\hat{S}^L$ and S^L that will help us formulate the matrix-vector multiplication algorithm. To this end, let us denote by $\langle \cdot, \cdot \rangle_w$ the scalar product associated with the norm (36). $\langle \cdot, \cdot \rangle_w$ is obviously equivalent to the usual scalar product in $H_0^1(D)$ and $(\psi_{j,i})_{(j,i)\in\mathscr{F}}$ becomes an orthonormal basis of $H_0^1(D)$ equipped with $\langle \cdot, \cdot \rangle_w$. Let us denote by P_L and Q_L the orthogonal projections in $H_0^1(D)$ w.r.t. $\langle \cdot, \cdot \rangle_w$, on V_L and W_L respectively, as they were defined in (31), (32), so that

$$P_L = \sum_{l=0}^{L} Q_l.$$

Correspondingly, we denote by $\hat{P}_L$ and $\hat{Q}_L$ the orthogonal projections on $\hat{V}_L$ and $\hat{W}_L$ (see (33), (34)) w.r.t. the scalar product on $H_0^{\mathbf{1}}(D^k)$ obtained by tensorizing $\langle \cdot, \cdot \rangle_w$ by itself.

On account of (33), (34), we have the multilevel decomposition

$$\hat{P}_L = \sum_{l=0}^{L} \hat{Q}_l, \tag{37}$$

as well as

$$\hat{Q}_L = \sum_{\substack{\underline{l}\in\mathbb{N}^k \\ |\underline{l}|=L}} Q_{\underline{l}} \;\text{ with }\; Q_{\underline{l}} := \bigotimes_{\nu=1}^{k} Q_{l_\nu} \tag{38}$$

the projection on the space $W_{\underline{l}}$ introduced in (35).

Let us further denote by $\mathscr{Q}^k$ the k-fold tensor product bilinear form of the moment problem (13),

$$\mathscr{Q}^k := \mathscr{Q} \otimes \mathscr{Q} \otimes \cdots \otimes \mathscr{Q}, \;\text{ with }\; \mathscr{Q}(u, v) := \langle A\nabla u, \nabla v\rangle_{L^2(D)}, \tag{39}$$

$\forall u, v \in H_0^1(D)$. Then the discrete problem in $\hat{V}_L$ is given by the bilinear form

$$\mathcal{Q}_L^k(u,v) := \mathcal{Q}^k(\hat{P}_L u, \hat{P}_L v) \quad \forall u, v \in \hat{V}_L \subset H_0^{\mathbf{1}}(D^k) \tag{40}$$

or, inserting (37) and (38) in (40), by

$$\mathcal{Q}_L^k(u,v) = \sum_{l,l'=0}^{L} \sum_{\substack{\underline{l},\underline{l}'\in\mathbb{N}^k \\ |\underline{l}|=l, |\underline{l}'|=l'}} \mathcal{Q}^k\left(Q_{\underline{l}}u, Q_{\underline{l}'}v\right) \;\; \forall u, v \in \hat{V}_L. \tag{41}$$

Recalling that $\Psi_{\underline{l}}$ is the vector containing the basis functions of $W_{\underline{l}}$ given in (35), we can write

$$Q_{\underline{l}}u = u_{\underline{l}}^\top \cdot \Psi_{\underline{l}}, \tag{42}$$

with real vector coefficients $u_{\underline{l}}$ and similarly for v.

Using (42) in (41), we obtain

$$\mathcal{Q}_L^k(u,v) = \sum_{l,l'=0}^{L} \sum_{\substack{\underline{l},\underline{l}'\in\mathbb{N}^k \\ |\underline{l}|=l, |\underline{l}'|=l'}} u_{\underline{l}}^\top \cdot \hat{S}_{\underline{l},\underline{l}'}^L \cdot v_{\underline{l}'}, \tag{43}$$

where the matrix $\hat{S}_{\underline{l},\underline{l}'}^L$ is given by evaluating the bilinear form on the basis functions,

$$\hat{S}_{\underline{l},\underline{l}'}^L := \mathcal{Q}^k(\Psi_{\underline{l}}, \Psi_{\underline{l}'}).$$

But, in view of (39) and (34), we have

$$\hat{S}_{\underline{l},\underline{l}'}^L = \mathcal{Q}^k(\Psi_{\underline{l}}, \Psi_{\underline{l}'}) = \bigotimes_{\nu=1}^{k} \mathcal{Q}(\Psi_{l_\nu}, \Psi_{l'_\nu}) = \bigotimes_{\nu=1}^{k} S_{l_\nu, l'_\nu}^L, \tag{44}$$

where $S_{l,l'}^L := \mathcal{Q}(\Psi_l, \Psi_{l'}), \forall 0 \le l, l' \le L$ are the blocks of the stiffness matrix S^L corresponding to the mean field problem (2) in D (or, equivalently, to the simple case $k = 1$).

The representation formulas (43) and (44) show that

$$\mathcal{Q}_L^k(u,v) = \sum_{l,l'=0}^{L} \sum_{\substack{\underline{l},\underline{l}'\in\mathbb{N}^k \\ |\underline{l}|=l, |\underline{l}'|=l'}} u_{\underline{l}}^\top \cdot \left(\bigotimes_{\nu=1}^{k} S_{l_\nu, l'_\nu}^L\right) \cdot v_{\underline{l}'}, \tag{45}$$

that is the stiffness matrix $\hat{S}^L$ of the k-th moment problem computed w.r.t. the basis (33) of the FE space $\hat{V}_L$ has a block structure

$$\hat{S}^L = (\hat{S}^L_{\underline{l},\underline{l}'})_{\substack{\underline{l},\underline{l}'\in\mathbb{N}^k\\ |\underline{l}|=l\le L\\ |\underline{l}'|=l'\le L}},$$

and each block is a tensor product of certain blocks of the stiffness matrix of the mean field problem, that is, $k = 1$.

Moreover, S^L is almost sparse, once for the basis $(\psi_{j,i})_{(j,i)\in\mathscr{F}}$ the following 'local support' assumption holds true. We remark that the above-mentioned examples as well as similar wavelet-type constructions are in this category.

Assumption 5.6. *There exists $p \in \mathbb{N}^*$ such that for all $1 \le i \le 2^j \in \mathbb{N}^d$ and $j' \in \mathbb{N}^d$, the set $\operatorname{supp}(\psi_{j,i}) \cap \operatorname{supp}(\psi_{j',i'})$ has nonempty interior for at most $p^d \cdot \prod_{q=1}^d \max(1, 2^{j'_q - j_q})$ values of i'.*

Remark 5.7. *From Assumption 5.6 it follows by a simple counting argument that*

$$\operatorname{nnz}(S^L_{l,l'}) \le p^d \cdot (\min(l,l')+1)^{d-1} \cdot 2^{d\cdot\max\{l,l'\}} \quad \forall 0 \le l, l' \le L. \tag{46}$$

To formulate the matrix-vector multiplication algorithm, we shall also need, for each pair $\underline{l} = (l_v)_{v=1}^k, \underline{l}' = (l'_v)_{v=1}^k$, a reordering $\sigma_{\underline{l},\underline{l}'}$ of $\{1, 2, \ldots, k\}$ such that

$$\sum_{v=1}^{q} l_{\sigma(v)} + \sum_{v=q+1}^{k} l'_{\sigma(v)} \le \max\left\{\sum_{v=1}^{k} l_v, \sum_{v=1}^{k} l'_v\right\} \quad \forall 1 \le q \le k. \tag{47}$$

The existence of such a permutation σ is easy to prove, by choosing $x_v = l_v$, $y_v = l'_v$, $\forall 1 \le v \le k$ in the following Lemma.

Lemma 5.8. *If $(x_v)_{1\le v\le k}$ and $(y_v)_{1\le v\le k}$ are two families of positive real numbers, then there exists a permutation σ of the set $\{1, 2, \ldots, k\}$ such that*

$$\sum_{v=1}^{q} x_{\sigma(v)} + \sum_{v=q+1}^{k} y_{\sigma(v)} \le \max\left\{\sum_{v=1}^{k} x_v, \sum_{v=1}^{k} y_v\right\} \quad \forall 1 \le q \le k. \tag{48}$$

Proof. We use induction on k. Since for $k = 1$ the claim is trivial, assume that it holds also for some $k \ge 1$. Consider $(x_v)_{1\le v\le k+1}$ and $(y_v)_{1\le v\le k+1}$ two families of positive real numbers and define $z_v := x_v$ for $1 \le v \le k-1$ and $z_k := x_k + x_{k+1}$, as well as $t_v := y_v$ for $1 \le v \le k-1$ and $t_k := y_k + y_{k+1}$. The induction assumption ensures the existence of a permutation τ of $\{1, 2, \ldots, k\}$ such that $\sum_{v=1}^q z_{\tau(v)} + \sum_{v=q+1}^k t_{\tau(v)} \le \max\left\{\sum_{v=1}^k x_v, \sum_{v=1}^k y_v\right\}$, $\forall 1 \le q \le k$. We define then $\sigma(v) := \tau(v)$ for all $v < \tau^{-1}(k)$ and $\sigma(v) := \tau(v-1)$ for all $v > \tau^{-1}(k)+1$. Now, if $y_k + x_{k+1} \le x_k + y_{k+1}$ holds true, we set $\sigma(\tau^{-1}(k)) := k$, $\sigma(\tau^{-1}(k)+1) := k+1$, otherwise, that is if $y_k + x_{k+1} > x_k + y_{k+1}$, we define $\sigma(\tau^{-1}(k)) := k+1$ and $\sigma(\tau^{-1}(k)+1) := k$. With this choice for σ one can easily check the inequalities (48). □

To simplify the exposition of the algorithm, let us introduce, for an arbitrary pair $(\underline{l}, \underline{l}')$ of indices, $1 \le q \le k$, and a permutation $\sigma = \sigma_{\underline{l},\underline{l}'}$ associated to it in the sense explained above, the following tensor product matrices

$$T^L_{\underline{l},\underline{l}',q} := \bigotimes_{\nu=1}^{k} U_\nu, \tag{49}$$

where

$$U_\nu := \begin{cases} \mathrm{Id}_{l_\nu, l_\nu}, & \nu \in \{\sigma(1), \sigma(2), \ldots, \sigma(q-1)\} \\ S^L_{l_{\sigma(q)}, l'_{\sigma(q)}}, & \nu = \sigma(q), \\ \mathrm{Id}_{l'_\nu, l'_\nu}, & \nu \in \{\sigma(q+1), \sigma(q+2), \ldots, \sigma(k)\} \end{cases} \tag{50}$$

and $\mathrm{Id}_{l,l}$ denotes for $l \ge 0$ the identity matrix of size $\dim W_l$. With these notations, each block in (45) can be expressed as a product of simpler matrices, of the type introduced in (49),

$$\bigotimes_{\nu=1}^{k} S^L_{l_\nu, l'_\nu} = T^L_{\underline{l},\underline{l}',k} \cdot T^L_{\underline{l},\underline{l}',k-1} \cdots T^L_{\underline{l},\underline{l}',1}. \tag{51}$$

For later use, let us remark that (46) entails the following estimate concerning the sparsity of the matrices $T^L_{\underline{l},\underline{l}',q}$.

Remark 5.9.

$$\begin{aligned} \mathrm{nnz}(T^L_{\underline{l},\underline{l}',q}) \lesssim & \prod_{\nu=1}^{q-1} (l_{\sigma(\nu)} + 1)^d \cdot (\min(l_{\sigma(q)}, l'_{\sigma(q)}) + 1)^{d-1} \\ & \times \prod_{\nu=q+1}^{k} (l_{\sigma(\nu)} + 1)^d \cdot 2^{d \cdot \left(\sum_{\nu=1}^{q-1} l_{\sigma(\nu)} + \max\{l_{\sigma(q)}, l'_{\sigma(q)}\} + \sum_{\nu=q+1}^{k} l'_{\sigma(\nu)} \right)} \end{aligned} \tag{52}$$

Proof. This follows immediately from the obvious equality

$$\mathrm{nnz}(T^L_{\underline{l},\underline{l}',\mu}) = \prod_{q=1}^{\mu-1} \dim W_{l_{\sigma(q)}} \cdot \mathrm{nnz}(S^L_{l_{\sigma(\mu)}, l'_{\sigma(\mu)}}) \cdot \prod_{q=\mu+1}^{k} \dim W_{l'_{\sigma(q)}},$$

the asymptotic estimate $\dim W_l \simeq (L+1)^d \cdot 2^{dL}$ and (46). □

Based on the factorization formula (51), we can develop now the multiplication algorithm of the matrix $\hat{S}^L$ by a vector x.

Algorithm 5.10.

store $(S^L_{l,l'})_{0 \le l, l' \le L}$ (sparse), $(x_{\underline{l}})_{l_1 + l_2 + \cdots l_k \le L}$

for $\underline{l}$ satisfying $\sum_{v=1}^{k} l_v \leq L$

initialize $(\hat{S}^L x)_{\underline{l}} := 0$

for $\underline{l}'$ satisfying $\sum_{v=1}^{k} l'_v \leq L$

compute $y_{\underline{l}} := T^L_{\underline{l},\underline{l}',k} \cdot T^L_{\underline{l},\underline{l}',k-1} \cdots T^L_{\underline{l},\underline{l}',1} \cdot x_{\underline{l}'}$

update $(\hat{S}^L x)_{\underline{l}} := (\hat{S}^L x)_{\underline{l}} + y_{\underline{l}}$

end % for

end % for

Remark 5.11. *The order in the multiplication giving* $y_{\underline{l}}$ *is essential for the efficiency of the algorithm. To implement the multiplication of* $T^L_{\underline{l},\underline{l}',q}$ *by a vector, one should not build* $T^L_{\underline{l},\underline{l}',q}$*, but, due to* (49), (50), *split the vector into blocks and multiply each of them by* $S^L_{l_{\sigma(q)},l'_{\sigma(q)}}$.

The estimate of the complexity of Algorithm 5.10 can be carried out as in [9]. The result reads:

Theorem 5.12. *The algorithm* (5.10) *performs the matrix-vector multiplication* $x \to \hat{S}^L x$ *using at most* $O((\log N_L)^{kd+2k-2} N_L)$ *floating point operations. Besides, it requires only storage of the stiffness matrix* S^L *of the mean field problem and of* x.

Proof. Due to (45), (51) we can write

$$(\hat{S}^L x)_{\underline{l}} = \sum_{\substack{\underline{l}' \\ |\underline{l}'| \leq L}} \bigotimes_{v=1}^{k} S^L_{l_v,l'_v} \cdot x_{\underline{l}'} = \sum_{\substack{\underline{l}' \\ |\underline{l}'| \leq L}} T^L_{\underline{l},\underline{l}',k} \cdot T^L_{\underline{l},\underline{l}',k-1} \cdots T^L_{\underline{l},\underline{l}',1} \cdot x_{\underline{l}'}.$$

The multiplication under the summation above can be then performed using at most

$$\#_{\underline{l},\underline{l}'} := \sum_{q=1}^{k} \mathrm{nnz}(T^L_{\underline{l},\underline{l}',q}) \tag{53}$$

floating point operations. From (52) we obtain that

$$\#_{\underline{l},\underline{l}'} \lesssim \sum_{q=1}^{k} (l_{\sigma(1)} + 1)^d \cdots (l_{\sigma(q-1)} + 1)^d \cdot (\min(l_{\sigma(q)}, l'_{\sigma(q)}) + 1)^{d-1}$$
$$\cdot (l'_{\sigma(q+1)} + 1)^d \cdots (l'_{\sigma(k)} + 1)^d \cdot 2^{d \cdot (\sum_{v=1}^{q-1} l_{\sigma(v)} + \max\{l_{\sigma(q)}, l'_{\sigma(q)}\} + \sum_{v=q+1}^{k} l'_{\sigma(v)})}.$$

From this estimate and the defining property (47) of $\sigma = \sigma_{\underline{l},\underline{l}'}$, we deduce that for $L \geq 1$,

$$\#_{\underline{l},\underline{l}'} \lesssim (\max\{|\underline{l}|, |\underline{l}'|\})^{dk-1} \cdot 2^{d \cdot \max\{|\underline{l}|,|\underline{l}'|\}}.$$

Then the computation of the block $(\hat{S}^L x)_{\underline{l}}$ can be done using $\sum_{\underline{l}'} \#_{\underline{l},\underline{l}'}$ operations. Finally, the number of operations needed to perform $x \to \hat{S}^L x$ (collect all blocks $(\hat{S}^L x)_{\underline{l}}$ for all $\underline{l}$) admits the asymptotic upper bound, as $L \to \infty$,

$$\sum_{\substack{\underline{l} \\ |\underline{l}| \leq L}} \sum_{\substack{\underline{l}' \\ |\underline{l}'| \leq L}} (\max\{|\underline{l}|, |\underline{l}'|\})^{dk-1} \cdot 2^{d \cdot \max\{|\underline{l}|,|\underline{l}'|\}}.$$

Since for a given $l \geq 0$ the equation $|\underline{l}| = l$ has exactly $\binom{l+k-1}{k-1} = O(l^{k-1})$ (as $l \to \infty$) solutions $\underline{l} \in \mathbb{N}^k$, we conclude

$$\#\text{flops}(x \to \hat{S}^L x) \lesssim \sum_{l=0}^{L} l^{dk+2k-2} \cdot 2^{dl} = O((\log N_L)^{kd+2k-2} N_L). \qquad \square$$

Due to Proposition 5.5, the number of steps required by the CG algorithm to compute the discrete solution up to a prescribed accuracy is bounded once we use the solution at level $L-1$ as initial guess of the solution at level L. Thus it holds.

Theorem 5.13. *The deterministic problem* (13) *for the k-point correlation function* $\mathscr{M}^k(u) \in H^{\mathbf{s+1}}(D^k) \cap H_0^{\mathbf{1}}(D^k)$ *of the random solution u to* (1) *is numerically solvable at a cost of*

$$O((\log N_L)^{kd+2k-2} N_L) \tag{54}$$

floating point operations, with a

$$O((\log N_L)^{k-1} N_L) \tag{55}$$

needed amount of memory, for a relative accuracy of

$$O((\log N_L)^{(k-1)/2} N_L^{-\delta}), \tag{56}$$

where $\delta = \min\{p, s\}/d$.

Up to the logarithmic terms, the estimates (54), (55), (56) are similar to the ones of the mean field problem (2).

6 Numerical Examples

We present here some elementary one-dimensional examples concerning the 2-point correlation ($D = (-1, 1)$ and $k = 2$ throughout this section) and numerical experiments we have performed in order to validate our main theoretical result, Theorem 5.13. We mention that for the following computations we have used the Riesz basis in Example 5.2 (piecewise linear elements, $p = 1$).

Let us consider first (1) with $g = 0$, $\Gamma_1 = \emptyset$ and a random field $f(x, \omega)$ which is completely uncorrelated, the so-called ‘white-noise’. This amounts formally to

$$\mathcal{M}^2(f) = \delta(x - y) \tag{57}$$

where $\delta(x - y)$ is the Dirac distribution supported on the diagonal in $D \times D$,

$$\langle \delta(x - y), \phi \rangle = \int_D \phi(x, x)dx \quad \forall \phi \in C_0^\infty(D \times D).$$

One can see that the functional $\delta(x - y)$ admits a unique extension to $H^{\mathbf{1/4}+\varepsilon}(D^2)$ $\forall \varepsilon > 0$. It follows, via Theorem (3.3), that the 2-nd moment of u solution to (1) has the following regularity on the anisotropic Sobolev scale:

$$\mathcal{M}^2(u) \in H^{\mathbf{7/4}-\varepsilon}(D^2) \cap H_0^{\mathbf{1}}(D^2). \tag{58}$$

Taking the coefficient A constant, equal to 1, the assumption of Theorem 5.13 holds true with $s = 3/4$. The expected convergence rate (expressed in terms of the number of dofs N) of the discrete solution is therefore $O((\log N)^{5/4} N^{-3/4})$ (compare (55), (56) and note that $N = (\log N_L) N_L$). The observed rate matches the predicted one in Figure 1.

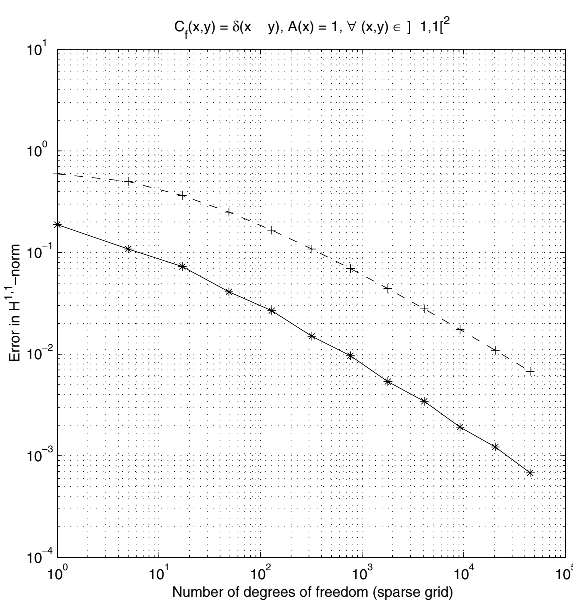

Fig. 1. Convergence in the 1D white-noise case with constant coefficient (solid) and the predicted rate (dashed)

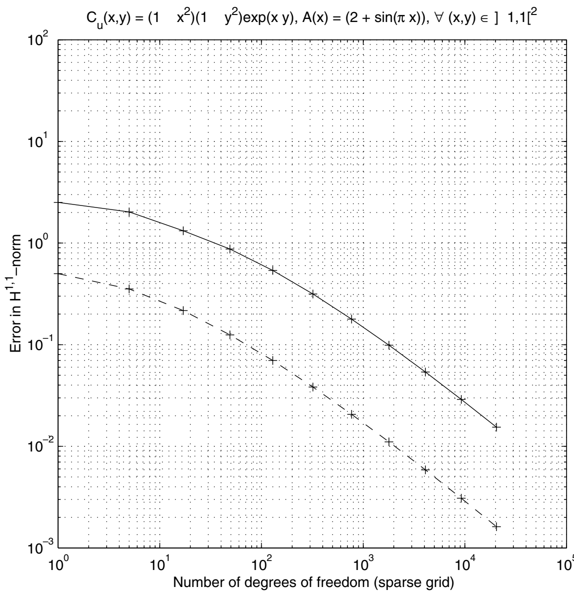

Fig. 2. Convergence in the case of a non-constant coefficient A (solid) and the predicted rate (dashed)

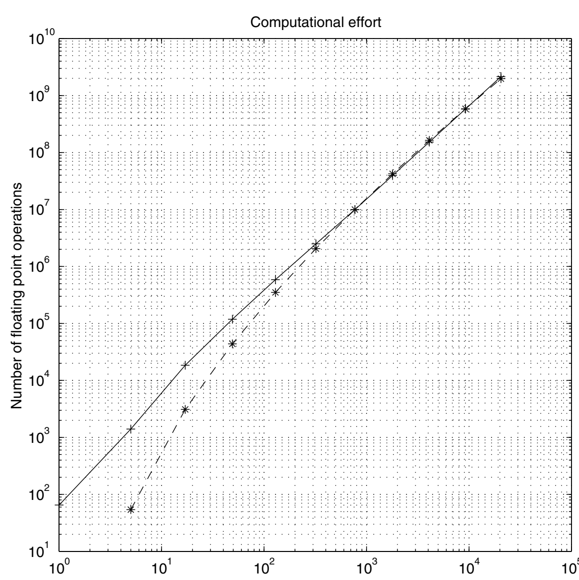

Fig. 3. Comparison between the effort required by the standard CG method based on Algorithm 5.10 (solid) and its theoretical estimate (dashed)

We consider a second example on which we test our complexity estimate (54). Let the coefficient A be given by $A(x) = 2 + \sin(\pi x),\ x \in (-1, 1)$, and the solution to the two-point correlation problem be

$$\mathscr{M}^2(u)(x,y) = (1-x^2)(1-y^2)e^{xy} \in C^\infty(\mathbb{R}^2). \tag{59}$$

A and $\mathscr{M}^2(u)$ being smooth, the assumptions of Theorem 5.13 are satisfied $\forall s > 0$. As a consequence, the expected convergence rate of the discrete solution (again expressed in terms of number of dofs N) is $O((\log N)^{3/2} N^{-1})$. The expected asymptotic behaviour of the computational effort (flops) is $O((\log N_L)^5 N_L)$ for a direct computation of the solution at each level. The observed rates confirm these estimates in Figure 2 and Figure 3.

References

[1] Babuška, I., Guo, B. Q.: Regularity of the solution of elliptic problems with piecewise analytic data. Part I. Boundary value problems for linear elliptic equation of second order. SIAM J. Math. Anal. *19*, 172–203 (1988).

[2] Bungartz, H.-J., Griebel, M.: A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives. Journal of Complexity *15*, 167–199 (1999).

[3] Cohen, A., Daubechies, I., Vial, P.: Wavelets on the interval and fast wavelet transforms. Appl. Comput. Harm. Anal. *1*, 54–81 (1993).

[4] Dahmen, W.: Wavelet and multiscale methods for operator equations. Acta Numerica 55–228 (1997).

[5] Dahmen, W., Schneider, R.: Composite wavelet bases for operator equations. Math. Comp. *68*, 1533–1567 (1999).

[6] Dahmen, W., Stevenson, R.: Element-by-element construction of wavelets satisfying stability and moment conditions. SIAM J. Numer. Anal. *37*, 319–352 (1999).

[7] Griebel, M., Oswald, P., Schiekofer, T.: Sparse grids for boundary integral equations. Numer. Mathematik *83*(2), 279–312 (1999).

[8] Griebel, M., Oswald, P.: Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems. Adv. Comput. Math *4*, 171–206 (1999).

[9] Schwab, Ch., Todor, R. A.: Sparse finite elements for elliptic problems with stochastic data. SAM ETHZ Report 2002–05.

[10] Todor, R. A.: Numerical methods for solving stochastic PDEs. Doctoral Dissertation ETHZ, in preparation.

[11] Yosida, K.: Functional Analysis. Grundlehren, 123, Springer-Verlag, 1964.

[12] Zenger, Ch.: Sparse Grids. "Parallel Algorithms for PDE's ", Proceedings of the 6th GAMM-Seminar, Kiel, 1990.

Christoph Schwab
Seminar for Applied Mathematics ETHZ
Rämistrasse 101
CH-8092 Zürich
Switzerland
e-mail: schwab@math.ethz.ch

Radu Alexandru Todor
Seminar for Applied Mathematics ETHZ
Rämistrasse 101
CH-8092 Zürich
Switzerland
e-mail: todor@math.ethz.ch