Eidgenössische
Technische Hochschule
Zürich

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

hp-FEM for second moments of elliptic PDEs with stochastic data
Part 1: Analytic regularity

B. Pentenrieder and Ch. Schwab

Research Report No. 2010-08
March 2010

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

hp-FEM FOR SECOND MOMENTS OF ELLIPTIC PDES WITH STOCHASTIC DATA PART 1: ANALYTIC REGULARITY

BASTIAN PENTENRIEDER AND CHRISTOPH SCHWAB

Seminar for Applied Mathematics, ETH Zentrum, Zürich, Switzerland

Abstract. For a linear second order elliptic partial differential operator $A : V \to V'$, we consider the boundary value problems $Au = f$ with stationary Gaussian random data f over the dual V' of the separable Hilbert space V in which the solution u is sought. The operator A is assumed to be deterministic and bijective. The unique solution $u = A^{-1}f$ is a Gaussian random field over V. It is characterized by its mean field E_u and its covariance $C_u \in V \otimes V$. For a class of piecewise analytic covariance kernels $C_f \in V' \otimes V'$ for Gaussian data f, we prove analytic regularity of the covariance C_u of the Gaussian solution u in families of countably normed spaces. To this end, we investigate shift theorems for the (non-hypoelliptic) deterministic tensor PDEs $(A \otimes A)C_u = C_f$ proposed in [14] for the covariance C_u. The non-hypoelliptic nature of $A \otimes A$ implies that $\operatorname{sing\,supp}(C_u)$ is in general strictly larger than $\operatorname{sing\,supp}(C_f)$. Based on our regularity results, we outline an hp-Finite Element strategy from [7, 8] to approximate C_u stemming from covariances of stationary Gaussian data f. In the second part [8] of this work, we prove that this discretization gives exponential rates of convergence of the FE approximations, in terms of the number of degrees of freedom.

AMS Subject classification: 65N30

1. Introduction

The efficient numerical solution of partial differential equations (PDEs) with random data has drawn substantial interest in recent years, in particular due to its significance for uncertainy quantification (UQ) in engineering. In such problems, one can roughly distinguish between two types of data uncertainty: uncertainty in loadings, uncertainty in coefficients and even in domains of problem definitions. Whereas in the mathematical theory of such PDEs the focus has been on rather rough noise in data (such as, e.g., white noise) without correlation, in engineering applications the focus is on so-called "colored" noise, where spatial and temporal correlations in input data are estimated from observations.

Computational strategies for this kind of problem can be roughly divided into two classes: sampling methods and parametric, deterministic methods. The former class contains, in particular, Monte Carlo (MC) as well as Quasi-Monte Carlo (QMC) methods. It is based on the computation of a solution "ensemble" based on a suitable data "ensemble" of M i.i.d. draws of the problem data, and amounts to the solution of M many, independent deterministic PDE problems. Deterministic, statistical quantities of engineering interest are then estimated from the solution ensemble by means of statistical estimators such as e.g. sample average or empirical variance estimators. The latter class consists of parametric methods which aim at the deterministic approximation of the law of the unknown random solution. We mention only the so-called *polynomial chaos approach* and its generalizations which

Date: March 3, 2010.

aim at computing a spectral representation of the unknown random fields in terms of polynomials of the (assumed i.i.d) random parameters in the input data. Once again, from the computed probability densities, the computation of moments and other quantities of statistical interest is achieved by numerical integration.

Alternatively, one could attempt to *directly compute statistical moments of the random solution in terms of the corresponding moments of the random data*. This is, in general, not possible due to the nonlinear dependence of the moments of the random solution on the corresponding moments of the random data so that *closure hypotheses* must be imposed which are, to some extent heuristic. Also, a finite number of statistical moments is in general insufficient to determine a random field completely with the notable exception of *Gaussian Random Fields*: these are completely determined by their first and second moments, i.e. by their mean fields and their covariances. In the so-called FOSM[1] ("FOSM" for short) perturbation analysis, explicit and linear PDEs are obtained for the direct computation of the second moments of the random solution which are, of course, only accurate up to terms which are of second and higher order in the perturbation amplitude.

Whereas the mean field corresponds to the solution of a deterministic PDE in the physical domain D of the problem, the second moment or covariance (resp. two-point correlation) is a deterministic quantity which is to be computed in *twice* the number of variables, i.e. in cartesian product domain $\square := D \times D$. For physical domains $D \subset \mathbb{R}^3$, the efficient numerical solution of problems in $\square = D \times D$ is, at first sight, prohibitive due to the so-called *curse of dimensionality*. In [14, 15] and in [8], however, the use of *sparse tensor products* of multilevel Finite Element Spaces in the domain D has been shown to be able to avoid the complexity increase in the deterministic covariance computations due to the doubling of the dimension of the computational domain. Deterministic second moment solvers which scale log-linearly in the number N of degrees of freedom for the solution of the mean field problem have been given in [14, 15]. The reduction of complexity is, however, achieved only for rather smooth covariance functions (belonging to so-called "spaces of mixed highest derivatives"). This assumption is not realistic: two-point spatial correlation functions widely used in statistical modeling do exhibit singular supports on the diagonal of the domain $D \times D$. This implies only rather limited regularity in scales of Sobolev spaces of mixed dominating derivative and, accordingly, poor approximation properties of the sparse tensor product approximations in [14, 15] of such functions.

To exhibit a new, tensorized *hp*-Finite Element approximation scheme for such covariance functions for an elliptic model problem in one spatial dimension is the purpose of the present paper and the related reference [8]. We prove in [8] that it converges exponentially, for *covariance functions with singular support on the diagonal* $\Delta := \{(x, y) : x, y \in D, x = y\} \subset \square$ *which are analytic in* $\overline{D \times D} \backslash \Delta$. Key ingredient in achieving this is the assumption of *stationarity* of the random input data f. We present a Finite Element approximation algorithm for the computation of the solution which only requires *hp*-discretization of the (deterministic) elliptic differential operator. We state an exponential convergence result which will be proved in [8]. It is based on an analytic regularity theory for the tensor differential operator $A \otimes A$ where A is a second order differential operator acting in a one dimensional domain D. The presentation of this regularity theory is the principal purpose of the present paper. It is structured as follows: in Section 2, we present the class of PDEs of interest and define, in particular, Gaussian data f and Gaussian solution fields u, and derive tensorized equations of the k-point correlation functions of u. Section 3 introduces a specific class of elliptic problems in the unit interval

[1] First Order, Second Moment

$D = (0,1)$ and the associated mean and covariance equations to be investigated in the following. Several classes of parametric, analytic data covariance functions C_f which are frequently used in practice are also presented. We also introduce the tool of countably normed, weighted Sobolev spaces in order to quantify the analytic regularity of these covariances. Section 4 is devoted to a detailed description of the regularity of the solution covariance C_u, for the data covariance classes introduced in Section 3, in particular to the classification of the singular supports of C_u and its regularity in terms of the countably normed spaces. Section 5 discusses the behaviour of the analytic regularity bounds for the covariance C_u in the case of *small correlation length*, i.e. the (co)variance of u is concentrated near the diagonal Δ of $\Box = D \times D$. Section 6 outlines the *hp*-Galerkin Finite Element approximation of the solution's covariance C_u by a *hp*-Finite Element approximation in $D \times D$ from [8] designed to resolve the singular support of C_u while maintaining exponential convergence.

2. Elliptic linear operator equations with Gaussian data

In this section, we provide the general problem framework within which the present paper resides. At first, the notion of a Gaussian measure is introduced leading us to the definition of Gaussian random fields. Then, we will admit these ones as right hand side to a linear operator equation and state properties of the corresponding solution. Finally, the last subsection gives a definition of the statistical moments associated with a random field and supplies deterministic equations that allow to compute the moments of the solution from the moments of the right hand side. The presentation of the material follows [3, 13].

2.1. **Gaussian random fields.** Throughout this work, $(\Omega, \mathcal{F}, \mathbb{P})$ shall denote a generic probability space.

Definition 2.1 (Borel algebra $\mathfrak{B}(V)$)**.** Let V be any complete metric space. Then, we define $\mathfrak{B}(V)$ to be the σ-algebra which is generated by the open subsets of V, i.e. $\mathfrak{B}(V)$ is the smallest σ-algebra containing all open subsets of V.

With $\mathfrak{B}(V)$ at hand, the generalization of the standard definition of $\mathbb{R}$-valued random variables is straightforward:

Definition 2.2 (V-valued random variable in $(\Omega, \mathcal{F}, \mathbb{P})$)**.** Let V be a complete metric space. A V-valued random variable in $(\Omega, \mathcal{F}, \mathbb{P})$ is a mapping $X : \Omega \to V$ which is $\mathcal{F}$-$\mathfrak{B}(V)$-measurable:

$$I \in \mathfrak{B}(V) \quad \Rightarrow \quad X^{-1}(I) \in \mathcal{F}$$

Remark 2.3 (random vector, stochastic process, random field). If $V = \mathbb{N}^d$, $\mathbb{Z}^d$, $\mathbb{R}^d$, etc. with $d > 1$, it is common to call the random variable X a *random vector*, since the values that are assumed by X are d-dimensional vectors. Analogously, in case of V being a function space such as $C^0([0,T])$ or $L^2(D)$ with a domain $D \subset \mathbb{R}^d$, $d \geq 1$, the terms *stochastic process* and *random field* are often used—depending on whether the argument of the functions in V is a time or space variable.

Definition 2.4 (law/distribution of X)**.** Let X be a V-valued random variable in $(\Omega, \mathcal{F}, \mathbb{P})$. Then, the law (or distribution) of X is the measure $X_\#\mathbb{P}$ on $(V, \mathfrak{B}(V))$ given by

$$X_\#\mathbb{P}(A) := \mathbb{P}(X^{-1}(A)), \quad A \in \mathfrak{B}(V).$$

In particular, $X_\#\mathbb{P}(V) = \mathbb{P}(\Omega) = 1$ so that $(V, \mathfrak{B}(V), X_\#\mathbb{P})$ is a probability space.

Since they require only the notion of an open set, the above definitions are meaningful whenever V is equipped with a metric. However, in the remainder of this work, we will focus on the particular case where V is a separable Hilbert space. We consider random data and solutions by introducing probability measures on the pair $(V, \mathfrak{B}(V))$ which are *Gaussian*, i.e. we consider data and solutions which are *Gaussian random fields*. In the following, we will mostly speak of random fields instead of random variables, since the spaces V we have in mind are function spaces. The norm and inner product in V are written as $\|\cdot\|_V$ and $\langle\cdot,\cdot\rangle$, respectively.

Definition 2.5 (operator sets $L(V)$, $L^+(V)$ and $L_1^+(V)$; trace class)**.** Let V be a separable Hilbert space. By $L(V)$, we mean the Banach space of continuous linear operators from V into itself. Furthermore, $L^+(V)$ shall be the set of all symmetric and non-negative $T \in L(V)$,

$$\begin{aligned} &\langle Tx, y\rangle = \langle x, Ty\rangle \quad && \forall x, y \in V, \\ &\langle Tx, x\rangle \geq 0 && \forall x \in V. \end{aligned}$$

Finally, we denote by $L_1^+(V)$ the set of all operators $Q \in L^+(V)$ of trace class, i.e. of all $Q \in L^+(V)$ for which

$$\operatorname{Tr} Q := \sum_{k=1}^{\infty} \langle Qe_k, e_k\rangle < \infty$$

holds for a complete orthonormal system $(e_k)_{k\in\mathbb{N}}$ in V. (Notice that $\operatorname{Tr} Q$ equals the sum of Q's eigenvalues repeated according to their multiplicity.)

In the following, we give definitions for mean and covariance of a *probability measure* on the measurable space $(V, \mathfrak{B}(V))$, regardless whether this one is the law of some random field $X : \Omega \to V$ or not. The relation of these quantities to the so-called mean-field and covariance kernel of a *random field* will be established in Remark 2.16.

Definition 2.6 (mean of a measure)**.** Let μ be a probability measure on $(V, \mathfrak{B}(V))$ such that

$$\int_V \|x\|_V \, \mu(dx) < \infty.$$

Then,

$$v \mapsto \int_V \langle x, v\rangle \, \mu(dx)$$

is a continuous linear functional from V into $\mathbb{R}$. Thus, according to the Riesz representation theorem, there exists a unique $a \in V$ such that

$$\int_V \langle x, v\rangle \, \mu(dx) = \langle a, v\rangle \quad \forall v \in V.$$

a is called the mean of μ. We will write

$$a = \int_V x \, \mu(dx).$$

Definition 2.7 (covariance of a measure)**.** Let μ be a probability measure on $(V, \mathfrak{B}(V))$ with mean a and

$$\int_V \|x\|_V^2 \, \mu(dx) < \infty.$$

Then,

$$(v_1, v_2) \mapsto \int_V \langle v_1, x - a\rangle\langle v_2, x - a\rangle \, \mu(dx)$$

defines a continuous bilinear form on $V \times V$. Thus, again by the Riesz theorem, for every fixed $v_1 \in V$, there exists a unique $Qv_1 \in V$ such that

$$\int_V \langle v_1, x-a\rangle\langle v_2, x-a\rangle\, \mu(dx) = \langle Qv_1, v_2\rangle$$

for all $v_2 \in V$. The mapping $v \mapsto Qv$ is continuous and linear, i.e. $Q \in L(V)$. We call Q the covariance or covariance operator of μ.

Proposition 2.8. *Let μ be a probability measure on $(V, \mathfrak{B}(V))$ with mean a and covariance operator Q. Then, $Q \in L_1^+(V)$, i.e. Q is symmetric, non-negative and of trace class.*

Proof. See e.g. [3, Proposition 1.8]. □

In general, there will be many different probability measures having the same mean and covariance. In order for one of these to be Gaussian, it must possess a particular characteristic function:

Definition 2.9 (Gaussian measure)**.** Let V be a separable Hilbert space, $a \in V$ and $Q \in L_1^+(V)$. A Gaussian measure $N_{a,Q}$ on $(V, \mathfrak{B}(V))$ is a probability measure with mean a, covariance operator Q and characteristic function

$$\widehat{N_{a,Q}}(v) := \int_V e^{i\langle v,x\rangle} N_{a,Q}(dx) = e^{i\langle a,v\rangle - \frac{1}{2}\langle Qv,v\rangle} \quad \forall v \in V.$$

Theorem 2.10. *For every pair $(a, Q) \in V \times L_1^+(V)$, there exists a unique Gaussian measure with mean a and covariance operator Q.*

Proof. See e.g. [3, Theorem 1.12]. □

Definition 2.11 (Gaussian random field)**.** A V-valued random field X in $(\Omega, \mathcal{F}, \mathbb{P})$ is called Gaussian, if its law $X_\#\mathbb{P}$ is a Gaussian measure on $(V, \mathfrak{B}(V))$.

Lemma 2.12. *Let $X : \Omega \to V$ be a Gaussian random field with measure $\mu = N_{a,Q}$ on $(V, \mathfrak{B}(V))$. Furthermore, let $T \in L(V, K)$ be a continuous linear mapping from V into a Hilbert space K. Then, $T \circ X$ is a K-valued Gaussian random field with law $T_\#\mu = N_{Ta,TQT^*}$, where T^* denotes the transpose of T.*

Proof. See e.g. [3, Proposition 1.18]. □

2.2. **Linear operator equations.** Let V be a separable Hilbert space with its dual V'. Furthermore, let $A \in L(V, V')$ be a bounded linear operator from V into V' with associated bilinear form

$$a : V \times V \to \mathbb{R}, \qquad a(w, v) := {}_{V'}\langle Aw, v\rangle_V,$$

where ${}_{V'}\langle \cdot, \cdot\rangle_V$ denotes the standard duality pairing. We consider the problem: Given an $f \in V'$,

$$\text{find } u \in V \text{ such that} \quad a(u, v) = {}_{V'}\langle f, v\rangle_V \quad \forall v \in V. \tag{2.1}$$

By the Lax–Milgram lemma, the existence of constants $K < \infty$, $\gamma > 0$ with

$$|a(w, v)| \le K\|w\|_V\|v\|_V \qquad \forall w, v \in V, \tag{2.2}$$

$$a(v, v) \ge \gamma\, \|v\|_V^2 \qquad \forall v \in V, \tag{2.3}$$

guarantees that $A \in L(V, V')$ is boundedly invertible and that the operator norm of the inverse A^{-1} is not larger than γ^{-1}, i.e.: For every $f \in V'$, problem (2.1) has a unique solution $u = A^{-1}f \in V$ satisfying the a priori estimate

$$\|u\|_V = \|A^{-1}f\|_V \le \frac{1}{\gamma}\, \|f\|_{V'}. \tag{2.4}$$

We are interested in operator equations of the form $Au(\omega) = f(\omega)$, where the right hand side is a random field $f : \Omega \to V'$. Suitable function spaces for random data and solution are provided by the following generalization of the classical L^p-spaces:

Definition 2.13 (Bochner spaces $L^k(\Omega; H)$)**.** Let $k \geq 1$. For any separable Hilbert space H, we denote by $L^k(\Omega, \mathcal{F}, \mathbb{P}; H)$ the Banach space of all random fields $X : \Omega \to H$ for which

$$\|X\|_{L^k(\Omega,\mathcal{F},\mathbb{P};H)} := \left(\int_\Omega \|X(\omega)\|_H^k \, \mathbb{P}(d\omega) \right)^{\frac{1}{k}}$$

is finite. In particular, the Hölder inequality implies the inclusion

$$L^m(\Omega, \mathcal{F}, \mathbb{P}; H) \subset L^k(\Omega, \mathcal{F}, \mathbb{P}; H) \qquad \forall m \geq k.$$

If $k = 2$, $L^k(\Omega, \mathcal{F}, \mathbb{P}; H)$ is a Hilbert space. As an abbreviation for $L^k(\Omega, \mathcal{F}, \mathbb{P}; H)$, we shall write $L^k(\Omega; H)$.

For the important special case of Gaussian random data $f : \Omega \to V'$, the application of Lemma 2.12 with operator $T = A^{-1} \in L(V', V)$ implies that the solution $u : \Omega \to V$, $u(\omega) \mapsto (A^{-1} \circ f)(\omega)$, is again Gaussian:

Theorem 2.14. *Let $f \in L^2(\Omega; V')$ be a Gaussian random field with mean $a_f \in V'$ and covariance operator $Q_f \in L_1^+(V')$. If (2.2) and (2.3) hold, the equation*

$$Au = f \quad \text{in } L^2(\Omega; V')$$

admits a unique solution $u \in L^2(\Omega; V)$ which is Gaussian as well. In particular, its mean and covariance are given by

$$a_u = A^{-1} a_f \in V, \qquad Q_u = A^{-1} Q_f (A^{-1})^* \in L_1^+(V).$$

2.3. **Deterministic moment equations.** This subsection investigates the problem $Au = f$ with $A \in L(V, V')$ and a general (not necessarily Gaussian) random field $f \in L^k(\Omega; V')$, $k \geq 1$. If (2.2) and (2.3) hold for the bilinear form associated with A, then $u \in L^k(\Omega; V)$ since we have estimate (2.4) for $\mathbb{P}$-a.e. $\omega \in \Omega$:

$$\|u\|_{L^k(\Omega;V)}^k = \int_\Omega \|u(\omega)\|_V^k \, \mathbb{P}(d\omega) \leq \gamma^{-k} \int_\Omega \|f(\omega)\|_{V'}^k \, \mathbb{P}(d\omega) = \gamma^{-k} \|f\|_{L^k(\Omega;V')}^k$$

Typically, one deals with statistical information on the random fields $X : \Omega \to H$, especially with their moments (here, X represents u or f, and H stands for V or V', respectively). The moments of order k of a random field taking values in H are sometimes also referred to as k-point correlations. They are deterministic quantities taking values in the k-fold tensor product spaces

$$H^{(k)} = \underbrace{H \otimes \cdots \otimes H}_{k \text{ times}}. \tag{2.5}$$

The natural norm of $H^{(k)}$ is denoted by $\|\cdot\|_{H^{(k)}}$. It satisfies

$$\|v_1 \otimes \cdots \otimes v_k\|_{H^{(k)}} = \|v_1\|_H \cdots \|v_k\|_H \qquad \forall v_1, \ldots, v_k \in H. \tag{2.6}$$

(For tensor products of Hilbert spaces and related norms, see e.g. [10, Chapter 2.4].) Now, let $X^{(k)} : \Omega \to H^{(k)}$ be defined by $X^{(k)}(\omega) := X(\omega) \otimes \cdots \otimes X(\omega)$. Then, property (2.6) ensures $X^{(k)} \in L^1(\Omega; H^{(k)})$:

$$\begin{aligned} \|X^{(k)}\|_{L^1(\Omega;H^{(k)})} &= \int_\Omega \|X(\omega) \otimes \cdots \otimes X(\omega)\|_{H^{(k)}} \, \mathbb{P}(d\omega) \\ &= \int_\Omega \|X(\omega)\|_H \cdots \|X(\omega)\|_H \, \mathbb{P}(d\omega) = \|X\|_{L^k(\Omega;H)}^k \end{aligned}$$

In particular, this implies that the expectation of $X^{(k)} : \Omega \to H^{(k)}$ (i.e. the integral of $X^{(k)}$ over the sample space Ω) is a well-defined quantity in the space $H^{(k)}$:

Definition 2.15 (statistical moments)**.** Let $X \in L^k(\Omega; H)$ with $k \in \mathbb{N}$. By the k-th moment of the random field X, we mean the expectation of $X^{(k)}$:

$$\mathcal{M}^k X := \int_\Omega X^{(k)}(\omega)\,\mathbb{P}(d\omega) = \int_\Omega X(\omega) \otimes \cdots \otimes X(\omega)\,\mathbb{P}(d\omega) \in H^{(k)}$$

Remark 2.16 (connection between moments and mean/covariance). For every random field $X \in L^1(\Omega; H)$, we have the integral transformation

$$\mathcal{M}^1 X = \int_\Omega X(\omega)\,\mathbb{P}(d\omega) = \int_H v\, X_{\#}\mathbb{P}(dv).$$

This equation in the space H shows that the first moment of X is identical with the mean of its law $X_{\#}\mathbb{P}$ (cf. Definition 2.6). If $X \in L^2(\Omega; H)$, then $\mathcal{M}^2 X - \mathcal{M}^1 X \otimes \mathcal{M}^1 X \in H \otimes H$, and

$$(v_1, v_2) \mapsto \left\langle \mathcal{M}^2 X - \mathcal{M}^1 X \otimes \mathcal{M}^1 X, v_1 \otimes v_2 \right\rangle_{H \otimes H}$$

defines a continuous bilinear form on $H \times H$ (here, $\langle \cdot, \cdot \rangle_{H\otimes H}$ denotes the inner product on the space $H \otimes H$). One can show that this bilinear form is identical with the one in Definition 2.7. For this reason, we have

$$\left\langle \mathcal{M}^2 X - \mathcal{M}^1 X \otimes \mathcal{M}^1 X, v_1 \otimes v_2 \right\rangle_{H \otimes H} = \langle Q v_1, v_2 \rangle \qquad \forall v_1, v_2 \in H,$$

where $Q \in L_1^+(H)$ is the covariance operator of the law $X_{\#}\mathbb{P}$. It is common to call $\mathcal{M}^1 X$ the *mean-field* and $\mathcal{M}^2 X - \mathcal{M}^1 X \otimes \mathcal{M}^1 X$ the *covariance kernel* of the random field X. In the next section, we will give a concrete example for the representation of the covariance operator by means of its kernel.

Remark 2.17. If $f \in L^2(\Omega; V')$ is a Gaussian random field, the solution $u \in L^2(\Omega; V)$ is guaranteed to be Gaussian as well by Theorem 2.14. In this case, the law of u is uniquely determined by its mean and covariance (cf. Definition 2.9 and Theorem 2.10). Hence, according to Remark 2.16, it is sufficient to compute only the first two moments of u.

In order to derive an equation for $\mathcal{M}^k u$, we consider the k-fold tensor product of $A \in L(V, V')$:

$$A^{(k)} := \underbrace{A \otimes \cdots \otimes A}_{k \text{ times}}$$

$A^{(k)}$ is a mapping from $V^{(k)}$ into $(V')^{(k)}$. Thus, the tensorized problem corresponding to

$$Au = f \qquad \text{in } L^k(\Omega; V')$$

reads

$$A^{(k)} u^{(k)} = f^{(k)} \quad \text{in } L^1(\Omega; (V')^{(k)}).$$

Integrating both sides over Ω yields the *deterministic k-th moment equation*

$$\int_\Omega A^{(k)} u^{(k)}(\omega)\,\mathbb{P}(d\omega) = \int_\Omega f^{(k)}(\omega)\,\mathbb{P}(d\omega) \qquad \text{in } (V')^{(k)}\ .$$

Since $A^{(k)}$ is a continuous linear operator, it commutes with the integration. For this reason, the above identity is equivalent to the operator equation

$$A^{(k)} \mathcal{M}^k u = \mathcal{M}^k f. \tag{2.7}$$

With (2.7), we have finally obtained an equation that allows to compute the k-th moment of u from the k-th moment of f in a fully deterministic way (see also [16]).

3. Problem setting

3.1. **Stochastic model equation.** Let $D = [0, 1]$ be the unit interval and $b > 0$ a constant. We consider the stochastic model equation

$$\left.\begin{aligned} Au(\omega) = -u_{xx}(\cdot,\omega) + b^2 u(\cdot,\omega) = f(\cdot,\omega) \quad &\text{in } (H^1(D))' \\ u_x(0,\omega) = u_x(1,\omega) = 0 \quad & \end{aligned}\right\} \quad \text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega, \tag{3.1}$$

where the load term f and therefore the solution u randomly depend on $\omega \in \Omega$. The functions u and f are construed as random fields in $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $H^1(D)$ and its dual $(H^1(D))'$, respectively (cf. Section 2). The bilinear form $a : H^1(D) \times H^1(D) \to \mathbb{R}$ associated with the linear operator equation (3.1) is given by

$$a(w, v) = \int_D \big(w'(x)\, v'(x) + b^2 w(x)\, v(x)\big)\, dx. \tag{3.2}$$

For the bilinear form a, we shall also use the notation

$$a(w, v) = \int_D \left\langle \begin{pmatrix} d/dx \\ b^2 \,\mathrm{id} \end{pmatrix} w, \begin{pmatrix} d/dx \\ \mathrm{id} \end{pmatrix} v \right\rangle dx.$$

The bilinear form $a : H^1(D) \times H^1(D) \to \mathbb{R}$ satisfies (2.2), (2.3) with continuity constant $K_b = 1 + b^2$ and coercivity constant $\gamma_b = \min\{1, b^2\} > 0$. Thus, all results from Sections 2.2 and 2.3 apply.

Assumption 3.1. We assume f to be a Gaussian random field, and let $\mathcal{M}^1 f = 0$ without loss of generality. Then, equation (2.7) immediately implies $\mathcal{M}^1 u = 0$, and due to Remark 2.17, the law of u is completely determined by $\mathcal{M}^2 u$.

In the context of the model problem, the Hilbert space V from Sections 2.2 and 2.3 is given by $H^1(D)$. Correspondingly, the following definition is helpful with regard to the tensor product spaces $V^{(k)}$:

Definition 3.2 (tensor product Sobolev spaces)**.** Let $k \in \mathbb{N}$ and $\boldsymbol{\nu} = (\nu, \ldots, \nu)$ be a k-tuple of non-negative integers ν. Then, the tensor product Sobolev space of order $\boldsymbol{\nu}$ on $D^k = D \times \cdots \times D$ is defined by

$$H^{\boldsymbol{\nu}}(D^k) := H^\nu(D) \otimes \cdots \otimes H^\nu(D).$$

The space $H^{\boldsymbol{\nu}}(D^k)$ can be characterized as the set of all functions $g : D^k \to \mathbb{R}$ for which

$$\big\|\partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \cdots \partial_{x_k}^{\alpha_k} g\big\|_{L^2(D^k)} < \infty, \quad \text{if } \max\big\{\alpha_1, \alpha_2, \ldots, \alpha_k\big\} \le \nu.$$

Notice that–in comparison to the classical Sobolev spaces–we have the inclusions:

$$H^{\boldsymbol{\nu}}(D^k) \subset H^l(D^k), \text{ if } \nu \ge l \qquad \text{and} \qquad H^m(D^k) \subset H^{\boldsymbol{\nu}}(D^k), \text{ if } m \ge k\,\nu$$

With $\nu = 1$, the spaces $H^{\boldsymbol{\nu}}(D^k)$ allow us to state the variational equivalent to the moment equations (2.7):

$$\begin{aligned} \int_{D^k} \left\langle \begin{pmatrix} d/dx_1 \\ b^2 \,\mathrm{id} \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} d/dx_k \\ b^2 \,\mathrm{id} \end{pmatrix} \mathcal{M}^k u, \begin{pmatrix} d/dx_1 \\ \mathrm{id} \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} d/dx_k \\ \mathrm{id} \end{pmatrix} \mathcal{M} \right\rangle d\mathbf{x} \\ = \int_{D^k} \mathcal{M}^k f\, \mathcal{M}\, d\mathbf{x} \qquad \forall \mathcal{M} \in H^{1,\ldots,1}(D^k) \end{aligned} \tag{3.3}$$

The formulation (3.3) supplies a deterministic way for computing the moments of the solution u from the moments of the right hand side f (independent of f being Gaussian or not). Eliminating randomness in this fashion comes at a price, however: whereas the stochastic problem (3.1) was posed on the physical domain D, the dimension of the domain in the k-th moment equation is the hypercube D^k. Nevertheless, at least in principle, approach (3.3) works for arbitrary k as long as $f \in L^k(\Omega; V')$ (cf. Section 2.3). For Gaussian data and for a FOSM[2] perturbation analysis of nonlinear problems, only first and second moments, the so-called mean-field and the 2-point correlation, of random fields are of interest.

Definition 3.3 (mean-field, 2-point correlation/correlation kernel)**.** For $k = 1$ and $k = 2$ in Definition 2.15, we set $E_u := \mathcal{M}^1 u$ and $C_u := \mathcal{M}^2 u$, respectively.

$$E_u \in H^1(D), \qquad E_u(x) = \int_\Omega u(x,\omega)\, \mathbb{P}(d\omega) \tag{3.4}$$

is called mean-field of u.

$$C_u \in H^{1,1}(D^2), \qquad C_u(x,y) = \int_\Omega u(x,\omega)\, u(y,\omega)\, \mathbb{P}(d\omega) \tag{3.5}$$

is called 2-point correlation or correlation kernel of u.

By Assumption 3.1, the mean-field E_u of the solution to our model problem (3.1) is zero. For this reason, the correlation kernel C_u coincides with the covariance kernel $C_u - E_u \otimes E_u$ mentioned in Remark 2.16. In this situation, we may represent the covariance operator $Q_u \in L_1^+(H^1(D))$ by means of the kernel $C_u \in H^{1,1}(D^2)$ as follows:

$$(Q_u v)(x) = \int_D \big(C_u(x,y)\, v(y) + \partial_y C_u(x,y)\, v'(y)\big)\, dy$$

From now on, we shall be concerned with C_u only.

3.2. **Problem formulation.**

Notation 3.4. For the unit square, its diagonal and boundary, we will write:

$$\Box := \Big\{(x,y) \in \mathbb{R}^2 \,:\, 0 \le x, y \le 1\Big\}, \quad \Delta := \Big\{(x,y) \in \Box \,:\, x = y\Big\},$$
$$\Gamma := \Big\{(x,y) \in \Box \,:\, x = 0 \vee x = 1 \vee y = 0 \vee y = 1\Big\} = \partial\Box$$

This paper is dedicated to assessing the regularity of the 2-point correlation $C_u \in H^{1,1}(\Box)$ for the stochastic model equation described in Section 3.1. C_u is the solution to the following variational problem (cf. (3.3) with $k = 2$):

Problem 3.5. Find $C_u \in H^{1,1}(\Box)$ such that

$$B(C_u, \mathcal{C}) = F(\mathcal{C}) \quad \forall \mathcal{C} \in H^{1,1}(\Box),$$

where

$$B(C_u, \mathcal{C}) := \iint_\Box \big(\partial_x\partial_y C_u\, \partial_x\partial_y \mathcal{C} + b^2 \partial_x C_u\, \partial_x \mathcal{C} + b^2 \partial_y C_u\, \partial_y \mathcal{C} + b^4 C_u\, \mathcal{C}\big)\, dy\, dx,$$
$$F(\mathcal{C}) := \iint_\Box C_f\, \mathcal{C}\, dy\, dx.$$

Problem 3.5 is well-posed since B is continuous with constant $K_b^2 = (1 + b^2)^2$ and coercive with $\gamma_b^2 = \min\{1, b^4\} > 0$. Major challenges for the analysis to follow will arise from C_f being allowed to feature a singularity on the diagonal Δ. Before specifying precise assumptions on the 2-point correlation C_f, we introduce countably normed spaces based on weighted Sobolev spaces (see e.g. [2]).

[2] FOSM = First Order, Second Moment

Definition 3.6 (spaces $\mathcal{B}^l_{\beta,d}(0,1)$)**.** Let $0 \le \beta < 1$. Then, for all $l \in \mathbb{N}$ and natural numbers $k \ge l$,

$$|v|_{H^{k,l}_\beta(0,1)} := \left\| x^{\beta+k-l} v^{(k)} \right\|_{L^2(0,1)} \tag{3.6}$$

defines a seminorm. If $v \in H^{l-1}(0,1)$ and if there exist constants $C > 0$, $d \ge 1$ such that

$$|v|_{H^{k,l}_\beta(0,1)} \le C\, d^{k-l}(k-l)! \quad \forall k \ge l, \tag{3.7}$$

then we write $v \in \mathcal{B}^l_{\beta,d}(0,1)$, or simply $v \in \mathcal{B}^l_\beta(0,1)$.

The countably normed spaces $\mathcal{B}^l_{\beta,d}(0,1)$ contain functions that are analytic in $(0,1)$ and may feature an algebraic singularity at $x = 0$. Examples of functions in $\mathcal{B}^l_{\beta,d}(0,1)$ are $v(x) = x^\alpha$ with $\alpha > -\frac{1}{2}$ or $v(x) = \ln(x)$.

Lemma 3.7. *If $v \in \mathcal{B}^{l+1}_{\beta,d}(0,1)$, then v' belongs to the space $\mathcal{B}^l_{\beta,d}(0,1)$.*

Proof. According to Definition 3.6, $v \in \mathcal{B}^{l+1}_{\beta,d}(0,1)$ means

$$v \in H^l(0,1), \qquad \left\| x^{\beta+k-(l+1)} v^{(k)} \right\|_{L^2(0,1)} \le C\, d^{k-(l+1)}(k-(l+1))!$$

for all $k \ge l+1$. From this, it immediately follows

$$v' \in H^{l-1}(0,1), \qquad \left\| x^{\beta+(k-1)-l} (v')^{(k-1)} \right\|_{L^2(0,1)} \le C\, d^{(k-1)-l}((k-1)-l)!$$

for all $(k-1) \ge l$, which implies $v' \in \mathcal{B}^l_{\beta,d}(0,1)$. □

Lemma 3.8. *$\mathcal{B}^{l+1}_{\beta,d}(0,1) \subset \mathcal{B}^l_{\beta^\star,d}(0,1)$ with $\beta^\star \in [0,1)$ arbitrary.*

Proof. Let $v \in \mathcal{B}^{l+1}_{\beta,d}(0,1)$. Then, v satisfies

$$v \in H^l(0,1), \qquad \left\| x^{\beta+k-(l+1)} v^{(k)} \right\|_{L^2(0,1)} \le C\, d^{k-(l+1)}(k-(l+1))!$$

for all $k \ge l+1$. $v \in H^{l-1}(0,1)$ follows from the inclusion $H^l(0,1) \subset H^{l-1}(0,1)$. Now, let $\beta^\star \in [0,1)$ arbitrary. For the l-th derivative of v, it holds:

$$\left\| x^{\beta^\star+l-l} v^{(l)} \right\|_{L^2(0,1)} \le \|v^{(l)}\|_{L^2(0,1)} < \infty \tag{3.8}$$

For derivatives of order $k \ge l+1$, we have:

$$\begin{aligned} \left\| x^{\beta^\star+k-l} v^{(k)} \right\|_{L^2(0,1)} &\le \left\| x^{\beta+k-(l+1)} v^{(k)} \right\|_{L^2(0,1)} \qquad (\text{due to } \beta^\star > \beta - 1) \\ &\le C\, d^{k-(l+1)}(k-(l+1))! \le \frac{C}{d}\, d^{k-l}(k-l)! \end{aligned}$$

Combining this result with (3.8), we can write:

$$\left\| x^{\beta^\star+k-l} v^{(k)} \right\|_{L^2(0,1)} \le \max\left\{ \|v^{(l)}\|_{L^2(0,1)}, \frac{C}{d} \right\} d^{k-l}(k-l)! \qquad \forall k \ge l$$

Thus, $v \in \mathcal{B}^l_{\beta^\star,d}(0,1)$. □

Assumption 3.9. In addition to Assumption 3.1, we assume the Gaussian random field f to be *stationary*, i.e. its correlation kernel C_f is translation invariant:

$$C_f(x,y) = C_f(x+t, y+t) \quad \text{for all } t \text{ with } (x+t, y+t) \in \square$$

Thus, C_f can be written as a function of the difference $z = x - y$, $z \in [-1,1]$:

$$C_f = C_f(z), \quad z = x - y \tag{3.9}$$

Furthermore, it is assumed

$$C_f\big|_{[0,1]} \in \mathcal{B}^l_{\beta,d}(0,1) \cap C^0([0,1]) \tag{3.10}$$

with some $l \in \mathbb{N}$, $\beta \in [0,1)$ and $d \geq 1$.

Notice:

(1) From (3.5), it becomes clear that C_f in (3.9) has to be an even function so that it is sufficient to formulate assumption (3.10) with respect to the interval $[0,1]$ instead of $[-1,1]$.
(2) $C_f \in C^0([0,1])$ constitutes an extra requirement only in case $l = 1$; for $l \geq 2$, the continuity in $[0,1]$ follows from $C_f \in \mathcal{B}^l_{\beta,d}(0,1) \subset H^1(0,1) \subset C^0([0,1])$.

Assumption 3.9 is reasonable in the sense that one can find correlation models in the literature on spatial statistics which meet (3.9) and (3.10). Examples will be given at the end of the next subsection.

3.3. **Examples of correlation kernels.** It is important to realize that not every function $C : \square \to \mathbb{R}$ satisfying (3.9) and (3.10) is automatically 2-point correlation of some random field.

Definition 3.10 (positive definiteness)**.** Let T be a set and $C : T \times T \to \mathbb{R}$. The function C is positive (semi-)definite, if

$$\sum_{i=1}^{k} \sum_{j=1}^{k} a_i\, a_j\, C(t_i, t_j) \geq 0 \tag{3.11}$$

for all $k \in \mathbb{N}$, $\{t_1, \ldots, t_k\} \subset T$ and $a_1, \ldots, a_k \in \mathbb{R}$.

Every 2-point correlation $C_f : \square \to \mathbb{R}$ has to be positive definite: Let $k \in \mathbb{N}$ and $x_1, \ldots, x_k \in D$. Then, for any linear combination of the random variables $f(x_1, \cdot), \ldots, f(x_k, \cdot)$, it must hold:

$$\begin{aligned} &\int_\Omega \big(a_1 f(x_1,\omega) + \ldots + a_k f(x_k,\omega)\big)^2\, \mathbb{P}(d\omega) \geq 0 \\ \Leftrightarrow \quad &\sum_{i=1}^{k} \sum_{j=1}^{k} a_i\, a_j \underbrace{\int_\Omega f(x_i,\omega) f(x_j,\omega)\, \mathbb{P}(d\omega)}_{= C_f(x_i, x_j)} \geq 0 \end{aligned} \tag{3.12}$$

The opposite direction is also true, i.e. every positive definite function is correlation kernel for some random field.

Theorem 3.11 (cf. Theorem 1.2 in [1] and the references there)**.** *The class of positive definite functions coincides with the class of correlation kernels.*

Remark 3.12. Usually, the above theorem is stated in terms of covariance instead of correlation. In its proof, argument (3.12) is then replaced by the fact that the variance of $a_1 f(x_1, \cdot) + \ldots + a_k f(x_k, \cdot)$ is non-negative for any choice of the a_i. The opposite direction works with a centered random field so that covariance and correlation kernels are identical anyway (see [1, 11] and the references therein).

Testing a given function for positive definiteness by means of (3.11) is in general difficult and we refrain from listing sufficient conditions for positive definiteness. Instead, we give references to the literature allowing to check that the following examples actually are admissible correlation kernels.

Example 3.13. Let

$$C_f(x,y) := c_1 - c_2\,|x-y|^\gamma, \qquad \gamma \in (0,1),\ \ c_1 \geq c_2 > 0.$$

This function constitutes an admissible correlation model on the unit interval, whose positive definiteness can be deduced from Pólya's criterion (see [9]). C_f satisfies assumptions (3.9) and (3.10). In particular, one can show that

$$C_f \in \mathcal{B}^l_{\beta,d}(0,1) \quad \text{with} \quad \begin{cases} l=1,\ \beta \in \left(\frac{1}{2}-\gamma, 1\right),\ d=1 & \text{if } \gamma \in \left(0,\frac{1}{2}\right] \\ l=2,\ \beta \in \left(\frac{3}{2}-\gamma, 1\right),\ d=1+\epsilon & \text{if } \gamma \in \left(\frac{1}{2},1\right) \end{cases}$$

where the value $\epsilon > 0$ may be chosen arbitrarily small (see [7, Examples 2.11 and 2.16]).

Example 3.14 (damped sine or hole effect model; cf. Section 2.1 in [11]). Let

$$C_f(x,y) := \left(\frac{|x-y|}{\lambda}\right)^{-1} \sin\left(\frac{|x-y|}{\lambda}\right), \quad \lambda > 0.$$

This model is a special case of a more general class of correlation kernels called the Bessel family. Its positive definiteness follows from a theorem by Schoenberg (see, e.g., [12]). The data's spatial correlation function C_f obviously satisfies assumption (3.9). In order to verify (3.10), we set $z = x-y$ and write

$$C_f(z) = \left(\frac{z}{\lambda}\right)^{-1} \sin\left(\frac{z}{\lambda}\right) = \sum_{i=0}^{\infty} (-1)^i \left(\frac{1}{\lambda}\right)^{2i} \frac{z^{2i}}{(2i+1)!}, \quad z \in [0,1].$$

Now, we investigate the derivatives

$$C_f^{(k)}(z) = \sum_{\substack{i\in\mathbb{N}_0:\\ i\geq k/2}} (-1)^i \left(\frac{1}{\lambda}\right)^{2i} \frac{z^{2i-k}}{(2i+1)(2i-k)!}, \quad k \in \mathbb{N}_0.$$

Since $2i-k$ is non-negative for all summands, the $L^2(0,1)$-norm of $C_f^{(k)}$ can be bounded as follows:

$$\begin{aligned}
\|C_f^{(k)}\|_{L^2(0,1)} &\leq \sum_{\substack{i\in\mathbb{N}_0:\\ i\geq k/2}} \frac{\lambda^{-2i}}{(2i+1)(2i-k)!} \underbrace{\|z^{2i-k}\|_{L^2(0,1)}}_{\leq 1} \\
&\leq \frac{1}{k+1}\left(\frac{1}{\lambda}\right)^k \sum_{\substack{i\in\mathbb{N}_0:\\ i\geq k/2}} \frac{1}{(2i-k)!}\left(\frac{1}{\lambda}\right)^{2i-k} \leq \frac{1}{k+1}\left(\frac{1}{\lambda}\right)^k e^{\frac{1}{\lambda}}
\end{aligned}$$

Obviously, all derivatives are square-integrable on $(0,1)$ even without multiplication by the weighting factor in seminorm (3.6); C_f does not exhibit any singularity at $z=0$. In fact, we may choose l, β and d arbitrarily and will always find assumption (3.10) satisfied,

$$C_f\big|_{[0,1]} \in \mathcal{B}^l_{\beta,d}(0,1) \cap C^0([0,1]).$$

With regard to bound (3.7), the trivial estimate

$$\begin{aligned}
|C_f|_{H^{k,l}_\beta(0,1)} &\leq \|C_f^{(k)}\|_{L^2(0,1)} \leq \frac{1}{k+1}\left(\frac{1}{\lambda}\right)^k e^{\frac{1}{\lambda}} \\
&\leq \frac{e^{\frac{1}{\lambda}}}{(l+1)\,\lambda^l}\left(\frac{1}{d\lambda}\right)^{k-l} \frac{1}{(k-l)!}\, d^{k-l}(k-l)! \leq \frac{e^{\frac{d+1}{d\lambda}}}{(l+1)\,\lambda^l}\, d^{k-l}(k-l)!
\end{aligned}$$

holds for all $k \geq l$.

Example 3.15 (power-exponential models; cf. Section 2.1 in [11]). Let

$$C_f(x,y) := \exp\left(-\frac{|x-y|^\gamma}{\lambda^\gamma}\right), \qquad \gamma \in (0,2],\ \ \lambda > 0.$$

A proof that these models are positive definite is given in [5]. Setting $z = x - y$, the derivatives of $C_f(z)$ can be written as

$$C_f^{(k)}(z) = \lambda^{-k} g^{(k)}\Big(\frac{z}{\lambda}\Big), \quad z \in [0,1], \quad k \in \mathbb{N}_0,$$

with the auxiliary function $g(t) := \exp(-t^\gamma)$. C_f is in $L^2(0,1)$; thus, we choose $l = 1$ and estimate the seminorm (3.6) by:

$$\begin{aligned}
|C_f|^2_{H^{k,1}_\beta(0,1)} &= \big\|z^{\beta+k-1} C_f^{(k)}\big\|^2_{L^2(0,1)} = \int_0^1 z^{2\beta+2k-2} \lambda^{-2k} g^{(k)}\Big(\frac{z}{\lambda}\Big)^2 \, dz \\
&= \lambda^{2\beta-1} \int_0^{1/\lambda} t^{2\beta+2k-2} g^{(k)}(t)^2 \, dt \\
&= \lambda^{2\beta-1} \int_0^{1/\lambda} \Big[t^{k-\gamma} g^{(k)}(t)\Big]^2 t^{2\gamma+2\beta-2} \, dt \\
&\le \lambda^{2\beta-1} \max_{t\ge 0} \Big[t^{k-\gamma} g^{(k)}(t)\Big]^2 \int_0^{1/\lambda} t^{2\gamma+2\beta-2} \, dt
\end{aligned}$$

For $\beta > \frac{1}{2} - \gamma$, we obtain:

$$\begin{aligned}
|C_f|_{H^{k,1}_\beta(0,1)} &\le \frac{1}{\sqrt{2\gamma+2\beta-1}} \Big(\frac{1}{\lambda}\Big)^\gamma \max_{t\ge 0} \big|t^{k-\gamma} g^{(k)}(t)\big| \\
&= \frac{1}{\sqrt{2\gamma+2\beta-1}} \Big(\frac{1}{\lambda}\Big)^\gamma \underbrace{\max_{t\ge 0} \frac{\big|t^{k-\gamma} g^{(k)}(t)\big|}{(k-1)!}}_{=:\ m_\gamma(k)} (k-1)!
\end{aligned}$$

The maximum in the last line was evaluated for different values of γ up to an order of $k = 60$ (with the aid of Mathematica). The results are shown in Figure 3.1. We observe that the curves on the left hand side are decreasing, whereas the curves on the right hand side are increasing while approaching straight lines. In any case, this suggests that $m_\gamma(k)$ can be bounded by a term of the form $C\,d^{k-1}$, i.e. that assumption (3.10) is satisfied by the function C_f. In particular, the constant d need not be chosen larger than approximately 1.5.

For further examples of stationary correlation models, see [1, 11].

4. Regularity of C_u

In this section, we examine location and order of the singularities in C_u which arise from the singularity in C_f (cf. Assumption 3.9). The results to be found are fundamental for the design of appropriate *hp*-finite element spaces for the approximation of C_u (see the second part [8] of this paper or [7, Chapters 4 and 5]).

Proposition 4.1 (strong solution). *If $C_u : \square \to \mathbb{R}$ satisfies*

$$\left.\begin{aligned}
\big(\partial_x^2\partial_y^2 - b^2\partial_x^2 - b^2\partial_y^2 + b^4\,\mathrm{id}\big)\, C_u &= C_f && \text{in } \square \\
\partial_x C_u &\equiv 0 && \text{on } \{0,1\}\times I \\
\partial_y C_u &\equiv 0 && \text{on } I\times\{0,1\}
\end{aligned}\right\}, \tag{4.1}$$

then C_u is the solution of Problem 3.5.

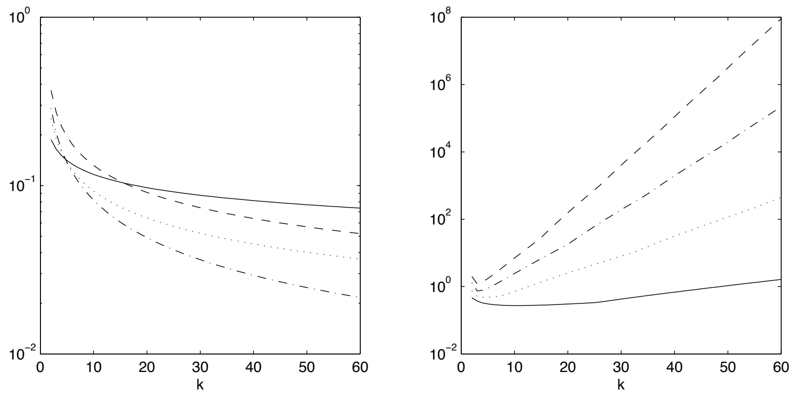


FIGURE 3.1. Logarithmic plot of $m_\gamma(k)$ for $k \leq 60$
Left: $\gamma = 0.25$ (solid), 0.5 (dotted), 0.75 (dash-dot), 1.0 (dashed)
Right: $\gamma = 1.25$ (solid), 1.5 (dotted), 1.75 (dash-dot), 2.0 (dashed)

Remark 4.2 (non-hypoellipticity; cf. Section 1 in [4]). The symbol associated with the differential operator in (4.1) is given by

$$P(i\boldsymbol{\xi}) = \xi_1^2\xi_2^2 + b^2\xi_1^2 + b^2\xi_2^2 + b^4 = (\xi_1^2 + b^2)(\xi_2^2 + b^2).$$

According to [6, Theorem 11.1.1], hypoellipticity is equivalent to

$$\frac{\partial_{\xi_1}^{\alpha_1}\partial_{\xi_2}^{\alpha_2} P(i\boldsymbol{\xi})}{P(i\boldsymbol{\xi})} \to 0 \qquad \text{as} \qquad \boldsymbol{\xi} \to \infty \ \text{ in } \mathbb{R}^2 \tag{4.2}$$

for all $(\alpha_1, \alpha_2) \in \mathbb{N}_0^2 \setminus \{(0,0)\}$. We select $\alpha_1 = 2$, $\alpha_2 = 0$ in order to obtain $\partial_{\xi_1}^2 P(i\boldsymbol{\xi})/P(i\boldsymbol{\xi}) = 2/(\xi_1^2 + b^2)$. With $\xi_1 = $ const, $\xi_2 \to \infty$, it is obvious that (4.2) does not hold. Thus, the differential operator in (4.1) is *not* hypoelliptic. We shall see the consequences of this fact on the singular support of C_u later.

Neglecting the boundary conditions for a moment, the next two lemmas provide an integral representation of stationary solutions $C_u(z)$, $z = x - y$, to the differential equation in (4.1) and assess the regularity of these solutions.

Lemma 4.3. *Let $g \in C^0([-1,1])$ and $b > 0$. Then, the general solution to the ordinary differential equation*

$$v^{(4)}(z) - 2b^2\, v''(z) + b^4\, v(z) = g(z) \qquad \forall z \in (-1,1) \tag{4.3}$$

is given by

$$v(z) = J_g(z) + c_1 \cosh(b\,z) + c_2 \sinh(b\,z) + c_3\, z \cosh(b\,z) + c_4\, z \sinh(b\,z) \tag{4.4}$$

with $c_i \in \mathbb{R}$ and

$$J_g(z) = \int_0^z \left(\frac{z-t}{2\,b^2} \cosh\big(b\,(z-t)\big) - \frac{\sinh\big(b\,(z-t)\big)}{2\,b^3} \right) g(t)\,dt. \tag{4.5}$$

In particular: If g is an even function, J_g is even as well—in this case, v is even, iff $c_2 = c_3 = 0$.

Proof. Equation (4.3) is a linear ODE of fourth order with constant coefficients and an inhomogeneity g. The functions $\cosh(b\,z)$, $\sinh(b\,z)$, $z\cosh(b\,z)$ and $z\sinh(b\,z)$ span the space of solutions to the corresponding homogeneous equation. The linear combination

$$v_h(z) := \frac{z}{2\,b^2}\cosh(b\,z) - \frac{\sinh(b\,z)}{2\,b^3}$$

satisfies

$$v_h^{(k)}(0) = \delta_{k,3} \qquad \forall k \in \{0,1,2,3\},$$

where δ_{ij} denotes the Kronecker delta. For such a function, it is known that

$$v_p(z) := \int_0^z v_h(z-t)\, g(t)\, dt$$

is a particular solution of the inhomogeneous equation. Setting $J_g = v_p$, the proof is complete. $\square$

Lemma 4.4. *Let $g \in C^0([-1,1])$. In addition, let $g|_{[0,1]}$ belong to $\mathcal{B}^l_{\beta,d}(0,1)$ and v be any solution to the ordinary differential equation (4.3),*

$$v^{(4)}(z) - 2b^2\, v''(z) + b^4\, v(z) = g(z) \qquad \forall z \in (-1,1).$$

Then, $v \in \mathcal{B}^{l+4}_{\beta,d}(0,1) \cap C^4([-1,1])$.

Proof. The proof is split into several steps.
First step. We claim that the derivatives of v are given by

$$v^{(2n)} = \sum_{j=0}^{n-2} (j+1)\, b^{2j} g^{(2n-4-2j)} + n\, b^{2n-2} v'' + (1-n)\, b^{2n} v, \tag{4.6a}$$

$$v^{(2n+1)} = \sum_{j=0}^{n-2} (j+1)\, b^{2j} g^{(2n-3-2j)} + n\, b^{2n-2} v^{(3)} + (1-n)\, b^{2n} v'. \tag{4.6b}$$

The second line follows from the first one. The first line is proven by induction. For $n = 0, 1$, we verify directly that

$$v^{(2\cdot 0)} = 0 + 0\, b^{-2} v'' + (1-0)\, b^0 v, \qquad v^{(2\cdot 1)} = 0 + 1\, b^0 v'' + (1-1)\, b^2 v\ .$$

The induction step "$n \to n+1$" for $n \geq 1$ is obtained as follows:

$$\begin{aligned} v^{(2(n+1))} &= \left(v^{(2n)}\right)'' \\ &= \sum_{j=0}^{n-2} (j+1)\, b^{2j} g^{(2n-4-2j+2)} + n\, b^{2n-2} v^{(4)} + (1-n)\, b^{2n} v''\ . \end{aligned}$$

Now, use (4.3) and $n - 1 \geq 0$:

$$\begin{aligned} v^{(2(n+1))} &= \sum_{j=0}^{n-2} (j+1)\, b^{2j} g^{(2(n+1)-4-2j)} \\ &\quad + n\, b^{2n-2} \left(g + 2b^2\, v'' - b^4\, v\right) + (1-n)\, b^{2n} v'' \\ &= \sum_{j=0}^{n-1} (j+1)\, b^{2j} g^{(2(n+1)-4-2j)} + (n+1)\, b^{2n} v'' + (1-(n+1))\, b^{2n+2} v\ . \end{aligned}$$

For convenience of notation, the two expressions (4.6a) and (4.6b) are merged into one:

$$v^{(k)} = \sum_{i=0}^{k-4} \alpha_i\, g^{(k-4-i)} + \sum_{i=0}^{3} \gamma_{i,k}\, v^{(i)} \qquad \forall k \in \mathbb{N}_0, \tag{4.7}$$

where

$$\alpha_i := \begin{cases} \left(\frac{i}{2}+1\right) b^i & \text{if } i \text{ is even,} \\ 0 & \text{if } i \text{ is odd,} \end{cases} \tag{4.8}$$

$$\begin{aligned} \gamma_{0,k} &:= \mathbf{1}_{2\mathbb{Z}}(k)\Big(1-\frac{k}{2}\Big)\, b^k, \quad \gamma_{1,k} := \mathbf{1}_{2\mathbb{Z}+1}(k)\Big(1-\frac{k-1}{2}\Big)\, b^{k-1}, \\ \gamma_{2,k} &:= \mathbf{1}_{2\mathbb{Z}}(k)\,\frac{k}{2}\, b^{k-2}, \quad \gamma_{3,k} := \mathbf{1}_{2\mathbb{Z}+1}(k)\,\frac{k-1}{2}\, b^{k-3}. \end{aligned} \tag{4.9}$$

Writing $v^{(k)}$ in this manner, both cases which were distinguished in (4.6) (i.e., between even/odd differentiation order) are "hidden" in α_i and $\gamma_{i,k}$.
Second step. $v|_{[0,1]} \in H^{l+3}(0,1)$
From $g \in C^0([-1,1])$, it follows that $v \in C^4([-1,1])$ and in particular v, v', $\ldots$, $v^{(4)} \in L^2(0,1)$. With $g \in \mathcal{B}^l_{\beta,d}(0,1)$, we have $g^{(j)} \in L^2(0,1)$ for all $j \le l-1$. Thus, the representation of $v^{(k)}$ by (4.7) is a linear combination of L^2-functions as long as $k \le l+3$. This implies $v^{(k)} \in L^2(0,1)$ for all $k \le l+3$, which is equivalent to $v \in H^{l+3}(0,1)$.
Third step. We show $|v|_{H^{k,l+4}_\beta(0,1)} \le \bar{C}\, d^{k-(l+4)}(k-(l+4))!$ for all $k \ge l+4$.
For $k \ge l+4$, we split (4.7) as follows:

$$v^{(k)} = \underbrace{\sum_{i=0}^{k-(l+4)} \alpha_i\, g^{(k-4-i)}}_{=:\, s_1} + \underbrace{\sum_{i=k-(l+4)+1}^{k-4} \alpha_i\, g^{(k-4-i)} + \sum_{i=0}^{3} \gamma_{i,k}\, v^{(i)}}_{=:\, s_2}\ .$$

In general, s_1 is a linear combination of non-square-integrable functions on $(0,1)$, whereas all functions occurring in s_2 belong to $L^2(0,1)$. By the triangle inequality, we obtain the following bound:

$$\begin{aligned} |v|_{H^{k,l+4}_\beta(0,1)} &= \left\| z^{\beta+k-(l+4)} v^{(k)} \right\|_{L^2(0,1)} \\ &\le \left\| z^{\beta+k-(l+4)} s_1 \right\|_{L^2(0,1)} + \left\| z^{\beta+k-(l+4)} s_2 \right\|_{L^2(0,1)}\ . \end{aligned} \tag{4.10}$$

Next, a bound for the first summand on the right hand side is derived.

$$\begin{aligned} \left\| z^{\beta+k-(l+4)} s_1 \right\|_{L^2(0,1)} &\le \sum_{i=0}^{k-(l+4)} |\alpha_i| \cdot \left\| z^{\beta+k-(l+4)} g^{(k-4-i)} \right\|_{L^2(0,1)} \\ &\le \sum_{i=0}^{k-(l+4)} |\alpha_i| \cdot \left\| z^{\beta+(k-4-i)-l} g^{(k-4-i)} \right\|_{L^2(0,1)} \\ &\le C \sum_{i=0}^{k-(l+4)} |\alpha_i|\, d^{(k-4-i)-l}((k-4-i)-l)! \\ &= C\, d^{k-(l+4)}(k-(l+4))! \sum_{i=0}^{k-(l+4)} |\alpha_i|\, d^{-i}\, \frac{(k-l-4-i)!}{(k-l-4)!} \\ &\le C\, d^{k-(l+4)}(k-(l+4))! \sum_{i=0}^{k-(l+4)} |\alpha_i|\, d^{-i}\, \frac{1}{i!}\ . \end{aligned}$$

With definition (4.8), we can estimate further:

$$
\begin{aligned}
\left\| z^{\beta+k-(l+4)} s_1 \right\|_{L^2(0,1)} &\le C\, d^{k-(l+4)} (k-(l+4))! \sum_{j=0}^{\infty} |\alpha_{2j}|\, d^{-2j} \,\frac{1}{(2j)!} \\
&= C\, d^{k-(l+4)} (k-(l+4))! \sum_{j=0}^{\infty} (j+1) \left(\frac{b^2}{d^2}\right)^{j} \frac{1}{(2j)!} \\
&= C\, d^{k-(l+4)} (k-(l+4))! \sum_{j=0}^{\infty} \underbrace{\frac{(j+1)!}{(2j)!}}_{\le 1} \left(\frac{b^2}{d^2}\right)^{j} \frac{1}{j!} \\
&\le C \,\exp\left(\frac{b^2}{d^2}\right) d^{k-(l+4)} (k-(l+4))!\,. \qquad (4.11)
\end{aligned}
$$

It remains to investigate the second summand in (4.10):

$$
\begin{aligned}
\left\| z^{\beta+k-(l+4)} s_2 \right\|_{L^2(0,1)} &\le \|s_2\|_{L^2(0,1)} \\
&\le \sum_{i=k-(l+4)+1}^{k-4} |\alpha_i| \cdot \|g^{(k-4-i)}\|_{L^2(0,1)} + \sum_{i=0}^{3} |\gamma_{i,k}| \cdot \|v^{(i)}\|_{L^2(0,1)}\,.
\end{aligned}
$$

Defining $\tilde{C} := \max\left\{\|g\|_{H^{l-1}(0,1)}, \|v\|_{H^3(0,1)}\right\}$ and $\tilde{b} := \max\{1, b\}$, we obtain:

$$
\begin{aligned}
\left\| z^{\beta+k-(l+4)} s_2 \right\|_{L^2(0,1)} &\le \tilde{C}\left(\sum_{i=k-(l+4)+1}^{k-4} |\alpha_i| + \sum_{i=0}^{3} |\gamma_{i,k}| \right) \\
&\le \tilde{C}\Big(l \max\left\{ |\alpha_{k-(l+4)+1}|, \dots, |\alpha_{k-4}| \right\} \\
&\qquad\qquad + 4 \max\left\{ |\gamma_{0,k}|, \dots, |\gamma_{3,k}| \right\} \Big)\,.
\end{aligned}
$$

With $k \ge l+4 \ge 5$ in definition (4.9):

$$
\begin{aligned}
\left\| z^{\beta+k-(l+4)} s_2 \right\|_{L^2(0,1)} &\le \tilde{C}\left(l \left(\frac{k}{2} - 1\right) \tilde{b}^{k-4} + 4\,\frac{k}{2}\,\tilde{b}^{k} \right) \le \frac{1}{2}\,\tilde{C}\,\tilde{b}^4\,(l+4)\,\tilde{b}^{k-4}\,k \\
&= \frac{1}{2}\,\tilde{C}\,\tilde{b}^{l+4}\,(l+4)\,\underbrace{\frac{\left(\frac{\tilde{b}}{d}\right)^{k-(l+4)} k}{(k-(l+4))!}}_{(\star)}\, d^{k-(l+4)} (k-(l+4))!\,.
\end{aligned}
$$

Since the expression $(\star)$ converges to 0 for $k \to \infty$, it is in particular bounded for all $k \ge l+4$. Together with (4.10) and (4.11), this means that there exists a constant $\bar{C}$ such that

$$
|v|_{H^{k,l+4}_{\beta}(0,1)} \le \bar{C}\, d^{k-(l+4)} (k-(l+4))! \quad \text{for all } k \ge l+4.
$$

$v \in H^{l+3}(0,1)$ was already proven above. Thus, we have shown $v \in \mathcal{B}^{l+4}_{\beta,d}(0,1)$. □

Theorem 4.5. *Let C_f satisfy Assumption 3.9. With J_{C_f} defined by (4.5), the unique solution C_u to Problem 3.5 admits the representation*

$$
C_u(x,y) = C_u^{\Delta}(x-y) + C_u^{\Gamma}(x,y), \qquad (4.12)
$$

where

$$C_u^\Delta(z) = J_{C_f}(z) + \frac{J''_{C_f}(1)}{b^3 \sinh b}\big(2\cosh(b\,z) - b\,z\sinh(b\,z)\big), \tag{4.13a}$$

$$\begin{aligned} C_u^\Gamma(x,y) = &-(C_u^\Delta)'(1-x)\,\frac{\cosh(b\,y)}{b\sinh b} - (C_u^\Delta)'(x)\,\frac{\cosh\big(b\,(1-y)\big)}{b\sinh b} \\ &-(C_u^\Delta)'(1-y)\,\frac{\cosh(b\,x)}{b\sinh b} - (C_u^\Delta)'(y)\,\frac{\cosh\big(b\,(1-x)\big)}{b\sinh b}\,. \end{aligned} \tag{4.13b}$$

Furthermore, $C_u^\Delta(z)$ is an even function of the difference $z = x - y$.

Proof. Due to Proposition 4.1, it suffices to verify that the C_u defined above solves problem (4.1).

Differential equation. First, we check that C_u satisfies the differential equation in (4.1). To this end, the differential operator is applied separately to the summands C_u^Δ and C_u^Γ: From (4.13a) and Lemma 4.3, it follows

$$\begin{aligned} &\big(\partial_x^2\partial_y^2 - b^2\partial_x^2 - b^2\partial_y^2 + b^4\,\mathrm{id}\big)\,C_u^\Delta(x-y) \\ &\qquad = (C_u^\Delta)^{(4)}(x-y) - 2b^2\,(C_u^\Delta)''(x-y) + b^4\,C_u^\Delta(x-y) = C_f(x-y). \end{aligned}$$

For C_u^Γ, we find

$$\begin{aligned} &\big(\partial_x^2\partial_y^2 - b^2\partial_x^2 - b^2\partial_y^2 + b^4\,\mathrm{id}\big)\,C_u^\Gamma(x,y) \\ &\qquad = \left\{ \begin{array}{l} \big(-\partial_x^2 + b^2\,\mathrm{id}\big)\big(-\partial_y^2 + b^2\,\mathrm{id}\big)\,C_u^\Gamma(x,y) \\ \big(-\partial_y^2 + b^2\,\mathrm{id}\big)\big(-\partial_x^2 + b^2\,\mathrm{id}\big)\,C_u^\Gamma(x,y) \end{array} \right\} = 0 \end{aligned}$$

as a consequence from the special structure in (4.13b) and the fact that

$$\big(-\partial_z^2 + b^2\,\mathrm{id}\big)\cosh(b\,z) = \big(-\partial_z^2 + b^2\,\mathrm{id}\big)\cosh\big(b\,(1-z)\big) = 0 \quad \forall z.$$

Boundary conditions. We verify that the homogeneous Neumann boundary conditions in (4.1) are fulfilled: The second derivative of C_u^Δ is

$$(C_u^\Delta)''(z) = J''_{C_f}(z) - \frac{J''_{C_f}(1)}{\sinh b}\,z\,\sinh(b\,z).$$

Because of $J''_{C_f}(0) = 0$ (cf. the proof of Lemma 4.3), it follows that $(C_u^\Delta)''(0) = (C_u^\Delta)''(1) = 0$. For this reason, we obtain from (4.13b):

$$\begin{aligned} \partial_x C_u^\Gamma(0,y) &= +(C_u^\Delta)'(y), & \partial_x C_u^\Gamma(1,y) &= -(C_u^\Delta)'(1-y) \\ \partial_y C_u^\Gamma(x,0) &= +(C_u^\Delta)'(x), & \partial_y C_u^\Gamma(x,1) &= -(C_u^\Delta)'(1-x)\,. \end{aligned}$$

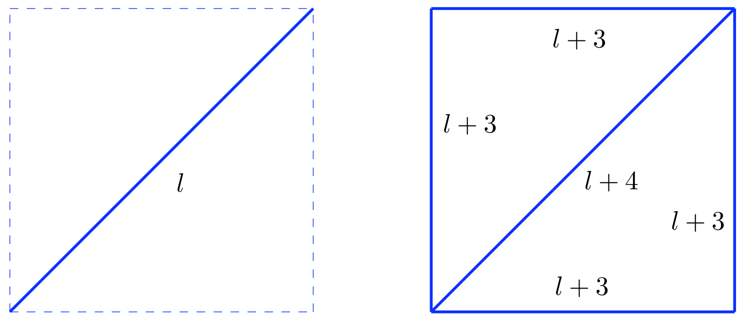
By Lemma 4.3, J_{C_f} is an even function, because $C_f(z)$ is even. Thus, $C_u^\Delta(z)$ is even as well, and $(C_u^\Delta)'(z)$ is odd. Along with the above identities, this yields:

$$\begin{aligned} \partial_x C_u(0,y) &= (C_u^\Delta)'(-y) + \partial_x C_u^\Gamma(0,y) = -(C_u^\Delta)'(y) + \partial_x C_u^\Gamma(0,y) &&= 0\,, \\ \partial_x C_u(1,y) &= (C_u^\Delta)'(1-y) + \partial_x C_u^\Gamma(1,y) &&= 0\,, \\ \partial_y C_u(x,0) &= -(C_u^\Delta)'(x) + \partial_y C_u^\Gamma(x,0) &&= 0\,, \\ \partial_y C_u(x,1) &= -(C_u^\Delta)'(x-1) + \partial_y C_u^\Gamma(x,1) = (C_u^\Delta)'(1-x) + \partial_y C_u^\Gamma(x,1) &&= 0\,. \end{aligned}$$

This completes the proof. □

Remark 4.6. From the proof of Theorem 4.5, one can see which roles the individual summands C_u^Δ and C_u^Γ in (4.12) play in the context of problem (4.1),

$$\begin{aligned} \big(\partial_x^2\partial_y^2 - b^2\partial_x^2 - b^2\partial_y^2 + b^4\,\mathrm{id}\big)\,C_u &= C_f && \text{in } \square, \\ \partial_x C_u &\equiv 0 && \text{on } \{0,1\}\times I, \\ \partial_y C_u &\equiv 0 && \text{on } I\times\{0,1\}. \end{aligned}$$



sing supp C_f sing supp C_u

FIGURE 4.1. Illustration of Corollary 4.7. The singularity of C_f on the diagonal (order l) gives rise to singularities in C_u on the diagonal (order $l+4$) and the boundary (order $l+3$).

C_u^Δ is a stationary function satisfying the differential equation; C_u^Γ lies in the kernel of the differential operator and compensates for the boundary conditions which C_u^Δ cannot satisfy on its own (notice: A stationary function satisfying the homogeneous Neumann conditions would have to be constant!).

The main result of this paper is:

Corollary 4.7 (singularities in C_u)**.** *Let C_f satisfy Assumption 3.9, i.e. in particular $C_f \in \mathcal{B}^l_{\beta,d}(0,1)$. Then, the unique solution C_u to Problem 3.5 admits a splitting*

$$C_u = C_u^\Delta + C_u^\Gamma \tag{4.14}$$

with

$$C_u^\Delta(x,y) = w_1(x-y), \quad w_1(z) \text{ an even function of } z = x-y, \tag{4.15a}$$

$$\begin{aligned} C_u^\Gamma(x,y) &= w_2(1-x)\, w_3(y) + w_2(x)\, w_3(1-y) \\ &\quad + w_3(x)\, w_2(1-y) + w_3(1-x)\, w_2(y), \end{aligned} \tag{4.15b}$$

where $w_1 \in \mathcal{B}^{l+4}_{\beta,d}(0,1) \cap C^4([-1,1])$, $w_2 \in \mathcal{B}^{l+3}_{\beta,d}(0,1)$, and w_3 is an analytic function on $[0,1]$ satisfying

$$\max_{x\in[0,1]} \left| w_3^{(k)}(x) \right| \le b^{-1} \coth(b)\, b^k \qquad \forall k \in \mathbb{N}_0. \tag{4.16}$$

Proof. The representation of C_u by (4.14), (4.15) follows from Theorem 4.5. Since $C_u^\Delta(z)$ solves equation (4.3) with right hand side $g = C_f(z) \in \mathcal{B}^l_{\beta,d}(0,1)$ (cf. the proof of Theorem 4.5), Lemma 4.4 yields $C_u^\Delta(z) \in \mathcal{B}^{l+4}_{\beta,d}(0,1) \cap C^4([-1,1])$. By Lemma 3.7, we have $(C_u^\Delta)'(z) \in \mathcal{B}^{l+3}_{\beta,d}(0,1)$, and thus, with (4.13b), we obtain that $w_2 = -(C_u^\Delta)'$ belongs to $\mathcal{B}^{l+3}_{\beta,d}(0,1)$. The claim (4.16) follows from $w_3(x) = \frac{\cosh(b\,x)}{b \sinh b}$ and

$$\begin{aligned} \left| \frac{d^{2k}}{dx^{2k}} \cosh(b\,x) \right| &= \left| b^{2k} \cosh(b\,x) \right| && \le b^{2k} \cosh(b) \\ \left| \frac{d^{2k+1}}{dx^{2k+1}} \cosh(b\,x) \right| &= \left| b^{2k+1} \sinh(b\,x) \right| \le b^{2k+1} \sinh(b) && \le b^{2k+1} \cosh(b) \end{aligned}$$

for all $x \in [0,1]$ and $k \in \mathbb{N}_0$. $\square$

Remark 4.8 (enlargement of the singular support). Corollary 4.7 shows

$$\operatorname{sing\,supp} C_u = \Delta \cup \Gamma \supsetneq \Delta = \operatorname{sing\,supp} C_f.$$

This increase of the solution's singular support is a consequence of the non-hypoelliptic nature of the differential operator $A \otimes A$ (see Remark 4.2). The principal consequence of this observation for the design of efficient finite element approximations of C_u is that not only the singularity on Δ (see [8]) has to be resolved, but in addition those on the boundary Γ (see [7, Chapter 4]).

Remark 4.9 (Dirichlet boundary conditions). An explicit representation of C_u can also be derived for homogeneous Dirichlet boundary conditions. In this case, proceeding in an analogous way as above yields

$$C_u(x,y) = C_u^{\Delta}(x-y) + C_u^{\Gamma}(x,y),$$

where

$$\begin{aligned} C_u^{\Delta}(z) &= J_{C_f}(z) - \frac{J_{C_f}(1)}{\sinh b}\, z \sinh(b\, z), \\ C_u^{\Gamma}(x,y) &= -C_u^{\Delta}(1-x)\,\frac{\sinh(b\,y)}{\sinh b} - C_u^{\Delta}(x)\,\frac{\sinh\big(b\,(1-y)\big)}{\sinh b} \\ &\quad - C_u^{\Delta}(1-y)\,\frac{\sinh(b\,x)}{\sinh b} - C_u^{\Delta}(y)\,\frac{\sinh\big(b\,(1-x)\big)}{\sinh b}\,. \end{aligned}$$

The equivalent to Corollary 4.7 then has $w_2 \in \mathcal{B}^{l+4}_{\beta,d}(0,1)$ instead of $w_2 \in \mathcal{B}^{l+3}_{\beta,d}(0,1)$ and a different bound (4.16) for the derivatives of the analytic function w_3.

5. Small correlation lengths

Assume we are given an even template function $\widehat{C}_f : \mathbb{R} \to \mathbb{R}$, from which we may generate 2-point correlations $C_f^{\lambda} : [-1,1] \to \mathbb{R}$ by

$$C_f^{\lambda}(x-y) := \widehat{C}_f\left(\frac{x-y}{\lambda}\right), \quad \lambda > 0. \tag{5.1}$$

In this context, the parameter λ is sometimes referred to as *correlation length*. A situation as in (5.1) was already encountered before in Examples 3.14 and 3.15. With regard to establishing Assumption 3.9 for C_f^{λ}, i.e. in particular $C_f^{\lambda}|_{[0,1]} \in \mathcal{B}^{l}_{\beta,d}(0,1)$, we are interested in the behavior of the values l, β and d when λ becomes smaller and smaller.

The application of the chain rule for differentiation yields the equation

$$\|z^{\beta+k-l}(C_f^{\lambda})^{(k)}\|_{L^2(0,1)} = \lambda^{\beta+l-\frac{1}{2}}\,\|z^{\beta+k-l}\,\widehat{C}_f^{(k)}\|_{L^2(0,\frac{1}{\lambda})} \quad \forall k \geq l. \tag{5.2}$$

An immediate consequence is the following:

Proposition 5.1. *Let $l \in \mathbb{N}$, $\beta \in [0,1)$ and $d \geq 1$. Furthermore, let $\widehat{C}_f : \mathbb{R} \to \mathbb{R}$ be an even function with*

$$\widehat{C}_f \in H^{l-1}(\mathbb{R}) \quad \text{and} \quad \|z^{\beta+k-l}\,\widehat{C}_f^{(k)}\|_{L^2(0,\infty)} \leq C\, d^{k-l}(k-l)! \quad \forall k \geq l. \tag{5.3}$$

Then, the C_f^{λ} obtained from (5.1) satisfy Assumption 3.9 with the same values l, β and d independent of λ.

Under condition (5.3), the influence of small correlation lengths λ is completely "absorbed" by the factor $\lambda^{\beta+l-\frac{1}{2}}$ in (5.2), which exclusively affects the constant C in the upper bound (3.7) for $|C_f^{\lambda}|_{H^{k,l}_{\beta}(0,1)}$. In particular, the above proposition implies that the regularity results for C_u from the previous section do not deteriorate as λ goes to 0. For this reason, a numerical method for the computation of C_u based on these regularity results can in principle remain stable for arbitrarily small correlation lengths.

6. Galerkin FEM for C_u

6.1. **Galerkin projection.** Let us recall Problem 3.5,

$$\text{find } C_u \in H^{1,1}(\Box) \text{ such that } B(C_u, \tilde{C}) = F(\tilde{C}) \quad \forall \tilde{C} \in H^{1,1}(\Box), \tag{6.1}$$

with B and F being defined there. We consider the Galerkin discretizations

$$\text{find } s_L \in S_L \text{ such that } B(s_L, \tilde{C}) = F(\tilde{C}) \quad \forall \tilde{C} \in S_L, \tag{6.2}$$

with nested finite-dimensional ansatz/test spaces

$$S_{L_0} \subset S_{L_0+1} \subset \ldots \subset S_{L-1} \subset S_L \subset S_{L+1} \subset \ldots \subset H^{1,1}(\Box). \tag{6.3}$$

Here, the subscript L denotes a "level of refinement" to be specified below. Existence and uniqueness of the Galerkin solutions s_L follow from the continuity and the coercivity of the bilinear form $B(\cdot,\cdot)$ in Problem 3.5. The properties imply that $B(\cdot,\cdot)$ is regular on $S_L \times S_L$. Moreover, the unique Galerkin approximations s_L satisfy the property of *Galerkin orthogonality*,

$$B(C_u - s_L, \tilde{C}) = 0 \quad \forall \tilde{C} \in S_L. \tag{6.4}$$

By *Galerkin finite element method*, we refer to the computation of approximate solutions s_L to C_u based on (6.2), (6.3).

6.2. **Quasi-optimality.**

Definition 6.1 (best approximation error)**.** For $w \in H^{1,1}(\Box)$ and a finite dimensional subspace S_L, we define the best approximation error

$$Z(w, S_L, H^{1,1}(\Box)) := \inf_{\tilde{C} \in S_L} \|w - \tilde{C}\|_{H^{1,1}(\Box)}.$$

$Z(C_u, S_L, H^{1,1}(\Box))$ provides a measure for the quality by which C_u can be approximated from the space S_L with respect to the $H^{1,1}$-norm. Since the solution s_L to (6.2) is itself an element of S_L, the best approximation error is obviously a *lower* bound for the error $\|C_u - s_L\|_{H^{1,1}(\Box)}$:

$$\|C_u - s_L\|_{H^{1,1}(\Box)} \geq \inf_{\tilde{C} \in S_L} \|C_u - \tilde{C}\|_{H^{1,1}(\Box)} = Z(C_u, S_L, H^{1,1}(\Box))$$

On the other hand, $\|C_u - s_L\|_{H^{1,1}(\Box)}$ cannot be worse than a certain multiple of the best approximation error:

Lemma 6.2 (Céa)**.** *Let $s_L \in S_L$ be the solution to (6.2). Then, it holds*

$$\|C_u - s_L\|_{H^{1,1}(\Box)} \leq \frac{K_b^2}{\gamma_b^2} Z(C_u, S_L, H^{1,1}(\Box)),$$

where K_b^2 is the continuity constant of the bilinear form B and γ_b^2 its coercivity constant (see Section 3.2). In particular, these constants do not depend on the ansatz space S_L.

Remark 6.3 (quasi-optimality). Céa's lemma shows that the approximate solution s_L is, up to a constant, as good as the best approximation of C_u from the subspace S_L. In this sense, s_L is said to be quasi-optimal.

6.3. **Ansatz spaces and approximability.** In [7, Chapters 4 and 5], it is explained in detail how to construct appropriate *hp*-finite element spaces $S^{\Delta}_{\mu,L}$ and $S^{\Gamma}_{p,L}$ for the approximation of C_u's summands C^{Δ}_u and C^{Γ}_u, respectively. (Furthermore, the second part [8] of this paper is entirely dedicated to the approximation of functions as C^{Δ}_u.) The numbers $\mu \geq 1$ and $p \in \mathbb{N}$ are parameters that control the polynomial degrees in the *hp*-approximations. They have to be chosen in dependence of C_f (see [7, Remark 4.31 and proof of Theorem 5.49]). Due to space constraints, we refrain from replicating the definitions of $S^{\Delta}_{\mu,L}$ and $S^{\Gamma}_{p,L}$ here and only cite the respective approximation results from [8] or Theorem 5.49 and Corollary 5.50 in [7].

Theorem 6.4. *Let $S^{\Delta}_{\mu,L}$ be the sequence of hp-FE spaces from* [7, Definition 5.19]. *Then, there exists a constant $\mu \geq 1$ such that within the sequence $\left(S^{\Delta}_{\mu,L}\right)_{L\in\mathbb{N}}$ there are hp covariance approximations $v^{\Delta}_L \in S^{\Delta}_{\mu,L}$ that converge, as $N = \dim(S_L) \to \infty$, at an exponential rate towards C^{Δ}_u in $H^{1,1}(\Box)$:*

$$\|C^{\Delta}_u - v^{\Delta}_L\|_{H^{1,1}(\Box)} \leq c_1 \exp\left(-c_2 \sqrt[3]{N}\right), \tag{6.5}$$

where $N = \dim S^{\Delta}_{\mu,L} = \mathcal{O}(L^3)$ and c_1, c_2 are positive constants independent of N.

Theorem 6.5 (cf. Theorem 4.30, Corollary 4.32 in [7])**.** *Let $S^{\Gamma}_{p,L}$ be the spaces from* [7, Definition 4.28]. *Then, there exists choice of polynomial degrees $p = p(L) \in \mathbb{N}$ such that the hp-FE spaces in the sequence $\left(S^{\Gamma}_{p,L}\right)_{L\in\mathbb{N}}$ contain approximations $v^{\Gamma}_L \in S^{\Gamma}_{p,L}$ which exhibit exponential convergence towards C^{Γ}_u in $H^{1,1}(\Box)$:*

$$\|C^{\Gamma}_u - v^{\Gamma}_L\|_{H^{1,1}(\Box)} \leq c_1 \exp\left(-c_2 \sqrt[4]{N}\right), \tag{6.6}$$

where $N = \dim S^{\Gamma}_{p,L} = \mathcal{O}(L^4)$ and c_1, c_2 are positive constants independent of N.

6.4. **Exponential convergence.**

Theorem 6.6 (convergence rate of the Galerkin FEM)**.** *Let f be stationary, V'-valued random loads with finite second moments and with piecewise analytic covariance function $C_f(x-y)$ that satisfies Assumption 3.9. Let further*

$$S_L := S^{\Delta}_{\mu,L} + S^{\Gamma}_{p,L} \tag{6.7}$$

be the ansatz space for the Galerkin FEM (6.2), (6.3). Then, one can find $\mu \geq 1$ constant and $p = p(L) \in \mathbb{N}$ such that the approximate solutions $s_L \in S_L$ to Problem 3.5 obtained from the Galerkin FEM, converge at exponential rate towards the exact solution C_u:

$$\|C_u - s_L\|_{H^{1,1}(\Box)} \leq c_1 \exp\left(-c_2 \sqrt[4]{N}\right),$$

where $N := \dim S_L$ denotes the number of degrees of freedom and c_1, c_2 are positive constants independent of N.

Proof. s_L is the solution of problem (6.2). Thus, Lemma 6.2 yields:

$$\begin{aligned} \|C_u - s_L\|_{H^{1,1}(\Box)} &= \frac{K_b^2}{\gamma_b^2} \inf_{\mathcal{C}_L \in S_L} \|C_u - \mathcal{C}_L\|_{H^{1,1}(\Box)} \\ &\leq \frac{K_b^2}{\gamma_b^2} \|C_u - \mathcal{C}_L\|_{H^{1,1}(\Box)} \quad \forall \mathcal{C}_L \in S_L \end{aligned}$$

We choose $\mu \geq 1$, $p = p(L) \in \mathbb{N}$ as in Theorems 6.4, 6.5 and assume the approximations $v^{\Delta}_L \in S^{\Delta}_{\mu,L}$, $v^{\Gamma}_L \in S^{\Gamma}_{p,L}$ from there in order to set

$$\mathcal{C}_L = v^{\Delta}_L + v^{\Gamma}_L \in S_L.$$

With $C_u = C_u^\Delta + C_u^\Gamma$, the triangle inequality provides the error bound

$$\|C_u - s_L\|_{H^{1,1}(\Box)} \le \frac{K_b^2}{\gamma_b^2}\Big(\|C_u^\Delta - v_L^\Delta\|_{H^{1,1}(\Box)} + \|C_u^\Gamma - v_L^\Gamma\|_{H^{1,1}(\Box)}\Big).$$

By Theorems 6.4 and 6.5, we obtain:

$$\begin{aligned}\|C_u - s_L\|_{H^{1,1}(\Box)} &\le \frac{K_b^2}{\gamma_b^2}\Big(c_1^\Delta \exp\big(-c_2^\Delta L\big) + c_1^\Gamma \exp\big(-c_2^\Gamma L\big)\Big)\\ &\le \frac{K_b^2(c_1^\Gamma + c_1^\Delta)}{\gamma_b^2}\exp\Big(-\min\big\{c_2^\Gamma, c_2^\Delta\big\}\, L\Big) \qquad (6.8)\end{aligned}$$

Because of $N = \dim S_L = \mathcal{O}(L^4)$, there is a constant $c > 0$ such that $L \ge \sqrt[4]{\frac{N}{c}}$ for $L \to \infty$. Inserting this into (6.8) concludes the proof. $\Box$

References

[1] P. Abrahamsen. A review of Gaussian random fields and correlation functions. Technical Report 917, Norwegian Computing Center, Oslo, 1997.

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

[3] G. Da Prato. *An introduction to infinite-dimensional analysis*. Universitext. Springer-Verlag, Berlin, 2006. Revised and extended from the 2001 original by Da Prato.

[4] K. Doppel and R. Hochmuth. On the regularity of solutions of a homogeneous Dirichlet problem for a nonhypoelliptic linear partial differential operator. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 16(1):183–198, 1991.

[5] W. Feller. *An introduction to probability theory and its applications. Vol. II.* John Wiley & Sons Inc., New York, second edition, 1971.

[6] L. Hörmander. *The analysis of linear partial differential operators. II*, volume 257 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1983. Differential operators with constant coefficients.

[7] B. Pentenrieder. *Exponential convergence of hp-FEM for second moments of elliptic PDEs with stochastic data*. PhD thesis, ETH Zürich Diss. No. 18729, 2009.

[8] B. Pentenrieder and Ch. Schwab. *hp*-FEM for second moments of elliptic PDEs with stochastic data. Part 2: Exponential convergence. Research Report 2010-??, Seminar for Applied Mathematics, ETH Zürich, 2010. In review.

[9] G. Pólya. Remarks on characteristic functions. In *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, 1945, 1946*, pages 115–123, Berkeley and Los Angeles, 1949. University of California Press.

[10] M. Reed and B. Simon. *Methods of modern mathematical physics. I.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, second edition, 1980. Functional analysis.

[11] M. Schlather. Introduction to positive definite functions and to unconditional simulation of random fields. Technical Report ST-99-10, Lancaster University, UK, 1999.

[12] I.J. Schoenberg. Metric spaces and completely monotone functions. *Ann. of Math. (2)*, 39(4):811–841, 1938.

[13] Ch. Schwab. Lecture notes on Numerical Analysis of Stochastic PDEs. Department of Mathematics, ETH Zürich, 2009.

[14] Ch. Schwab and R.A. Todor. Sparse finite elements for elliptic problems with stochastic loading. *Numer. Math.*, 95(4):707–734, 2003.

[15] Ch. Schwab and R.A. Todor. Sparse finite elements for stochastic elliptic problems - higher order moments. *Computing*, 71:43–63, 2003.

[16] T. von Petersdorff and Ch. Schwab. Sparse finite element methods for operator equations with stochastic data. *Applications of Mathematics*, 52:145–180, 2006.

Research Reports

No.	Authors/Title
10-08	*B. Pentenrieder and C. Schwab* *hp*-FEM for second moments of elliptic PDEs with stochastic data. Part 1: Analytic regularity
10-07	*C. Jerez-Hanckes and J.-C. Nédélec* Asymptotics for Helmholtz and Maxwell solutions in 3-D open waveguides
10-06	*C. Schwab and O. Reichmann* Numerical analysis of additive, Lévy and Feller processes with applications to option pricing
10-05	*C. Schwab and R. Stevenson* Fast evaluation of nonlinear functionals of tensor product wavelet expansions
10-04	*B.N. Khoromskij and C. Schwab* Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs
10-03	*A. Cohen, R. DeVore and C. Schwab* Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs
10-02	*V. Gradinaru, G.A. Hagedorn, A. Joye* Tunneling dynamics and spawning with adaptive semi-classical wave-packets
10-01	*N. Hilber, S. Kehtari, C. Schwab and C. Winter* Wavelet finite element method for option pricing in highdimensional diffusion market models
09-41	*E. Kokiopoulou, D. Kressner, N. Paragios, P. Frossard* Optimal image alignment with random projections of manifolds: algorithm and geometric analysis
09-40	*P. Benner, P. Ezzatti, D. Kressner, E.S. Quintana-Ortí, A. Remón* A mixed-precision algorithm for the solution of Lyapunov equations on hybrid CPU-GPU platforms
09-39	*V. Wheatley, P. Huguenot, H. Kumar* On the role of Riemann solvers in discontinuous Galerkin methods for magnetohydrodynamics
09-38	*E. Kokiopoulou, D. Kressner, N. Paragios, P. Frossard* Globally optimal volume registration using DC programming