Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics SAM

Multilevel approximation of Gaussian random fields: fast simulation

L. Herrmann and K. Kirchner and Ch. Schwab

Research Report No. 2019-32

June 2019
Latest revision: October 2019

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

Mathematical Models and Methods in Applied Sciences

Multilevel Approximation of Gaussian Random Fields: Fast Simulation

Lukas Herrmann*, Kristin Kirchner†, and Christoph Schwab‡

Seminar for Applied Mathematics, ETH Zürich,
Rämistrasse 101, CH-8092 Zürich, Switzerland
**lukas.herrmann@sam.math.ethz.ch*
†*kristin.kirchner@sam.math.ethz.ch*
‡*christoph.schwab@sam.math.ethz.ch*

We propose and analyze several multilevel algorithms for the fast simulation of possibly non-stationary Gaussian random fields (GRFs) indexed, e.g., by the closure of a bounded domain $\mathcal{D} \subset \mathbb{R}^n$ or, more generally, by a compact metric space $\mathcal{X}$ such as a compact n-manifold $\mathcal{M}$. A *colored* GRF $\mathcal{Z}$, admissible for our algorithms, solves the stochastic fractional-order equation $\mathcal{A}^\beta \mathcal{Z} = \mathcal{W}$ for some $\beta > n/4$, where $\mathcal{A}$ is a linear, local, second-order elliptic self-adjoint differential operator in divergence form and $\mathcal{W}$ is white noise on $\mathcal{X}$. We thus consider GRFs on $\mathcal{X}$ with covariance operators of the form $\mathcal{C} = \mathcal{A}^{-2\beta}$.

The proposed algorithms numerically approximate samples of $\mathcal{Z}$ on nested sequences $\{\mathcal{T}_\ell\}_{\ell \geq 0}$ of regular, simplicial partitions $\mathcal{T}_\ell$ of $\mathcal{D}$ and $\mathcal{M}$, respectively. Work and memory to compute one approximate realization of the GRF $\mathcal{Z}$ on the triangulation $\mathcal{T}_\ell$ of $\mathcal{X}$ with consistency $\mathcal{O}(N_\ell^{-\rho})$, for some consistency order $\rho > 0$, scale essentially linear in $N_\ell = \#(\mathcal{T}_\ell)$, independent of the possibly low regularity of the GRF. The algorithms are based on a sinc quadrature for an integral representation of (the application of) the negative fractional-order elliptic "coloring" operator $\mathcal{A}^{-\beta}$ to white noise $\mathcal{W}$. For the proposed numerical approximation, we prove bounds of the computational cost and the consistency error in various norms.

Keywords: Gaussian random fields, Matérn covariances, spatial statistics, fractional operators, multilevel methods.

AMS Subject Classification: 35S15, 65C30, 65C60, 65N12, 65N30.

1. Introduction

1.1. *Gaussian random fields in computational uncertainty quantification*

Gaussian and transformed Gaussian random fields (GRFs) are widely used to model spatial or spatiotemporal real-world phenomena with uncertainty. We first mention their relevance for geosciences such as for particulate matter concentration, temperature distributions, precipitation, and subsurface flow.[13,27,57] Further applications include, e.g., biomedical imaging,[55] astrophysics,[4] and material models in computational mechanics.[63]

The efficient, approximate sampling of GRFs on unstructured meshes of various physical domains is a key ingredient for several computational methodologies in

spatial statistics and data science; it therefore has received increasing attention in recent years.[19,20,43,44,53,54] In addition, such samples are needed, e.g., in Markov chain Monte Carlo algorithms for Bayesian estimation or filtering subject to PDE constraints.[27,58]

In computational spatial statistics or computational uncertainty quantification (UQ), a GRF $\mathcal{Z}$, defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with expectation operator $\mathbb{E}$, typically needs to be sampled at only finitely many locations in a "physical domain" $\mathcal{X}$ (e.g., a bounded Euclidean domain $\mathcal{D} \subset \mathbb{R}^n$ or a closed, connected, orientable, smooth, compact 2-surface $\mathcal{M} \subset \mathbb{R}^3$ such as the 2-sphere $\mathbb{S}^2$) and not on the entire domain of its definition (e.g., $\mathbb{R}^n$ or $\mathcal{M}$). In this case, if the covariance function

$$\varrho(x, x') := \mathbb{E}\left[(\mathcal{Z}(x) - \mathbb{E}[\mathcal{Z}(x)])(\mathcal{Z}(x') - \mathbb{E}[\mathcal{Z}(x')])\right], \qquad x, x' \in \mathcal{X},$$

of $\mathcal{Z}$ is explicitly available, a direct approach to numerically simulate the GRF $\mathcal{Z}$ at N locations $x_1, \dots, x_N \in \mathcal{X}$ is to assemble the covariance matrix $\mathbf{C} \in \mathbb{R}^{N \times N}$ with entries $C_{ij} := \varrho(x_i, x_j)$. A sample of the random vector $(\mathcal{Z}(x_1), \dots, \mathcal{Z}(x_N))^\top$ can then be obtained by computing the square root $\sqrt{\mathbf{C}}$ of $\mathbf{C}$ and by multiplying the matrix $\sqrt{\mathbf{C}}$ with a realization $\mathbf{y}(\omega)$ of the random vector $\mathbf{y} = (y_1, \dots, y_N)^\top$ consisting of N i.i.d. standard normally distributed random variables. This direct approach, however, entails high computational effort, because the covariance matrix $\mathbf{C}$ is, in general, dense so that calculating the square root requires $\mathcal{O}(N^3)$ operations and additional $\mathcal{O}(N^2)$ operations are needed for the matrix-vector multiplication.

Reducing this computational complexity has been subject of recent research in both areas, computational statistics and numerical mathematics, and many different approaches have been suggested.

1.2. *Existing methods*

A widely used computational method is the *circulant embedding* of the covariance matrix $\mathbf{C}$.[14,26,33] Here, *stationarity*, i.e., invariance of the Gaussian distribution under translation in $\mathbb{R}^n$ enters in an essential fashion: It facilitates the applicability of the fast Fourier transform (FFT) to numerically approximate the square root $\sqrt{\mathbf{C}}$ and to perform the matrix-vector multiplication $\sqrt{\mathbf{C}}\mathbf{y}(\omega)$ for a collection of N sampling nodes at a computational cost of $\mathcal{O}(N \log(N))$ operations per sample. For many applications, stationarity severely limits the scope of modeling: spatial random data often does not satisfy tests for stationarity.[9] For a set $\mathcal{X}$ that does not afford invariance under translations, FFT-based methods require rather specific, other invariances of $\mathcal{X}$ to achieve linear scaling, e.g., isotropic GRFs on $\mathcal{X} = \mathbb{S}^2$.[19]

Furthermore, in order to apply such methods for the simulation of a GRF in a Euclidean domain $\mathcal{D} \subset \mathbb{R}^n$, as proposed in Ref. 26, the sampling locations $x_1, \dots, x_N \in \mathcal{D}$ have to be situated on a uniform grid. For the finite element discretization of a PDE with random coefficient or forcing, one typically needs to simulate a GRF on a triangulation $\mathcal{T}$ of $\mathcal{D}$ with local mesh refinement. Sampling based

on circulant embedding mandates to simulate the GRF on a uniform "background" mesh of $\mathcal{D}$ and to reinterpolate these samples to the unstructured triangulation $\mathcal{T}$.[33] The spatial resolution of this uniform "background" mesh therefore must equal the *smallest* mesh width of $\mathcal{T}$. This may obviate possible complexity gains from mesh adaptation in the physical domain.

Sampling algorithms based on *hierarchical matrices* ($\mathcal{H}$-matrices) overcome these limitations at the cost of requiring the covariance function ϱ to be asymptotically smooth.[15,28,29] In this case, it has been shown[29] that combining an approximation of the covariance matrix $\mathbf{C}$ in the $\mathcal{H}$-matrix format with an iterative algorithm for realizing the matrix-vector product $\sqrt{\mathbf{C}}\mathbf{y}(\omega)$ can reduce the computational effort for generating one sample of the field $\mathcal{Z}$ to an optimal cost of essentially $\mathcal{O}(N)$ operations. In this work we pursue a different idea which has its roots in computational spatial statistics,[46] and which will allow us to formulate an efficient sampling algorithm for GRFs which neither are stationary nor have asymptotically smooth covariance functions. To put the results of this article in context, in the following we first present the key idea formulated in Ref. 46 as well as related developments since then. Subsequently, we delineate our novel contributions.

The *stochastic partial differential equation approach* formulated in Ref. 46 exploits the following observation going back to Ref. 64: In the case that $\mathcal{D} = \mathbb{R}^n$ is the whole Euclidean space, $\mathcal{W}$ denotes Gaussian white noise on the Hilbert space $L_2(\mathcal{D})$, and $\kappa > 0$ is constant, then the solution $\mathcal{Z}$ to the fractional-order stochastic partial differential equation (SPDE)

$$(\kappa^2 - \Delta)^\beta \mathcal{Z} = \mathcal{W}, \qquad \text{a.e. in } \mathcal{D}, \qquad \mathbb{P}\text{-a.s.}, \tag{1.1}$$

has a stationary covariance function ϱ of Matérn type, given for $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ by

$$\varrho(\mathbf{x}, \mathbf{x}') = \varrho_0(\|\mathbf{x} - \mathbf{x}'\|), \qquad \varrho_0(r) := \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa r)^\nu K_\nu(\kappa r). \tag{1.2}$$

Here, $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^n$ and Γ, K_ν denote the gamma function and a modified Bessel function of the second kind, respectively. The marginal variance in (1.2) is given by $\sigma^2 = \Gamma(2\beta - n/2)\Gamma(2\beta)^{-1}(4\pi)^{-n/2}\kappa^{n-4\beta}$ and the parameter ν, related to the fractional exponent β in (1.1) via $\nu = 2\beta - n/2$, determines the smoothness of the Matérn field. With the objective of modeling non-stationary effects, in Ref. 46 the SPDE (1.1) has been considered on a bounded Euclidean domain $\mathcal{D} \subsetneq \mathbb{R}^n$, where the differential operator $\kappa^2 - \Delta$ is augmented with Neumann boundary conditions. Then, non-stationarity enters in two ways: firstly, a *(generalized) Whittle–Matérn field* with spatially varying correlation range can be obtained via a function-valued parameter $\kappa\colon \mathcal{D} \to \mathbb{R}$; secondly, for $\mathcal{D} \subsetneq \mathbb{R}^n$, the solution $\mathcal{Z}$ of (1.1) with Neumann or Dirichlet boundary conditions is, in general, non-stationary, even if $\kappa > 0$ is constant. Furthermore, note that the covariance operator $\mathcal{C}\colon L_2(\mathcal{D}) \to L_2(\mathcal{D})$ of the GRF $\mathcal{Z}$ is known, namely $\mathcal{C} = (\kappa^2 - \Delta)^{-2\beta}$, but the covariance function ϱ of $\mathcal{Z}$ is usually not explicitly available in this case.

In Refs. 7, 8 the numerical solution of fractional-order stochastic partial differential equations including (1.1) has been considered in the general framework of

the stochastic (pseudodifferential) equation $\mathcal{A}^\beta \mathcal{Z} = \mathcal{W}$, holding $\mathbb{P}$-almost surely in a separable Hilbert space H (here, $\mathcal{W}$ denotes white noise on H). Based on a sinc quadrature for the Balakrishnan integral representation of the fractional-order inverse $\mathcal{A}^{-\beta}$, for the first time a numerical approximation has been proposed, which is computable for *the whole range of admissible parameters* $\beta > n/4$. If $\mathcal{A}$ is a linear, symmetric, second-order elliptic differential operator in divergence form and $H := L_2(\mathcal{X}, \nu_\mathcal{X})$ for the closure of some spatial domain $\mathcal{X}$, this approach can be realized in practice by applying only inverses of discretized *local* linear operators to a consistent numerical realization of white noise. The method is formulated with respect to a subspace $V_h \subset H$ of finite dimension N_h (here and throughout, $h > 0$ shall denote the discretization parameter which could indicate, for instance, the mesh width in a finite element method). The corresponding random load vector becomes a centered, N_h-variate Gaussian distributed random vector, whose covariance matrix corresponds to the Gramian (aka "mass matrix") $\mathbf{M}$ of a basis for V_h. This numerical approximation has been justified by a rigorous error analysis providing explicit convergence rates for (i) the strong mean-square error with respect to H,[7] and (ii) the weak error.[8] For the case of a bounded Euclidean domain, $\mathcal{X} = \overline{\mathcal{D}}$, the error analysis of Refs. 7,8 recently has been complemented by bounds for the strong mean-square error with respect to the (fractional-order) Sobolev space $H^\sigma(\mathcal{D})$ and the Hölder space $C^\gamma(\overline{\mathcal{D}})$ for $\sigma \in [0,1]$ and $\gamma \in (0,1)$.[18] Furthermore, rates of convergence for the covariance function of the approximation have been derived.[18]

1.3. *Present work*

Until now the following computational aspects of the approximation introduced in Ref. 7 have remained open questions: (I) Assuming an asymptotic target accuracy $O(N_h^{-\rho})$ (in strong mean-square L_2-sense) for some convergence rate $\rho > 0$, generating one sample of $\mathcal{Z}$ via the approach of Ref. 7 will require the numerical solution of $\mathcal{O}\big(\log^2(N_h)\big)$ linear systems of size N_h, out of which some correspond to singularly perturbed problems. For this reason, the applicability and complexity of multilevel preconditioning methods for solving these linear systems is unclear. (II) The problem of calculating $\sqrt{\mathbf{C}}$ for the dense covariance matrix $\mathbf{C}$ has been reduced to the computation of the matrix square root $\sqrt{\mathbf{M}}$ needed to simulate the random load vector. Even though the Gramian $\mathbf{M}$ usually is sparse, it is not obvious how to numerically realize $\sqrt{\mathbf{M}}$ in order to minimize the computational cost.

The novel contributions of the present work are to address the problems (I)–(II) above in detail. Theorem 2.1 bounds under minimal assumptions on the operator $\mathcal{A}$ and under certain conditions on the iterative solver involved (see Assumption 2.7) the consistency error when approximating the GRF $\mathcal{Z}$ on a general compact metric space $\mathcal{X}$ via the approach of Ref. 7. In Propositions 3.2, 4.1, and 5.1 it is shown that these conditions hold on bounded Euclidean domains, $\mathcal{X} = \overline{\mathcal{D}}$, and on closed, connected, orientable, smooth, compact 2-surfaces, $\mathcal{X} = \mathcal{M}$, for finite element and wavelet methods combined with BPX[65] and diagonal preconditioning, respectively.

These multilevel techniques are combined with the approach formulated in Subsection 2.2.4 to approximate the action of the matrix square root $\sqrt{\mathbf{M}}$. Our main result is that we are thus able to generate one approximate sample of the (generalized) Whittle–Matérn field $\mathcal{Z}$ at a total cost of $\mathcal{O}\big(N_h \log^a(N_h)\big)$ operations for some $a \geq 2$ and at target accuracy of $\mathcal{O}(N_h^{-\rho})$ for a convergence order $\rho > 0$ admitted by regularity of $\mathcal{Z}$ and dimensionality of $\mathcal{X}$. The presented algorithms allow, e.g., for non-stationary GRFs in the Euclidean case $\mathcal{X} = \overline{\mathcal{D}}$, as well as for GRFs on the sphere $\mathcal{X} = \mathbb{S}^2$ which are non-isotropic, so that two-point correlations are not rotation-invariant.

1.4. *Outline*

After a brief comment on notation in Subsection 1.5, we introduce GRFs indexed by a compact metric space $\mathcal{X}$ in Section 2. For this general setting, we specify admissible covariance operators of GRFs on $\mathcal{X}$ in Subsection 2.1. In Subsection 2.2 we then recall the sinc-Galerkin approximation of Refs. 7, 8 and we furthermore address the numerical matrix-vector multiplication for the matrix square root of the Gramian, needed for simulating the random forcing. Subsection 2.3 is devoted to a unified numerical analysis of (i) the discretization error induced by the sinc-Galerkin approximation and (ii) the consistency error, assuming that the arising linear systems are approximately solved by an iterative solver which is robust under singular perturbations.

Subsequently, we explicitly formulate sinc-Galerkin approximations based on finite element discretizations on bounded Lipschitz polytopes $\mathcal{X} = \overline{\mathcal{D}} \subset \mathbb{R}^n$ in Section 3 and on closed, connected, orientable, smooth, compact 2-surfaces $\mathcal{X} = \mathcal{M} \subset \mathbb{R}^3$ in Section 4. For the efficient solution of the arising second-order, local, elliptic boundary value problems, we propose multilevel preconditioning algorithms on hierarchies of unstructured, regular simplicial triangulations. We prove their essentially linear complexity, uniformly with respect to singular perturbations originating from the sinc quadrature, which allows us to apply the unified error analysis of Subsection 2.3. The results of Sections 3–4 are based on BPX preconditioning and cover unstructured, shape-regular triangulations of $\mathcal{D}$ and $\mathcal{M}$, respectively. Section 5 develops the alternative approach of *wavelet preconditioning*. While providing a unified framework for optimal preconditioning on $\mathcal{X} \in \{\overline{\mathcal{D}}, \mathcal{M}\}$, it is based on the assumed availability of wavelet bases for Galerkin discretizations of the parametric elliptic problems stemming from the sinc quadrature. Their construction commonly requires hierarchic triangulations with some degree of regularity, whereas the BPX techniques are readily applicable also for discretizations which are based on unstructured, nested grids.

In Section 6 several numerical experiments for convex and non-convex polygons $\mathcal{D} \subset \mathbb{R}^2$ and the sphere $\mathcal{M} = \mathbb{S}^2$ are performed, which verify our theoretical results. Section 7 collects the principal contributions of this work, and indicates several generalizations and extensions which follow from the present results.

1.5. *Notation*

If not specified otherwise, E^* denotes the dual space of a Banach space E and $(\cdot\,,\cdot)_H$ is the inner product on a Hilbert space H. The domain of a linear operator $\mathcal{A}$, which is possibly unbounded on H, is denoted by $\mathscr{D}(\mathcal{A}) \subseteq H$.

Throughout, $(\Omega, \mathcal{F}, \mathbb{P})$ denotes a complete probability space with expectation operator $\mathbb{E}$. We mark equations which hold almost everywhere or $\mathbb{P}$-almost surely with a.e. and $\mathbb{P}$-a.s., respectively. For a Banach space E and $p \in (0,\infty)$, $L_p(\Omega; E)$ denotes the space of (equivalence classes of) E-valued random variables with finite p-th moment, equipped with the norm (respectively, quasi-norm if $p \in (0,1)$)

$$\|X\|_{L_p(\Omega;E)} := \big(\mathbb{E}\big[\|X\|_E^p\big]\big)^{1/p}, \qquad X \in L_p(\Omega; E).$$

For a compact, connected metric space $\mathcal{X}$, $\mathcal{B}(\mathcal{X})$ is the Borel σ-algebra on $\mathcal{X}$. If, in addition, $\nu_{\mathcal{X}}$ is a Borel measure on $\mathcal{X}$ then $L_p(\mathcal{X}, \nu_{\mathcal{X}})$ is the corresponding L_p-space of p-integrable real-valued functions, with norm

$$\|f\|_{L_p(\mathcal{X},\nu_{\mathcal{X}})} := \left(\int_{\mathcal{X}} |f(x)|^p \,\mathrm{d}\nu_{\mathcal{X}}(x)\right)^{1/p}, \qquad 1 \le p < \infty,$$

and with the usual modification for $p = \infty$.

For a Euclidean domain $\mathcal{D} \subset \mathbb{R}^n$, $\overline{\mathcal{D}}$ is its closure, and the L_p-space with respect to the Lebesgue measure is abbreviated by $L_p(\mathcal{D})$. Furthermore, for $\sigma > 0$, $H^\sigma(\mathcal{D})$ denotes the (integer- or fractional-order) Sobolev space.

For a closed, connected, orientable, smooth, compact 2-surface $\mathcal{M}$, the set $OPS^r_{1,0}(\mathcal{M})$ comprises all pseudodifferential operators of order $r \in \mathbb{R}$ in the sense of Def. II.1.4 in Ref. 62.

2. Gaussian random fields on compact metric spaces

2.1. *Admissible fields*

Let us first assume that $\mathcal{Z}$ is a centered (i.e., $\mathbb{E}[\mathcal{Z}] = 0$) GRF on a compact, connected metric space $(\mathcal{X}, d_{\mathcal{X}})$ and let $\nu_{\mathcal{X}}$ be a strictly positive and finite Borel measure on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. The covariance operator of the GRF $\mathcal{Z}$, denoted by $\mathcal{C}\colon L_2(\mathcal{X}, \nu_{\mathcal{X}}) \to L_2(\mathcal{X}, \nu_{\mathcal{X}})$, is then defined via

$$(\mathcal{C}\phi, \psi)_{L_2(\mathcal{X},\nu_{\mathcal{X}})} = \mathbb{E}\left[(\mathcal{Z}, \phi)_{L_2(\mathcal{X},\nu_{\mathcal{X}})}(\mathcal{Z}, \psi)_{L_2(\mathcal{X},\nu_{\mathcal{X}})}\right] \quad \forall \phi, \psi \in L_2(\mathcal{X}, \nu_{\mathcal{X}}).$$

This integral operator is self-adjoint and compact on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$. Furthermore, while being in general only positive semi-definite, the covariance operators considered in this work exhibit positive definiteness.

More precisely, we focus on the following class of admissible covariance structures: We consider GRFs, whose precision operator $\mathcal{C}^{-1}$ is a linear operator $\mathcal{A}$ as specified in the assumption below or—motivated by the example (1.1) of Whittle–Matérn fields—more generally, by a fractional power of such an operator.

Assumption 2.1. The linear operator $\mathcal{A}\colon \mathscr{D}(\mathcal{A}) \subseteq L_2(\mathcal{X}, \nu_{\mathcal{X}}) \to L_2(\mathcal{X}, \nu_{\mathcal{X}})$ is

densely defined, self-adjoint, and positive definite. Furthermore, it has an inverse $\mathcal{A}^{-1}\colon L_2(\mathcal{X},\nu_{\mathcal{X}}) \to L_2(\mathcal{X},\nu_{\mathcal{X}})$ which is compact.

Assuming compactness of $\mathcal{A}^{-1}$ ensures the existence of a countable, orthonormal basis $\{e_j\}_{j\in\mathbb{N}}$ of $L_2(\mathcal{X},\nu_{\mathcal{X}})$, consisting of (equivalence classes of) eigenvectors of $\mathcal{A}$. We let them be enumerated such that the corresponding eigenvalues $\{\lambda_j\}_{j\in\mathbb{N}} \subset \mathbb{R}$ of $\mathcal{A}$ are in nondecreasing order, i.e.,

$$0 < \lambda_1 \le \lambda_2 \le \ldots \le \lambda_j \le \lambda_{j+1} \le \ldots, \qquad \lim_{j\to\infty} \lambda_j = \infty,$$

and repeated counting multiplicity.

For any $\sigma > 0$, the fractional power operator $\mathcal{A}^{\sigma/2}$ is well-defined on the domain

$$\mathscr{D}\big(\mathcal{A}^{\sigma/2}\big) := \left\{ \psi \in L_2(\mathcal{X},\nu_{\mathcal{X}}) : \sum_{j\in\mathbb{N}} \lambda_j^{\sigma} (\psi, e_j)^2_{L_2(\mathcal{X},\nu_{\mathcal{X}})} < \infty \right\}$$

via the spectral representation $\mathcal{A}^{\sigma/2}\psi := \sum_{j\in\mathbb{N}} \lambda_j^{\sigma/2} (\psi, e_j)_{L_2(\mathcal{X},\nu_{\mathcal{X}})}\, e_j$. For $\sigma > 0$, we thus introduce the subspace

$$\big(\dot{H}^{\sigma}_{\mathcal{A}}, (\,\cdot\,,\,\cdot\,)_{\sigma}\big), \qquad \dot{H}^{\sigma}_{\mathcal{A}} := \mathscr{D}\big(\mathcal{A}^{\sigma/2}\big) \subseteq L_2(\mathcal{X},\nu_{\mathcal{X}}), \tag{2.1}$$

which is equipped with the inner product

$$(\phi,\psi)_{\sigma} := \big(\mathcal{A}^{\sigma/2}\phi, \mathcal{A}^{\sigma/2}\psi\big)_{L_2(\mathcal{X},\nu_{\mathcal{X}})} = \sum_{j\in\mathbb{N}} \lambda_j^{\sigma} (\phi, e_j)_{L_2(\mathcal{X},\nu_{\mathcal{X}})} (\psi, e_j)_{L_2(\mathcal{X},\nu_{\mathcal{X}})},$$

and the induced norm $\|\,\cdot\,\|_{\sigma}$. Note that $(\,\cdot\,,\,\cdot\,)_{\sigma}$ renders $\dot{H}^{\sigma}_{\mathcal{A}}$ a Hilbert space. In addition, we define $\dot{H}^{0}_{\mathcal{A}} := L_2(\mathcal{X},\nu_{\mathcal{X}})$ and, for $\sigma > 0$, we let $\dot{H}^{-\sigma}_{\mathcal{A}}$ denote the dual space $\big(\dot{H}^{\sigma}_{\mathcal{A}}\big)^*$ after identification via the inner product on $L_2(\mathcal{X},\nu_{\mathcal{X}})$ (continuously extended as duality pairing). In this way, we obtain, for $0 \le \sigma_0 \le \sigma_1$, the following chain of continuous and dense embeddings

$$\dot{H}^{\sigma_1}_{\mathcal{A}} \hookrightarrow \dot{H}^{\sigma_0}_{\mathcal{A}} \hookrightarrow \dot{H}^{0}_{\mathcal{A}} = L_2(\mathcal{X},\nu_{\mathcal{X}}) \cong L_2(\mathcal{X},\nu_{\mathcal{X}})^* \hookrightarrow \dot{H}^{-\sigma_0}_{\mathcal{A}} \hookrightarrow \dot{H}^{-\sigma_1}_{\mathcal{A}}.$$

We furthermore recall from [48, Thm. 4.36] that, for all $0 \le \sigma_0 < \sigma_1$ and every $\theta \in [0,1]$, the space $\dot{H}^{(1-\theta)\sigma_0+\theta\sigma_1}_{\mathcal{A}}$ is equivalent to the complex interpolation space $\big[\dot{H}^{\sigma_0}_{\mathcal{A}}, \dot{H}^{\sigma_1}_{\mathcal{A}}\big]_{\theta}$, i.e.,

$$\exists c_1 = c_1(\sigma_0,\sigma_1,\theta) > 0 : \quad \forall v \in \dot{H}^{\sigma_1}_{\mathcal{A}} : \quad \|v\|_{(1-\theta)\sigma_0+\theta\sigma_1} \le c_1 \|v\|_{\left[\dot{H}^{\sigma_0}_{\mathcal{A}}, \dot{H}^{\sigma_1}_{\mathcal{A}}\right]_{\theta}}, \tag{2.2}$$

$$\exists c_2 = c_2(\sigma_0,\sigma_1,\theta) > 0 : \quad \forall v \in \dot{H}^{\sigma_1}_{\mathcal{A}} : \quad \|v\|_{\left[\dot{H}^{\sigma_0}_{\mathcal{A}}, \dot{H}^{\sigma_1}_{\mathcal{A}}\right]_{\theta}} \le c_2 \|v\|_{(1-\theta)\sigma_0+\theta\sigma_1}. \tag{2.3}$$

We then consider the generalization of the Whittle–Matérn field (1.1) to a centered, possibly non-stationary GRF $\mathcal{Z}^{\beta}$, defined on the compact metric space $\mathcal{X}$, with distribution

$$\mathcal{Z}^{\beta} \sim \mathsf{N}(0, \mathcal{C}_{\beta}), \qquad \mathcal{C}_{\beta} := \mathcal{A}^{-2\beta}. \tag{2.4}$$

Here, we emphasize the dependence of the covariance structure on the fractional exponent by the index β. Note that any system of eigenvectors for $\mathcal{A}$ is one for the covariance operator $\mathcal{C}_{\beta}$ and, in particular, $\mathcal{C}_{\beta} e_j = \lambda_j^{-2\beta} e_j$. Furthermore, the

definition (2.1) of $\dot{H}^\sigma_{\mathcal{A}}$ renders $\mathcal{C}_\beta^{1/2} = \mathcal{A}^{-\beta}$ an isometric isomorphism as a mapping from $\dot{H}^\sigma_{\mathcal{A}}$ to $\dot{H}^{\sigma+2\beta}_{\mathcal{A}}$ for any $\sigma \in \mathbb{R}$.

This property and a Karhunen–Loève expansion of the GRF $\mathcal{Z}^\beta$ in terms of the eigenpairs $\{(\lambda_j, e_j)\}_{j\in\mathbb{N}}$ of the operator $\mathcal{A}$ can be exploited to show the following regularity results for $\mathcal{Z}^\beta$ in (2.4) with respect to the spaces $\big(\dot{H}^\sigma_{\mathcal{A}}\big)_{\sigma\geq 0}$ in (2.1).

Lemma 2.2. *Let $\mathcal{A}$ be an operator satisfying Assumption 2.1, $\beta > 0$, $p \in (0,\infty)$, $\sigma \geq 0$ and $\mathcal{Z}$ be a GRF with distribution* (2.4)*. Suppose furthermore that the eigenvalues $\{\lambda_j\}_{j\in\mathbb{N}}$ of $\mathcal{A}$ satisfy the spectral asymptotics*

$$\exists c_1, c_2 > 0 : \quad \forall j \in \mathbb{N} : \qquad c_1 j^\alpha \leq \lambda_j \leq c_2 j^\alpha, \tag{2.5}$$

for some $\alpha > 0$. Then, $\mathbb{E}\big[\|\mathcal{Z}^\beta\|_\sigma^p\big] < \infty$ if and only if $2\beta - \sigma > \alpha^{-1}$.

Proof. By Fubini's theorem and by (2.4), (2.5) we find

$$\mathbb{E}\left[\|\mathcal{Z}^\beta\|_\sigma^2\right] = \sum_{j\in\mathbb{N}} \lambda_j^\sigma (\mathcal{C}_\beta e_j, e_j)_0 = \sum_{j\in\mathbb{N}} \lambda_j^{-2\beta+\sigma} \leq c_1^{-(2\beta-\sigma)} \sum_{j\in\mathbb{N}} j^{-\alpha(2\beta-\sigma)},$$

and this series converges if and only if $\alpha(2\beta-\sigma) > 1$. By observing that similarly $c_2^{-(2\beta-\sigma)} \sum_{j\in\mathbb{N}} j^{-\alpha(2\beta-\sigma)} \leq \mathbb{E}\left[\|\mathcal{Z}^\beta\|_\sigma^2\right]$, this shows that $\mathcal{Z}^\beta$ is a zero-mean Gaussian random variable in $L_2(\Omega;\dot{H}^\sigma_{\mathcal{A}})$ if and only if $2\beta - \sigma > \alpha^{-1}$. By the Kahane–Khintchine inequalities [45, Thm. 4.7 and p. 103] this statement remains true for $L_p(\Omega;\dot{H}^\sigma_{\mathcal{A}})$ and an arbitrary $p \in (0,\infty)$. □

We close this subsection by observing that a centered GRF $\mathcal{Z}^\beta$ with distribution (2.4) can be represented as the solution to the stochastic equation

$$\mathcal{A}^\beta \mathcal{Z}^\beta = \mathcal{W}, \qquad \mathbb{P}\text{-a.s.} \tag{2.6}$$

Here, $\mathcal{W}$ denotes Gaussian white noise on $L_2(\mathcal{X},\nu_{\mathcal{X}})$, i.e., $\mathcal{W}$ is a weak random variable in the sense of [3, Ch. 6.4], with values in $\dot{H}^0_{\mathcal{A}} = L_2(\mathcal{X},\nu_{\mathcal{X}})$ and characteristic function given by $\dot{H}^0_{\mathcal{A}} \ni \psi \mapsto \mathbb{E}[\exp(i(\psi,\mathcal{W})_0)] = \exp(-\frac{1}{2}\|\psi\|_0^2)$.

2.2. *Sinc-Galerkin discretization*

The numerical solution of white noise driven, fractional-order equations of the form (2.6) for $\beta \in (0,1)$ has been analyzed in a general Hilbert space framework in Refs. 7, 8.

We adopt this approach, based on an integral representation of the negative fractional power $\mathcal{A}^{-\beta}$ due to Balakrishnan[2] and a quadrature for this integral with exponential convergence, first proposed in Ref. 10. This method reduces the application of $\mathcal{A}^{-\beta}$ to the numerical realization of finitely many inverses of *local* second-order elliptic operators. In the following, we clearly exhibit the structure of these local operators, since this will be key to the multilevel preconditioning techniques.

2.2.1. *Integral representation and sinc quadrature*

A linear operator $\mathcal{A}$ satisfying Assumption 2.1 is closed. Therefore, for $\beta \in (0,1)$, the negative fractional power $\mathcal{A}^{-\beta}$ of the operator $\mathcal{A}$ admits the *Balakrishnan representation*[2]

$$\mathcal{A}^{-\beta} = \frac{c_\beta}{2} \int_0^\infty t^{-\beta} (\mathcal{A} + t\mathcal{I})^{-1} \, \mathrm{d}t = c_\beta \int_{-\infty}^\infty e^{-2\beta y} \left(e^{-2y}\mathcal{A} + \mathcal{I}\right)^{-1} \, \mathrm{d}y, \tag{2.7}$$

where $c_\beta := \pi^{-1}(2\sin(\pi\beta))$ and $\mathcal{I}$ denotes the identity on $L_2(\mathcal{X}, \nu_\mathcal{X})$.

In Ref. 10 it has been proven that the latter integral in (2.7) fulfills the requirements for convergence with respect to the operator norm when approximated by a sinc quadrature (see, e.g., Refs. 49,60). For $K \in \mathbb{N}$ and $k \in \{-K, \dots, K\}$, we define the following quadrature parameters:

$$\delta_y := 1/\sqrt{K} > 0, \qquad y_k := k\delta_y, \qquad \varepsilon_k := \exp(-y_k). \tag{2.8}$$

If $f \in L_2(\mathcal{X}, \nu_\mathcal{X})$ and $\beta \in (0,1)$, the *sinc approximation* u_K of $u = \mathcal{A}^{-\beta} f$ with $\mathcal{A}^{-\beta}$ as represented in (2.7) then reads

$$u_K := Q_K^{-\beta}(\mathcal{A}) f, \qquad Q_K^{-\beta}(\mathcal{A}) := c_\beta \delta_y \sum_{|k| \le K} \varepsilon_k^{2\beta} \left(\varepsilon_k^2 \mathcal{A} + \mathcal{I}\right)^{-1}. \tag{2.9}$$

In [10, Lem. 3.4] exponential convergence for the sinc quadrature (2.9) (with $\mathcal{A}$ replaced with an element λ of the positive real-valued spectrum of $\mathcal{A}$) is proven. The next lemma is an immediate consequence of that result.

Lemma 2.3. *Let $\mathcal{A}\colon \mathscr{D}(\mathcal{A}) \subseteq L_2(\mathcal{X}, \nu_\mathcal{X}) \to L_2(\mathcal{X}, \nu_\mathcal{X})$ satisfy Assumption 2.1, $\beta \in (0,1)$ and $Q_K^{-\beta}(\mathcal{A})$ be the sinc quadrature in* (2.9)*. Then, for every $0 \le \sigma \le 1$, there exists a constant $C > 0$, independent of K, such that, for all $f \in \dot{H}^\sigma_\mathcal{A}$,*

$$\left\| \mathcal{A}^{-\beta} f - Q_K^{-\beta}(\mathcal{A}) f \right\|_\sigma \le C e^{-2\min\{\beta, 1-\beta\}\sqrt{K}} \, \|f\|_\sigma$$

holds with respect to the norm $\|\cdot\|_\sigma$ on $\dot{H}^\sigma_\mathcal{A}$, see (2.1)*.*

2.2.2. *Galerkin discretization*

For $f \in L_2(\mathcal{X}, \nu_\mathcal{X})$ and $K \in \mathbb{N}$, we can write the sinc quadrature approximation u_K from (2.9) as

$$u_K = Q_K^{-\beta}(\mathcal{A}) f = c_\beta \delta_y \sum_{|k| \le K} \varepsilon_k^{2\beta} u^k, \tag{2.10}$$

i.e., as a linear combination of the $2K+1$ solutions to the non-fractional equations

$$\left(\varepsilon_k^2 \mathcal{A} + \mathcal{I}\right) u^k = f, \qquad k = -K, \dots, K. \tag{2.11}$$

Recall the Hilbert space $\left(\dot{H}^\sigma_\mathcal{A}, (\,\cdot\,,\,\cdot\,)_\sigma\right)$ from (2.1), and suppose that we are given a family $(V_h)_{h>0}$ of finite-dimensional subspaces,

$$V_h \subset \dot{H}^1_\mathcal{A}, \qquad \dim(V_h) := N_h < \infty. \tag{2.12}$$

A V_h-valued Galerkin approximation for u_K in (2.10) is then based on weak formulations of the problems in (2.11). To state these, we first introduce, for $\eta > 0$, the parametric bilinear form

$$a_\eta \colon \dot{H}^1_{\mathcal{A}} \times \dot{H}^1_{\mathcal{A}} \to \mathbb{R}, \qquad a_\eta(w, v) := \eta(w, v)_1 + (w, v)_0, \tag{2.13}$$

where $(\,\cdot\,,\,\cdot\,)_1$ and $(\,\cdot\,,\,\cdot\,)_0$ are the inner products on $\dot{H}^1_{\mathcal{A}}$ and $\dot{H}^0_{\mathcal{A}} = L_2(\mathcal{X}, \nu_{\mathcal{X}})$, respectively, cf. (2.1). In addition, we let

$$\|v\|_{a,\eta} := \sqrt{a_\eta(v, v)}, \qquad v \in \dot{H}^1_{\mathcal{A}}, \tag{2.14}$$

be the corresponding η-dependent energy norm on $\dot{H}^1_{\mathcal{A}}$. Note that, for any $\eta > 0$, the norms $\|\cdot\|_{a,\eta}$ and $\|\cdot\|_1$ are equivalent (with equivalence constants depending on η). The Galerkin discretization of (2.11) with respect to V_h then reads, find

$$u^k_h \in V_h : \quad a_{\varepsilon_k^2}(u^k_h, v_h) = (f, v_h)_0 \quad \forall v_h \in V_h, \quad k = -K, \dots, K. \tag{2.15}$$

Since $\mathcal{A} \colon \dot{H}^1_{\mathcal{A}} \to \dot{H}^{-1}_{\mathcal{A}}$ is bounded, symmetric and positive definite, the Lax–Milgram lemma ensures that the numerical approximations $\{u^{-K}_h, \dots, u^K_h\} \subset V_h$ in (2.15) are well-defined. Furthermore, we have *optimality* with respect to the energy norm,

$$\|u^k - u^k_h\|_{a,\varepsilon_k^2} = \inf_{w_h \in V_h} \|u^k - w_h\|_{a,\varepsilon_k^2}, \qquad k = -K, \dots, K. \tag{2.16}$$

Finally, by exploiting the Galerkin approximations from (2.15) in the sinc quadrature (2.10), we arrive at the *sinc-Galerkin approximation* $u_{K,h}$ for $u = \mathcal{A}^{-\beta} f$,

$$u_{K,h} := c_\beta \delta_y \sum_{|k| \le K} \varepsilon_k^{2\beta} u^k_h,$$

with the sinc parameters $\delta_y, \varepsilon_k > 0$ from (2.8). We note that

$$u_{K,h} = c_\beta \delta_y \sum_{|k| \le K} \varepsilon_k^{2\beta} \left(\varepsilon_k^2 \mathcal{A}_h + \mathcal{I}_h\right)^{-1} f = Q_K^{-\beta}(\mathcal{A}_h) f, \tag{2.17}$$

where $\mathcal{I}_h$ denotes the identity on V_h and $\mathcal{A}_h \colon V_h \to V_h$ is the Galerkin discretization of the operator $\mathcal{A}$. Thus, the Galerkin discretization for the sinc quadrature equals the sinc quadrature formulated for the Galerkin operator $\mathcal{A}_h$.

2.2.3. *Fully discrete approximation for $\beta \in (0, 1)$*

Besides an approximation of the negative fractional power operator $\mathcal{A}^{-\beta}$ we need a ($\mathbb{P}$-a.s.) V_h-valued approximation of white noise $\mathcal{W}$ on the Hilbert space $L_2(\mathcal{X}, \nu_{\mathcal{X}})$ in order to solve (2.6) numerically for the GRF $\mathcal{Z}^\beta$.

To this end, let $\Pi_h \colon L_2(\mathcal{X}, \nu_{\mathcal{X}}) \to V_h$ denote the $L_2(\mathcal{X}, \nu_{\mathcal{X}})$-orthogonal projection onto the finite-dimensional subspace $V_h \subset \dot{H}^1_{\mathcal{A}}$. Although Gaussian white noise on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$ is only $\dot{H}^{-\frac{1}{\alpha}-\varepsilon}_{\mathcal{A}}$-regular for $\alpha > 0$ as in (2.5) and $\varepsilon > 0$ ($\mathbb{P}$-a.s. and in

L_p-sense, $p \in (0,\infty)$, see [7, Prop. 2.3]), the Gaussian random variable $\mathcal{W}_h := \Pi_h \mathcal{W}$ is a well-defined element of $L_p(\Omega; L_2(\mathcal{X}, \nu_\mathcal{X}))$ as shown in the following lemma.

Lemma 2.4. *Let the operator $\mathcal{A}$ satisfy Assumption 2.1 and $\Pi_h \colon L_2(\mathcal{X}, \nu_\mathcal{X}) \to V_h$ be the $L_2(\mathcal{X}, \nu_\mathcal{X})$-orthogonal projection onto $V_h \subset \dot{H}^1_\mathcal{A}$ with $\dim(V_h) = N_h$, and the space $\dot{H}^1_\mathcal{A}$ be defined as in (2.1). Then, for $\mathcal{W}_h := \Pi_h \mathcal{W}$, we have $\mathbb{E}\big[\|\mathcal{W}_h\|_0^2\big] = N_h$. Furthermore, for any $p \in (0,\infty)$, there exists a constant $C_p > 0$, depending only on p, such that $\big(\mathbb{E}\big[\|\mathcal{W}_h\|_0^p\big]\big)^{1/p} \leq C_p \sqrt{N_h}$ holds.*

Proof. Let $\{(\lambda_{k,h}, e_{k,h})\}_{k=1}^{N_h}$ be eigenpairs of the Galerkin operator $\mathcal{A}_h \colon V_h \to V_h$, with $\{e_{k,h}\}_{k=1}^{N_h}$ orthonormalized in $L_2(\mathcal{X}, \nu_\mathcal{X})$ so that $\|e_{k,h}\|_0^2 = 1$. Then, for $p = 2$, we find by Fubini's theorem,

$$\mathbb{E}\left[\|\mathcal{W}_h\|_0^2\right] = \mathbb{E} \sum_{j \in \mathbb{N}} (\Pi_h \mathcal{W}, e_j)_0^2 = \sum_{j \in \mathbb{N}} \|\Pi_h e_j\|_0^2 = \sum_{k=1}^{N_h} \|e_{k,h}\|_0^2 = N_h.$$

For a general $p \in (0,\infty)$, the assertion of this lemma follows from the Kahane–Khintchine inequalities [45, Thm. 4.7 and p. 103]. □

Thus, for any $\beta \in (0,1)$, we may approximate the $L_2(\mathcal{X}, \nu_\mathcal{X})$-valued Gaussian random variable $\mathcal{Z}^\beta$ in (2.6), with distribution (2.4), by the GRF

$$\mathcal{Z}^\beta_{K,h} := Q_K^{-\beta}(\mathcal{A}_h)\mathcal{W}_h, \qquad \mathbb{P}\text{-a.s.}, \tag{2.18}$$

taking values in the finite-dimensional space V_h ($\mathbb{P}$-a.s.).

As mentioned already in Subsection 2.2.2 for the deterministic setting, computing (a sample of) $Q_K^{-\beta}(\mathcal{A}_h)\mathcal{W}_h$ requires to solve $2K+1$ integer-order problems. More precisely, the sinc-Galerkin approximation in (2.18) equals the linear combination

$$\mathcal{Z}^\beta_{K,h} = c_\beta \delta_y \sum_{|k| \leq K} \varepsilon_k^{2\beta} \mathcal{Z}^k_h, \qquad \mathbb{P}\text{-a.s.}, \tag{2.19}$$

where the random fields $\left\{\mathcal{Z}^{-K}_h, \ldots, \mathcal{Z}^K_h\right\}$ are solutions to the stochastic equations

$$\left(\varepsilon_k^2 \mathcal{A}_h + \mathcal{I}_h\right) \mathcal{Z}^k_h = \mathcal{W}_h, \qquad \mathbb{P}\text{-a.s.}, \qquad k = -K, \ldots, K. \tag{2.20}$$

These discrete problems in turn lead to $2K+1$ systems of linear equations of size N_h, namely

$$\left(\varepsilon_k^2 \mathbf{A} + \mathbf{M}\right) \mathbf{Z}^k_h = \mathbf{b}, \qquad k = -K, \ldots, K, \tag{2.21}$$

where the symmetric matrices $\mathbf{M}, \mathbf{A} \in \mathbb{R}^{N_h \times N_h}_{\text{sym}}$ are the Gramian and the matrix representation of the discrete operator $\mathcal{A}_h \colon V_h \to V_h$, which both depend on the choice of basis $\mathbf{\Phi}_h := \{\phi_{1,h}, \ldots, \phi_{N_h,h}\}$ for $V_h \subset \dot{H}^1_\mathcal{A}$. Their elements are given by (recall the inner product $(\,\cdot\,,\,\cdot\,)_\sigma$ on $\dot{H}^\sigma_\mathcal{A}$ from (2.1))

$$\mathbf{M}_{ij} := (\phi_{j,h}, \phi_{i,h})_0, \qquad \mathbf{A}_{ij} := (\mathcal{A}_h \phi_{j,h}, \phi_{i,h})_0 = (\phi_{j,h}, \phi_{i,h})_1, \tag{2.22}$$

for $i, j \in \{1, \dots, N_h\}$. The load vector $\mathbf{b}$ has entries $b_i := \big(\mathcal{W}_h, \phi_{i,h}\big)_0$. Thus, it is multivariate Gaussian distributed with $\mathbf{b} \sim \mathsf{N}(\mathbf{0}, \mathbf{M})$, since

$$\mathbb{E}[(\mathcal{W}_h, \phi_{i,h})_0 (\mathcal{W}_h, \phi_{j,h})_0] = \mathbb{E}[(\Pi_h \mathcal{W}, \phi_{i,h})_0 (\Pi_h \mathcal{W}, \phi_{j,h})_0] = (\phi_{i,h}, \phi_{j,h})_0.$$

The vector $\mathbf{Z}_K^\beta$ of coefficients for the sinc-Galerkin approximation $\mathcal{Z}_{K,h}^\beta$ in (2.18), (2.19) with respect to basis $\mathbf{\Phi}_h$ is then given by

$$\mathbf{Z}_K^\beta = c_\beta \delta_y \sum_{|k| \le K} \varepsilon_k^{2\beta} \mathbf{Z}_h^k, \qquad \mathbf{Z}_K^\beta \sim \mathsf{N}\left(\mathbf{0}, \mathbf{Q}_{\beta,K} \mathbf{M} \mathbf{Q}_{\beta,K}\right).$$

Here, $\mathbf{Q}_{\beta,K} \in \mathbb{R}_{\mathrm{sym}}^{N_h \times N_h}$ is the matrix representation of the sinc-Galerkin operator $Q_K^{-\beta}(\mathcal{A}_h)$ from (2.17) with respect to the basis $\mathbf{\Phi}_h = \{\phi_{1,h}, \dots, \phi_{N_h,h}\}$, i.e.,

$$\mathbf{Q}_{\beta,K} := c_\beta \delta_y \sum_{|k| \le K} \varepsilon_k^{2\beta} \left(\varepsilon_k^2 \mathbf{A} + \mathbf{M}\right)^{-1}.$$

Remark 2.1. The random fields $\left\{\mathcal{Z}_h^{-K}, \dots, \mathcal{Z}_h^K\right\}$ in (2.20) are independent of the fractional exponent $\beta \in (0,1)$. Thus, having simulated $\mathcal{Z}_h^k$ for every k, we can generate samples of the sinc-Galerkin approximation $\mathcal{Z}_{K,h}^\beta$ in (2.18) simultaneously for different values of $\beta \in (0,1)$ by taking the corresponding linear combinations (2.19).

2.2.4. *Simulation of the noise $\mathcal{W}_h^\Phi$*

Simulating the random load vector $\mathbf{b} \sim \mathsf{N}(\mathbf{0}, \mathbf{M})$ in (2.21) requires the matrix-vector multiplication of (an approximation for) the matrix square root $\sqrt{\mathbf{M}}$ of the Gramian $\mathbf{M} \in \mathbb{R}_{\mathrm{sym}}^{N_h \times N_h}$ in (2.22) with a vector $\mathbf{y}$ of N_h i.i.d. $\mathsf{N}(0,1)$-distributed random variables. In this part we address how this matrix square root can be efficiently approximated. Specifically, we apply the method proposed in Ref. 35 based on contour integral representations to the well-conditioned Gramian $\mathbf{M}$ and are thus able to generate an approximation of the load vector $\mathbf{b} \sim \mathsf{N}(\mathbf{0}, \mathbf{M})$ at a complexity of $\mathcal{O}(N_h \log(N_h))$ and at certified accuracy.

In what follows, let $\widehat{m} > 0$ be a lower bound for smallest eigenvalue and $\widehat{M} > 0$ be an upper bound for the largest eigenvalue of the Gramian $\mathbf{M}$. Furthermore, we set $\widehat{\varkappa}_{\mathrm{M}} := \widehat{M}/\widehat{m}$. Following [35, Eq. (4.4) and comments below], for $\widetilde{K} \in \mathbb{N}$, an approximation $\mathbf{M}_{\widetilde{K}}^{\surd}$ of the matrix square root $\sqrt{\mathbf{M}}$ can be obtained via

$$\mathbf{M}_{\widetilde{K}}^{\surd} := \frac{2J'\sqrt{\widehat{m}}}{\pi \widetilde{K}} \mathbf{M} \sum_{j=1}^{\widetilde{K}} \frac{\mathrm{dn}\left(t_j | 1 - \widehat{\varkappa}_{\mathrm{M}}^{-1}\right)}{\mathrm{cn}^2\left(t_j | 1 - \widehat{\varkappa}_{\mathrm{M}}^{-1}\right)} \left(\mathbf{M} + w_j^2 \mathbf{I}\right)^{-1}. \tag{2.23}$$

Here, $\mathbf{I} \in \mathbb{R}_{\mathrm{sym}}^{N_h \times N_h}$ denotes the identity matrix, sn, cn and dn are the Jacobian elliptic functions [1, Ch. 16], J' is the complete elliptic integral of the second kind associated with the parameter $\widehat{\varkappa}_{\mathrm{M}}^{-1}$ [1, Ch. 17] and, for $j \in \{1, \dots, \widetilde{K}\}$,

$$w_j := \sqrt{\widehat{m}}\, \frac{\mathrm{sn}\left(t_j | 1 - \widehat{\varkappa}_{\mathrm{M}}^{-1}\right)}{\mathrm{cn}\left(t_j | 1 - \widehat{\varkappa}_{\mathrm{M}}^{-1}\right)} \quad \text{and} \quad t_j := \frac{\left(j - \frac{1}{2}\right) J'}{\widetilde{K}}.$$

Note that (2.23) was obtained by reformulating [35, Eq. (4.4)] in order to clearly exhibit symmetry, positive definiteness and condition of the matrices $\mathbf{M} + w_j^2\mathbf{I}$, whose inverses have to be applied to the vector $\mathbf{y}$. We furthermore emphasize that, even though (2.23) results from a quadrature for a contour integral with complex arguments, all quantities in (2.23) are real-valued. In addition, the Gramian $\mathbf{M}$ is symmetric, positive definite and, thus, efficient methods are available to compute approximations $\widehat{m}, \widehat{M}$ of the smallest and largest eigenvalue of $\mathbf{M}$, which are necessary for (2.23). We mention here the power method and the inverse iteration[a], respectively, which performed well in our numerical experiments, see Subsection 6.1.

The following result, taken from [35, Thm. 4.1], shows exponential convergence of the approximation in (2.23) to the matrix square root $\sqrt{\mathbf{M}}$ as $\widetilde{K} \to \infty$ with respect to the 2-norm $\|\cdot\|$ (i.e., the operator matrix norm).

Proposition 2.1. *Let $\mathbf{M}$ be a matrix with eigenvalues in the interval $\left[\widehat{m}, \widehat{M}\right]$ for some $0 < \widehat{m} < \widehat{M}$. Then $\mathbf{M}_{\widetilde{K}}^{\checkmark}$ defined for $\widetilde{K} \in \mathbb{N}$ in (2.23) converges to the matrix square root $\sqrt{\mathbf{M}}$ of $\mathbf{M}$ and there exist constants $\widetilde{c}, C > 0$, which depend only on $\widehat{\varkappa}_{\mathrm{M}} = \widehat{M}/\widehat{m}$, but not on the dimension of $\mathbf{M}$, such that*

$$\left\|\sqrt{\mathbf{M}} - \mathbf{M}_{\widetilde{K}}^{\checkmark}\right\| \leq Ce^{-\widetilde{c}\widetilde{K}}. \tag{2.24}$$

Remark 2.2. If $\left(\mathbf{M}^h\right)_h$ are mass matrices corresponding to finite element spaces $(V_h)_h$ induced by a quasi-uniform family of triangulations and indexed by the mesh width $h \in (0, h_0)$, it is well-known that the condition number $\varkappa_{\mathrm{M}}$ and, thus, the constants $\widetilde{c}, C$ in (2.24) are independent of $h \in (0, h_0)$, see, e.g., [65, Eq. (5.8)]. Since the sum of two well-conditioned, symmetric, positive definite matrices is again well-conditioned, the same is true for the matrices $\mathbf{M} + w_j^2\mathbf{I}$ appearing in (2.23).

2.2.5. *The case $\beta \geq 1$*

We briefly comment on the case of an exponent $\beta \geq 1$. If $\beta = n_\beta + \beta_\star \geq 1$ for some $n_\beta \in \mathbb{N}$ and $\beta_\star \in [0, 1)$, we approximate $\mathcal{Z}_\beta$ in (2.6) by

$$\mathcal{Z}_{K,h}^{\beta} := Q_K^{-\beta_\star}(\mathcal{A}_h)\, \mathcal{A}_h^{-n_\beta}\, \mathcal{W}_h,$$

cf. (2.18), where we set $Q_K^0(\mathcal{A}_h) := \mathcal{I}_h$.

If $\beta_\star = 0$, no sinc quadrature is needed, and we obtain the approximation $\mathcal{Z}_{K,h}^{\beta}$ from an iterated finite element approach, i.e.,

$$\mathbf{Z}_K^{\beta} = \mathbf{A}^{-1}\left(\mathbf{M}\mathbf{A}^{-1}\right)^{n_\beta - 1}\mathbf{M}_{\widetilde{K}}^{\checkmark}\,\mathbf{y}, \qquad \mathbf{y} \sim \mathsf{N}(\mathbf{0}, \mathbf{I}).$$

Here, $\mathbf{M}_{\widetilde{K}}^{\checkmark}$ denotes the approximation (2.23) of the matrix square root $\sqrt{\mathbf{M}}$. The inverse of $\mathbf{A}$ can be numerically approximated for example by means of multilevel methods as described in Sections 3–5.

[a]For a convergence analysis of the power method and the inverse iteration, see [32, Section 8.2].

For $\beta_\star \in (0,1)$, samples of the approximation $\mathcal{Z}_{K,h}^{\beta}$ can be efficiently obtained by solving the $2K+1$ linear systems in (2.21) for the random load vector, approximately given by $\mathbf{b} \approx \left(\mathbf{M}\mathbf{A}^{-1}\right)^{n_\beta} \mathbf{M}_{\widetilde{K}}^{\surd} \mathbf{y}$ for $\mathbf{y} \sim \mathsf{N}(\mathbf{0}, \mathbf{I})$, and by evaluating the linear combination (2.19) with β replaced by $\beta_\star$.

2.3. *Unified error analysis*

The purpose of this section is to provide a complete error analysis for the sinc-Galerkin approximation $\mathcal{Z}_{K,h}^{\beta}$ in (2.18), (2.19), including rigorous bounds for the consistency error if the linear systems (2.21) are approximately solved by an iterative method generating approximations $\widetilde{\mathcal{Z}}_h^k$ of $\mathcal{Z}_h^k$ in (2.20). For linear iterative methods, after any finite number of steps the resulting approximation $\widetilde{\mathcal{Z}}_h^k$ is again a centered GRF.

We first discuss the discretization error of the sinc-Galerkin approximation $\mathcal{Z}_{K,h}^{\beta}$ from (2.18) in a different manner than in Refs. 7, 18: In contrast to Ref. 18 we consider GRFs on general compact metric spaces instead of on bounded Euclidean domains, and compared to [7, Ass. 2.6], we impose less restrictive assumptions on the family of discrete spaces $(V_h)_{h>0}$.

We develop our error analysis under the following assumptions on $(V_h)_{h>0}$.

Assumption 2.5. Let $(V_h)_{h>0}$ be a one parameter family of subspaces $V_h \subset \dot{H}^1_{\mathcal{A}}$ of finite dimension $\dim(V_h) = N_h < \infty$ such that:

(i) There exists a linear projection $\mathcal{P}_h \colon L_2(\mathcal{X}, \nu_{\mathcal{X}}) \to V_h$, which is stable on $L_2(\mathcal{X}, \nu_{\mathcal{X}})$, uniformly in N_h, and whose restriction to $\dot{H}^1_{\mathcal{A}}$ is stable on $\dot{H}^1_{\mathcal{A}}$, uniformly in N_h. Moreover, it allows for convergence rates $\rho_{01}, \rho_{02}, \rho_{12} > 0$ such that

$$\|v - \mathcal{P}_h v\|_0 \leq C_{01} N_h^{-\rho_{01}} \|v\|_1 \qquad \forall v \in \dot{H}^1_{\mathcal{A}}, \tag{2.25}$$

$$\|v - \mathcal{P}_h v\|_0 \leq C_{02} N_h^{-\rho_{02}} \|v\|_2 \qquad \forall v \in \dot{H}^2_{\mathcal{A}}, \tag{2.26}$$

$$\|v - \mathcal{P}_h v\|_1 \leq C_{12} N_h^{-\rho_{12}} \|v\|_2 \qquad \forall v \in \dot{H}^2_{\mathcal{A}}, \tag{2.27}$$

with constants $C_{01}, C_{02}, C_{12} > 0$ which are independent of N_h.

(ii) There exists $\widetilde{\rho} > 0$ such that, for every $0 \leq \sigma \leq 1$, the inverse inequality

$$\|\phi_h\|_\sigma \leq \widetilde{C}_\sigma N_h^{\widetilde{\rho}\sigma} \|\phi_h\|_0 \qquad \forall \phi_h \in V_h, \tag{2.28}$$

holds. Here, the constant $\widetilde{C}_\sigma > 0$ is independent of N_h.

Lemma 2.6. *Suppose that Assumptions 2.1 and 2.5(i) are satisfied for the operator $\mathcal{A}$ and the family $(V_h)_{h>0}$, respectively. For $\beta \in (0,1)$, let $Q_K^{-\beta}(\mathcal{A})$ and $Q_K^{-\beta}(\mathcal{A}_h)$ be the sinc quadrature in* (2.10) *and the sinc-Galerkin approximation in* (2.17)*. Assume further that $\gamma, \gamma' \in (0, \beta)$ and $\sigma, \sigma' \geq 0$ are such that*

$$\sigma < (\beta - \gamma) \min\left\{\tfrac{1-\gamma}{\gamma}, \tfrac{\gamma}{1-\gamma}\right\}, \quad \sigma' < (\beta - \gamma') \min\left\{\tfrac{1-\gamma'}{\gamma'}, \tfrac{\gamma'}{1-\gamma'}\right\}. \tag{2.29}$$

There exists a constant $C > 0$, independent of K, N_h, such that, for all $f \in \dot{H}_{\mathcal{A}}^{-\sigma'}$,

$$\big\|Q_K^{-\beta}(\mathcal{A})f - Q_K^{-\beta}(\mathcal{A}_h)f\big\|_{\sigma} \le C N_h^{-(\gamma+\gamma')\min\{\rho_{01},\,\rho_{12}\}}\|f\|_{-\sigma'}.$$

Proof. Recall from (2.14) that $\|v\|_{a,\varepsilon_k^2}^2 = \varepsilon_k^2\|v\|_1^2 + \|v\|_0^2$ and, thus, $\|v\|_0 \le \|v\|_{a,\varepsilon_k^2}$, $\|v\|_1 \le \varepsilon_k^{-1}\|v\|_{a,\varepsilon_k^2}$, for all $v \in \dot{H}_{\mathcal{A}}^1$. For $0 \le \widetilde{\sigma} \le 1$, we exploit the equivalence of $\dot{H}_{\mathcal{A}}^{\widetilde{\sigma}}$ with the complex interpolation space $\big[\dot{H}_{\mathcal{A}}^0, \dot{H}_{\mathcal{A}}^1\big]_{\widetilde{\sigma}}$, see (2.2)–(2.3), and conclude that there exists a constant $C_{\widetilde{\sigma}} > 0$, independent of K, N_h, such that, for all $v \in \dot{H}_{\mathcal{A}}^1$,

$$\|v\|_{\widetilde{\sigma}} \le C_{\widetilde{\sigma}}\|v\|_{\left[\dot{H}_{\mathcal{A}}^0, \dot{H}_{\mathcal{A}}^1\right]_{\widetilde{\sigma}}} \le C_{\widetilde{\sigma}}\varepsilon_k^{-\widetilde{\sigma}}\|v\|_{a,\varepsilon_k^2}. \tag{2.30}$$

Thus, $u^k \in \dot{H}_{\mathcal{A}}^1$ defined for $k \in \{-K, \dots, K\}$ as in (2.11) satisfies, for $f \in \dot{H}_{\mathcal{A}}^{-\widetilde{\sigma}}$,

$$a_{\varepsilon_k^2}(u^k, u^k) = (f, u^k)_0 \le \|f\|_{-\widetilde{\sigma}}\|u^k\|_{\widetilde{\sigma}} \le C_{\widetilde{\sigma}}\varepsilon_k^{-\widetilde{\sigma}}\|f\|_{-\widetilde{\sigma}}\|u^k\|_{a,\varepsilon_k^2}. \tag{2.31}$$

In the case that $f \in L_2(\mathcal{X}, \nu_{\mathcal{X}})$, we furthermore obtain that $\|u^k\|_0 \le \|u^k\|_{a,\varepsilon_k^2} \le \|f\|_0$ and

$$\begin{aligned}\varepsilon_k^4\|u^k\|_2^2 = \|\varepsilon_k^2\mathcal{A}u^k\|_0^2 = \|f - u^k\|_0^2 &= \|f\|_0^2 - 2(f, u^k)_0 + \|u^k\|_0^2 \\ &= \|f\|_0^2 - 2\|u^k\|_{a,\varepsilon_k^2}^2 + \|u^k\|_0^2 \le \|f\|_0^2.\end{aligned}$$

This shows, for $0 \le \widetilde{\sigma} \le 1$, the ε_k-dependent stability estimates

$$\|u^k\|_0 \le \|f\|_0, \quad \|u^k\|_1 \le \varepsilon_k^{-1}\|u^k\|_{a,\varepsilon_k^2} \le C_{\widetilde{\sigma}}\varepsilon_k^{-1-\widetilde{\sigma}}\|f\|_{-\widetilde{\sigma}}, \quad \|u^k\|_2 \le \varepsilon_k^{-2}\|f\|_0.$$

Using the equivalence in (2.2)–(2.3) once more, we find that for all $0 \le \gamma', \widetilde{\sigma} \le 1$, there exist constants $C_j = C_j(\gamma', \widetilde{\sigma}) > 0$, $j = 1, 2$, independent of K, N_h, such that

$$\|u^k\|_{\gamma'} \le C_1\varepsilon_k^{-\gamma'(1+\widetilde{\sigma})}\|f\|_{-\gamma'\widetilde{\sigma}}, \quad \|u^k\|_{1+\gamma'} \le C_2\varepsilon_k^{-1-\gamma'-(1-\gamma')\widetilde{\sigma}}\|f\|_{-\widetilde{\sigma}+\gamma'\widetilde{\sigma}}. \tag{2.32}$$

Now fix $\gamma, \gamma' \in (0, \beta)$ and $\sigma, \sigma' \ge 0$ satisfying (2.29) and $f \in \dot{H}_{\mathcal{A}}^{-\sigma'}$. First, let $k \in \{0, \dots, K\}$ and u^k, u_h^k be defined as in (2.11), (2.15). Since $u_h^k \in V_h$ is the best approximation of u^k with respect to the energy norm, see (2.16), we find

$$\|u^k - u_h^k\|_{a,\varepsilon_k^2} \le \varepsilon_k\|u^k - \mathcal{P}_h u^k\|_1 + \|u^k - \mathcal{P}_h u^k\|_0,$$

where $\mathcal{P}_h \colon \dot{H}_{\mathcal{A}}^1 \to V_h$ is the operator from Assumption 2.5(i). The approximation properties (2.25), (2.27) of $\mathcal{P}_h$ may be interpolated to hold for elements in $\dot{H}_{\mathcal{A}}^{\gamma'}$ with rate $\gamma'\rho_{01}$ and in $\dot{H}_{\mathcal{A}}^{1+\gamma'}$ with rate $\gamma'\rho_{12}$, respectively. Thus, by (2.32) applied for the choice of parameter $\widetilde{\sigma} := \sigma'/(1 - \max\{1-\gamma', \gamma'\}) < 1$, there exists a constant $C > 0$, independent of K, N_h, such that, for all $k \in \{0, \dots, K\}$, we have

$$\begin{aligned}\|u^k - u_h^k\|_{a,\varepsilon_k^2} &\le C\varepsilon_k^{-\gamma'-\max\{1-\gamma',\gamma'\}\widetilde{\sigma}}\left(N_h^{-\gamma'\rho_{12}} + N_h^{-\gamma'\rho_{01}}\right)\|f\|_{-\widetilde{\sigma}(1-\max\{1-\gamma',\gamma'\})}, \\ &\le C\varepsilon_k^{-\gamma'-\bar{\sigma}'}N_h^{-\gamma'\min\{\rho_{01},\,\rho_{12}\}}\|f\|_{-\sigma'},\end{aligned} \tag{2.33}$$

where $\bar{\sigma}' := \sigma'\max\{1-\gamma', \gamma'\}/(1 - \max\{1-\gamma', \gamma'\})$.

For $k \in \{-K, \dots, -1\}$, we consider $v^k := \varepsilon_k^2 u^k$ and $v_h^k := \varepsilon_k^2 u_h^k$, where u^k, u_h^k are again as in (2.11), (2.15). Then v^k solves $(\mathcal{A} + \varepsilon_k^{-2}\mathcal{I})v_k = f$ (in weak sense) and

v_h^k is the respective Galerkin approximation. Since $\varepsilon_k^{-1} < 1$ for $k < 0$, the operator $(\mathcal{A}+\varepsilon_k^{-2}\mathcal{I})$ is not singularly perturbed. Thus, by a standard argument there exists a constant $C > 0$ that does not depend on K, N_h such that, for all $k \in \{-K, \dots, -1\}$

$$\|v^k - v_h^k\|_1 \leq C N_h^{-\rho_{12}(1-\sigma')} \|f\|_{-\sigma'} \leq C N_h^{-\gamma'\rho_{12}} \|f\|_{-\sigma'}, \tag{2.34}$$

where $1 - \sigma' > \gamma'$ is implied by the assumption on σ' in (2.29).

Let $G \in \dot{H}_{\mathcal{A}}^{-\sigma}$ be arbitrary such that $\|G\|_{-\sigma} = 1$ and denote by $u_G^k, u_{G,h}^k, v_G^k, v_{G,h}^k$ the respective solutions to the adjoint equation (2.11) with right-hand side G. The differences $u_G^k - u_{G,h}^k$, $v_G^k - v_{G,h}^k$ satisfy (2.33) and (2.34), respectively, with $\|f\|_{-\sigma'}$ replaced by $\|G\|_{-\sigma}$. By an Aubin–Nitsche duality argument, we then obtain

$$\begin{aligned}
\big|G(Q_K^{-\beta}(\mathcal{A})f - Q_K^{-\beta}(\mathcal{A}_h)f)\big| &\leq C\delta_y \Bigg(\sum_{k=0}^{K} \varepsilon_k^{2\beta} \|u^k - u_h^k\|_{a,\varepsilon_k^2} \|u_G^k - u_{G,h}^k\|_{a,\varepsilon_k^2} \\
&\qquad + \sum_{k=-K}^{-1} \varepsilon_k^{2\beta-2} \|v^k - v_h^k\|_1 \|v_G^k - v_{G,h}^k\|_1 \Bigg) \\
&\leq C N_h^{-(\gamma+\gamma')\min\{\rho_{01},\rho_{12}\}} \|f\|_{-\sigma'} \left(\delta_y + \delta_y \sum_{k=1}^{K} \varepsilon_k^{2\beta-\gamma-\bar\sigma-\gamma'-\bar\sigma'} + \delta_y \sum_{k=-K}^{-1} \varepsilon_k^{2\beta-2} \right).
\end{aligned}$$

By a geometric series argument and the fact that $1/(e^x - 1) \leq 1/x$ for every $x > 0$,

$$\delta_y \sum_{k=1}^{K} \varepsilon_k^{2\beta-\gamma-\bar\sigma-\gamma'-\bar\sigma'} \leq \tfrac{1}{2\beta-\gamma-\bar\sigma-\gamma'-\bar\sigma'} \quad \text{and} \quad \delta_y \sum_{k=-K}^{-1} \varepsilon_k^{2\beta-2} \leq \tfrac{1}{2(1-\beta)}.$$

The claim follows upon taking the supremum over $G \in \dot{H}_{\mathcal{A}}^{-\sigma}$ with $\|G\|_{-\sigma} = 1$. □

Exploiting Lemmata 2.3, 2.4 and 2.6 allows us to bound, for $\beta \in (0,1)$, the strong $L_p\big(\Omega; \dot{H}_{\mathcal{A}}^{\sigma}\big)$-error between the generalized Whittle–Matérn field $\mathcal{Z}^\beta$ in (2.4), indexed by a compact metric space $\mathcal{X}$, and the sinc-Galerkin approximation $\mathcal{Z}_{K,h}^\beta$ in (2.18). This generality comes at the expense that for the Euclidean case $\mathcal{X} = \overline{\mathcal{D}}$ in connection with *uniformly refined* finite element meshes, analyzed in detail in [18, Thm. 6.23], the error bound of the following proposition will be as strong only if $\rho_{01} = \rho_{12} = 1$ and $\rho_{02} = 2$, i.e., if the elliptic problem associated with the differential operator $\mathcal{A}$ is $H^2(\mathcal{D})$-regular in the sense of [18, Def. 6.20].

Proposition 2.2. *Let the operator $\mathcal{A}$ satisfy Assumption 2.1 and the spectral asymptotics* (2.5) *for some $\alpha > 0$. Suppose Assumptions 2.5(i)–(ii) for the discrete spaces $(V_h)_{h>0}$. For $\beta \in (0,1)$, let $\mathcal{Z}^\beta$ be a GRF with distribution* (2.4) *and $\mathcal{Z}_{K,h}^\beta$ be the sinc-Galerkin approximation in* (2.18)*. Assume further that $\gamma \in (0, \beta)$ and $\sigma \geq 0$ are such that $2\beta - \sigma > \alpha^{-1}$ and* (2.29) *hold. Then, for every $\delta > 0$ and all $p \in (0, \infty)$, there exists a constant $C > 0$, independent of K and $N_h := \dim(V_h)$,*

such that

$$\Big(\mathbb{E}\Big[\big\|\mathcal{Z}^\beta - \mathcal{Z}^\beta_{K,h}\big\|_\sigma^p\Big]\Big)^{1/p} \le CN_h^{-\rho_{02}/2\left(2\beta-\sigma-\alpha^{-1}-\delta\right)} + C\left(e^{-2\min\{\beta,1-\beta\}\sqrt{K}}N_h^{\widetilde{\rho}\sigma} + N_h^{-2\gamma\min\{\rho_{01},\rho_{12}\}}\right)N_h^{1/2}.$$

Here, $\rho_{01},\rho_{02},\rho_{12},\widetilde{\rho}>0$ are as in Assumption 2.5.

Proof. We consider the case $p=2$ first. Then we bound the error as follows,

$$\mathbb{E}\Big[\big\|\mathcal{Z}^\beta - \mathcal{Z}^\beta_{K,h}\big\|_\sigma^2\Big] \le 3\,\mathbb{E}\left[\big\|\mathcal{A}^{-\beta}(\mathcal{W}-\mathcal{W}_h)\big\|_\sigma^2\right] + 3\,\mathbb{E}\left[\big\|\big(\mathcal{A}^{-\beta}-Q_K^{-\beta}(\mathcal{A})\big)\mathcal{W}_h\big\|_\sigma^2\right] + 3\,\mathbb{E}\left[\big\|\big(Q_K^{-\beta}(\mathcal{A})-Q_K^{-\beta}(\mathcal{A}_h)\big)\mathcal{W}_h\big\|_\sigma^2\right] =: \mathrm{(I)}+\mathrm{(II)}+\mathrm{(III)}.$$

The terms (II) and (III) can be bounded by Lemma 2.3 and by Lemma 2.6 (applied for $\gamma=\gamma'$ and $\sigma'=0$), respectively, combined with Lemma 2.4 and Assumption 2.5(ii). Thus, there exists a constant $C>0$, independent of K,N_h such that

$$\mathrm{(II)} \le C^2 e^{-4\min\{\beta,1-\beta\}\sqrt{K}}N_h^{1+2\widetilde{\rho}\sigma}, \qquad \mathrm{(III)} \le C^2 N_h^{-4\gamma\min\{\rho_{01},\rho_{12}\}+1}.$$

It remains to estimate term (I). To this end, we recall the definition of $\dot{H}^\sigma_{\mathcal{A}}$ from (2.1) and find by Fubini's theorem

$$\mathrm{(I)} = \sum_{j\in\mathbb{N}}\lambda_j^\sigma\,\mathbb{E}\left[\big(\mathcal{A}^{-\beta}(\mathcal{I}-\Pi_h)\mathcal{W},e_j\big)_0^2\right] = \sum_{j\in\mathbb{N}}\lambda_j^{\sigma-2\beta}\|(\mathcal{I}-\Pi_h)e_j\|_0^2.$$

We use the fact that $\Pi_h e_j\in V_h$ is the $L_2(\mathcal{X},\nu_{\mathcal{X}})$-best approximation of e_j and we find with (2.26) of Assumption 2.5(i) that

$$\|(\mathcal{I}-\Pi_h)e_j\|_0 \le \|(\mathcal{I}-\mathcal{P}_h)e_j\|_0 \le C_{02}N_h^{-\rho_{02}}\|e_j\|_2.$$

Since also $\|(\mathcal{I}-\Pi_h)e_j\|_0 \le \|e_j\|_0 = 1$, we obtain, for $0\le\sigma'\le 2$, by the equivalence of $\dot{H}^{\sigma'}_{\mathcal{A}}$ with $\big[\dot{H}^0_{\mathcal{A}},\dot{H}^2_{\mathcal{A}}\big]_{\sigma'/2}$, cf. (2.2)–(2.3), that there exists a constant $C_{\sigma'}>0$, independent of K,N_h such that

$$\|(\mathcal{I}-\Pi_h)e_j\|_0 \le \|(\mathcal{I}-\mathcal{P}_h)e_j\|_0 \le C_{02}^{\sigma'/2}N_h^{-\rho_{02}\sigma'/2}\|e_j\|_{\left[\dot{H}^0_{\mathcal{A}},\dot{H}^2_{\mathcal{A}}\right]_{\sigma'/2}} \le C_{\sigma'}N_h^{-\rho_{02}\sigma'/2}\|e_j\|_{\sigma'} = C_{\sigma'}N_h^{-\rho_{02}\sigma'/2}\lambda_j^{\sigma'/2}.$$

Without loss of generality we may assume that $\delta\in(0,2\beta-\sigma-\alpha^{-1})$. We then conclude with the choice $\sigma' := 2\beta-\sigma-\alpha^{-1}-\delta\in(0,2)$ and the spectral asymptotics (2.5) that there exist constants $C,\widetilde{C}>0$, independent of K,N_h, such that

$$\mathrm{(I)} \le \widetilde{C}N_h^{-\rho_{02}\left(2\beta-\sigma-\alpha^{-1}-\delta\right)}\sum_{j\in\mathbb{N}}j^{-1-\alpha\delta} \le CN_h^{-\rho_{02}\left(2\beta-\sigma-\alpha^{-1}-\delta\right)}.$$

Applying the Kahane–Khintchine inequalities, see, e.g., [45, Thm. 4.7 and p. 103], for the zero-mean $\dot{H}^\sigma_{\mathcal{A}}$-valued Gaussian random variable $\mathcal{Z}^\beta-\mathcal{Z}^\beta_{K,h}$ shows the assertion for all $p\in(0,\infty)$. □

The final component in our unified error analysis is the consistency error. For this, we suppose that a bound for the consistency error of each Galerkin approximation u_h^k in (2.15) with respect to the respective energy norm $\|\cdot\|_{a,\varepsilon_k^2}$ holds, which is uniform in k. Below, we formulate this assumption for general $\varepsilon > 0$.

Assumption 2.7. Suppose that $\mathcal{A}\colon \mathscr{D}(\mathcal{A}) \subseteq L_2(\mathcal{X},\nu_\mathcal{X}) \to L_2(\mathcal{X},\nu_\mathcal{X})$ is a linear operator satisfying Assumption 2.1 and that $V_h \subset \dot{H}^1_\mathcal{A}$ has finite dimension $\dim(V_h) = N_h$. For $\varepsilon > 0$, let $a_{\varepsilon^2}(\cdot,\cdot)$ and $\|\cdot\|_{a,\varepsilon^2}$ denote the bilinear form in (2.13) and the corresponding energy norm in (2.14) with $\eta = \varepsilon^2$. Assume that $f \in L_2(\mathcal{X},\nu_\mathcal{X})$ and that $\widetilde{u}_h^\varepsilon \in V_h$ is an approximation to the Galerkin solution

$$u_h^\varepsilon \in V_h: \quad a_{\varepsilon^2}(u_h^\varepsilon, v_h) = (f, v_h)_0 \quad \forall v_h \in V_h,$$

which is obtained by approximate solution of the linear system of equations by an iterative solution algorithm, whose number of iterations is uniquely determined by $n_0 \in \mathbb{N}$ and N_h. Suppose, there exist constants $C_{\mathrm{con}} > 0$, $q \in (0,1)$, independent of ε, n_0, f and N_h, such that $\widetilde{u}_h^\varepsilon$ satisfies the consistency error bound

$$\|\widetilde{u}_h^\varepsilon - u_h^\varepsilon\|_{a,\varepsilon^2} \le C_{\mathrm{con}}\, q^{n_0} \|u_h^\varepsilon\|_{a,\varepsilon^2}. \tag{2.35}$$

Proposition 2.3. *Suppose that we are given an iterative solution algorithm satisfying Assumption 2.7 which generates approximations $\widetilde{u}_h^k$ of the Galerkin solutions u_h^k in (2.15). Let $c_\beta := \pi^{-1}(2\sin(\pi\beta))$ and δ_y, ε_k be as in (2.8). Then, $\widetilde{u}_{K,h}$ defined as the linear combination*

$$\widetilde{u}_{K,h} := c_\beta \delta_y \sum_{|k|\le K} \varepsilon_k^{2\beta} \widetilde{u}_h^k,$$

approximates the sinc-Galerkin solution $u_{K,h}$ in (2.17), and there exists a constant $C > 0$, independent of q, n_0, K and N_h, such that, for $\beta \in (0,1)$, $0 \le \sigma \le 1$ with $\sigma < 2\beta$, $0 \le \sigma' < \min\{1, 2\beta - \sigma\}$, and all $f \in \dot{H}^{-\sigma'}_\mathcal{A}$ the consistency error bound

$$\|\widetilde{u}_{K,h} - u_{K,h}\|_\sigma \le C q^{n_0} \|f\|_{-\sigma'}, \tag{2.36}$$

with respect to the norm on $\dot{H}^\sigma_\mathcal{A}$ from (2.1) and on the dual space $\dot{H}^{-\sigma'}_\mathcal{A}$ holds.

Proof. Let $0 \le \widetilde{\sigma} \le 1$ and $C_{1\hookrightarrow\widetilde{\sigma}} > 0$ be a constant such that $\|v\|_{\widetilde{\sigma}} \le C_{1\hookrightarrow\widetilde{\sigma}} \|v\|_1$ for all $v \in \dot{H}^1_\mathcal{A} \hookrightarrow \dot{H}^{\widetilde{\sigma}}_\mathcal{A}$. Then, by definition of the bilinear form $a_{\varepsilon_k^2}$, cf. (2.13), we obtain, for $v \in \dot{H}^1_\mathcal{A}$, the following relations for the corresponding energy norm (2.14),

$$\|v\|_0 \le \|v\|_{a,\varepsilon_k^2}, \quad \|v\|_1 \le \varepsilon_k^{-1}\|v\|_{a,\varepsilon_k^2}, \quad \|v\|_{\widetilde{\sigma}} \le \begin{cases} C_{1\hookrightarrow\widetilde{\sigma}}\, \varepsilon_k^{-1}\|v\|_{a,\varepsilon_k^2}, & k < 0, \\ C_{\widetilde{\sigma}}\, \varepsilon_k^{-\widetilde{\sigma}}\|v\|_{a,\varepsilon_k^2}, & k \ge 0, \end{cases}$$

where $C_{\widetilde{\sigma}} > 0$ is a constant, independent of K, N_h, see (2.30). The ε_k-dependent stability estimate (2.31) holds also for the Galerkin approximation u_h^k in (2.15), i.e.,

$$\|u_h^k\|_{a,\varepsilon_k^2} \le C_{\widetilde{\sigma}}\, \varepsilon_k^{-\widetilde{\sigma}} \|f\|_{-\widetilde{\sigma}}, \qquad 0 \le \widetilde{\sigma} \le 1. \tag{2.37}$$

With these preliminary observations on the energy norm $\|\cdot\|_{a,\varepsilon_k^2}$ at hand, we proceed similarly as in the proof of Lemma 2.6 and find

$$\begin{aligned}\|u_{K,h}-\widetilde{u}_{K,h}\|_\sigma &\le c_\beta\delta_y\sum_{|k|\le K}\varepsilon_k^{2\beta}\|u_h^k-\widetilde{u}_h^k\|_\sigma\\ &\le C'\delta_y\left(\sum_{k=-K}^{-1}\varepsilon_k^{2\beta-1}\|u_h^k-\widetilde{u}_h^k\|_{a,\varepsilon_k^2}+\sum_{k=0}^{K}\varepsilon_k^{2\beta-\sigma}\|u_h^k-\widetilde{u}_h^k\|_{a,\varepsilon_k^2}\right),\end{aligned}$$

where $C' := c_\beta\max\{C_{1\hookrightarrow\sigma}, C_\sigma\}$. We use the consistency error bound (2.35) of Assumption 2.7 for each term in the sums, $\|u_h^k-\widetilde{u}_h^k\|_{a,\varepsilon_k^2}\le C_{\mathrm{con}}q^{n_0}\|u_h^k\|_{a,\varepsilon_k^2}$, and set $C'' := C_{\mathrm{con}}C' > 0$. By recalling the definition of ε_k from (2.8), we then obtain with (2.37) (applied for $\widetilde{\sigma}=1$ and $\widetilde{\sigma}=\sigma'$, respectively) that, for $0\le\sigma<2\beta$, $0\le\sigma'<\min\{1,2\beta-\sigma\}$ and all $f\in\dot{H}_{\mathcal{A}}^{-\sigma'}$,

$$\begin{aligned}\|u_{K,h}-\widetilde{u}_{K,h}\|_\sigma &\le C''q^{n_0}\delta_y\left(C_1\sum_{k=-K}^{-1}\varepsilon_k^{-2(1-\beta)}\|f\|_{-1}+C_{\sigma'}\sum_{k=0}^{K}\varepsilon_k^{2\beta-\sigma-\sigma'}\|f\|_{-\sigma'}\right)\\ &\le \widetilde{C}q^{n_0}\|f\|_{-\sigma'}\left(\delta_y+\delta_y\sum_{k=1}^{K}e^{-2(1-\beta)k\delta_y}+\delta_y\sum_{k=1}^{K}e^{-(2\beta-\sigma-\sigma')k\delta_y}\right),\end{aligned}$$

where $\widetilde{C} := C''\max\{C_1, C_{\sigma'}\} > 0$. Finally, bounding the sums by geometric series and using the relation $\left(e^{a\delta_y}-1\right)^{-1}\le(a\delta_y)^{-1}$, which holds for any $a>0$, gives

$$\|u_{K,h}-\widetilde{u}_{K,h}\|_\sigma \le \widetilde{C}\left(\tfrac{1}{\sqrt{K}}+\tfrac{1}{2(1-\beta)}+\tfrac{1}{2\beta-\sigma-\sigma'}\right)q^{n_0}\|f\|_{-\sigma'},$$

and (2.36) follows for $C := \widetilde{C}\left(1+(2-2\beta)^{-1}+(2\beta-\sigma-\sigma')^{-1}\right)$. □

Theorem 2.1. *Let $\beta\in(0,1)$ and suppose that the operator $\mathcal{A}$ satisfies Assumption 2.1 and the spectral asymptotics (2.5) for $\alpha>0$. Assume that we are given an iterative algorithm satisfying Assumption 2.7 which generates approximations $\widetilde{\mathcal{Z}}_h^k$ of the GRFs $\mathcal{Z}_h^k$ in (2.20). Then, $\widetilde{\mathcal{Z}}_{K,h}^\beta$, defined as the linear combination*

$$\widetilde{\mathcal{Z}}_{K,h}^\beta := c_\beta\delta_y\sum_{|k|\le K}\varepsilon_k^{2\beta}\widetilde{\mathcal{Z}}_h^k,\qquad \mathbb{P}\text{-a.s.},\qquad c_\beta := \pi^{-1}(2\sin(\pi\beta)),\tag{2.38}$$

with δ_y,ε_k from (2.8), approximates the sinc-Galerkin GRF $\mathcal{Z}_{K,h}^\beta$ in (2.18)–(2.19). Furthermore, for all $p\in(0,\infty)$, there exists a constant $C>0$, which is independent of q, n_0, K, and N_h, such that

$$\left(\mathbb{E}\left[\left\|\widetilde{\mathcal{Z}}_{K,h}^\beta-\mathcal{Z}_{K,h}^\beta\right\|_\sigma^p\right]\right)^{1/p}\le Cq^{n_0}N_h^{1/2}\tag{2.39}$$

holds, with respect to the norm on $\dot{H}_{\mathcal{A}}^\sigma$ in (2.1) for any $0\le\sigma\le1$ with $\sigma<2\beta$.

Proof. The claim follows from Proposition 2.3 (with $\sigma'=0$) and Lemma 2.4. □

3. GRF simulation on bounded Euclidean domains

3.1. *Admissible fields*

In what follows, we let $\mathcal{D}$ be a bounded, connected, open polytope in $\mathbb{R}^n$ for $n \in \{1,2,3\}$, with Lipschitz continuous boundary $\partial\mathcal{D}$, and we consider GRFs as described in Subsection 2.1 on $\mathcal{X} = \overline{\mathcal{D}}$, equipped with the Euclidean distance and the Lebesgue measure.

We furthermore assume that the operator $\mathcal{A}\colon \mathscr{D}(\mathcal{A}) \subseteq L_2(\mathcal{D}) \to L_2(\mathcal{D})$ is a linear, self-adjoint, elliptic differential operator of second order. Specifically, given functions $A\colon \mathcal{D} \to \mathbb{R}^{n\times n}$ and $\kappa\colon \mathcal{D} \to \mathbb{R}$, we let $\mathcal{A}\colon \mathscr{D}(\mathcal{A}) \subseteq L_2(\mathcal{D}) \to L_2(\mathcal{D})$ be the maximal accretive operator on $L_2(\mathcal{D})$ associated with the coefficients A, κ^2, and with domain $\mathscr{D}(\mathcal{A}) \subseteq H^1_0(\mathcal{D})$. By this we mean that, for $u \in \mathscr{D}(\mathcal{A})$, the element $\mathcal{A}u \in L_2(\mathcal{D})$ is uniquely defined via the relation

$$(A\nabla u, \nabla v)_{L_2(\mathcal{D})^d} + (\kappa^2 u, v)_{L_2(\mathcal{D})} = (\mathcal{A}u, v)_{L_2(\mathcal{D})} \quad \forall v \in H^1_0(\mathcal{D}). \tag{3.1}$$

We work under the following *minimal assumptions* on the coefficients A, κ (we shall add further, more restrictive assumptions as required in the ensuing analysis).

Assumption 3.1. $\kappa \in L_\infty(\mathcal{D})$ and $A \in L_\infty\big(\mathcal{D}; \mathbb{R}^{n\times n}\big)$ is symmetric and uniformly positive definite, i.e.,

$$\exists a_- > 0 : \quad \forall \xi \in \mathbb{R}^n : \quad \operatorname*{ess\,inf}_{x\in\mathcal{D}} \xi^\top A(x)\xi \geq a_- |\xi|^2. \tag{3.2}$$

The corresponding bounds for the essential suprema on $\mathcal{D}$ are $\kappa_+ \geq 0$, $a_+ > 0$, i.e.,

$$\operatorname*{ess\,sup}_{x\in\mathcal{D}} |\kappa(x)| = \kappa_+, \qquad \forall \xi \in \mathbb{R}^n : \ \operatorname*{ess\,sup}_{x\in\mathcal{D}} \xi^\top A(x)\xi \leq a_+ |\xi|^2.$$

Under Assumption 3.1 the operator $\mathcal{A}\colon \mathscr{D}(\mathcal{A}) \to L_2(\mathcal{D})$ in (3.1) is densely defined, self-adjoint and by the Lax–Milgram theorem, it admits a bounded inverse $\mathcal{A}^{-1}\colon H^1_0(\mathcal{D})^* \to H^1_0(\mathcal{D})$. By the Rellich–Kondrachov theorem, $\mathcal{A}^{-1}$ is compact on $L_2(\mathcal{D})$, see, e.g., [31, Thm. 7.22]. Therefore, $\mathcal{A}$ satisfies Assumption 2.1 for $\mathcal{X} = \overline{\mathcal{D}}$ and we are in the setting of Subsection 2.1. For $\beta > 0$, we may then consider the GRF $\mathcal{Z}^\beta$ in (2.4)/(2.6) with covariance operator $\mathcal{C}_\beta = \mathcal{A}^{-2\beta}$.

By Weyl's law, the second-order differential operator $\mathcal{A}$ defined via (3.1) admits the spectral asymptotics (2.5) for $\alpha = 2/n$, see [23, Thm. 6.3.1]. Thus, Lemma 2.2 shows that

$$\beta > n/4, \tag{3.3}$$

is sufficient for $\mathcal{Z}^\beta \in L_2(\mathcal{D})$ to hold $\mathbb{P}$-a.s. and in L_p-sense for $p \in (0,\infty)$.

Remark 3.1. For the Whittle–Matérn field (1.1), the assumption (3.3) corresponds to a positive smoothness parameter $\nu > 0$, i.e., a non-degenerate random field.

Below, we briefly recall the relation between the space $\dot{H}^\sigma_{\mathcal{A}}$ in (2.1), defined for the differential operator $\mathcal{A}$ in (3.1), and the Sobolev space $H^\sigma(\mathcal{D})$, cf. [47, Ch. 11].

Lemma 3.2. *Let Assumption 3.1 be satisfied and* $\mathcal{A}\colon \mathscr{D}(\mathcal{A}) \to L_2(\mathcal{D})$ *be the differential operator in* (3.1)*. In this case, the space* $\dot{H}^\sigma_{\mathcal{A}}$ *defined via* (2.1) *satisfies*

$$\big(\dot{H}^\sigma_{\mathcal{A}}, \|\cdot\|_\sigma\big) \hookrightarrow \big(H^\sigma(\mathcal{D}), \|\cdot\|_{H^\sigma(\mathcal{D})}\big), \qquad 0 \le \sigma \le 1,$$

and the norms $\|\cdot\|_\sigma$, $\|\cdot\|_{H^\sigma(\mathcal{D})}$ *are equivalent on* $\dot{H}^\sigma_{\mathcal{A}}$ *for* $0 \le \sigma \le 1$ *and* $\sigma \neq 1/2$.

Remark 3.2. Lemma 3.2 shows that all results of the presently developed, unified error analysis in Subsection 2.3 can be formulated with respect to the norms on the Sobolev space $H^\sigma(\mathcal{D})$ and the dual $H^{-\sigma'}(\mathcal{D}) = H^{\sigma'}(\mathcal{D})^*$ instead of $\|\cdot\|_\sigma$ and $\|\cdot\|_{-\sigma'}$, respectively.

3.2. *Finite element discretization*

The discretization of the $2K+1$ reaction-diffusion equations (2.11) for $\mathcal{X} = \overline{\mathcal{D}} \subset \mathbb{R}^n$ proceeds by a standard Galerkin finite element method (FEM). We confine our error and preconditioning analysis to first order, Lagrangean nodal FEM on regular, simplicial partitions $\mathcal{T}_h$ of $\overline{\mathcal{D}}$ of mesh width $h := \max\{\mathrm{diam}(T) \,|\, T \in \mathcal{T}_h\}$. We add that higher order, Lagrangean FEM can be analyzed in exactly the same fashion. The corresponding subspace V_h of $\dot{H}^1_{\mathcal{A}} = H^1_0(\mathcal{D})$ in (2.12) is of finite dimension $N_h = \dim(V_h) = \mathcal{O}\,(h^{-n})$.

In the context of multilevel methods, we shall also stipulate the availability of a *nested sequence* $\{\mathcal{T}_\ell\}_{\ell\ge 0}$ of such triangulations, with mesh width h_ℓ and corresponding finite element (FE) space $V_\ell := V_{h_\ell}$ of dimension $N_\ell := \dim(V_\ell)$, with nodal FE basis $\mathbf{\Phi}_\ell$. We also write $\mathcal{A}_\ell$ and $\mathcal{I}_\ell$ for the corresponding Galerkin-FE operator $\mathcal{A}_{h_\ell}\colon V_\ell \to V_\ell$ and the identity on V_ℓ, respectively.

We note that in the Euclidean case with $\mathcal{A}$ defined via (3.1) the *local* linear, second-order divergence form differential operators in (2.11) are formally given by

$$\big(\varepsilon^2\mathcal{A} + \mathcal{I}\big)\, w(x) = -\varepsilon^2 \nabla \cdot A(x)\nabla w(x) + \big(1 + \varepsilon^2\kappa^2(x)\big)\, w(x) \tag{3.4}$$

where the parameter $\varepsilon > 0$ stemming from the sinc quadrature ranges from $\exp\big(-\sqrt{K}\big)$ to $\exp\big(\sqrt{K}\big)$, cf. (2.8). Consequently, if the coefficients A, κ satisfy Assumption 3.1, the differential operator in (3.4) is a) *regularly perturbed* for $1 \le \varepsilon \le \exp\big(\sqrt{K}\big)$, and b) *singularly perturbed* for $\exp\big(-\sqrt{K}\big) \le \varepsilon < 1$.

3.3. *Multilevel preconditioning*

The key to the efficient simulation of the sinc-Galerkin approximation $\mathcal{Z}^\beta_{K,h}$ in (2.18), (2.19) of the GRF $\mathcal{Z}^\beta$ in (2.6) with distribution (2.4) is the numerical solution of the linear systems (2.21). The corresponding solution vectors enter the sinc approximation (2.19) additively; accordingly, the $2K+1$ linear systems (2.21) can be solved in parallel. The main issue in proving linear complexity for

iterative solvers is *preconditioning* of the coefficient matrices in (2.21). Due to the above-mentioned range of the parameters ε_k standard preconditioning arguments for second-order, elliptic PDEs do not *robustly precondition* (2.21), i.e., the condition number of the preconditioned coefficient matrices cannot be bounded uniformly with respect to h and to ε_k.

In the following, we present and analyze multilevel preconditioning methods which are robust with respect to ε_k. Specifically, we consider the BPX multilevel preconditioner which is applicable to any available FEM implementation for the operator $\mathcal{A}$ that affords a multilevel structure. We remark that in this section we first develop the preconditioning techniques for the Euclidean case, but their scope is considerably wider: they extend to n-variate manifolds $\mathcal{X} = \mathcal{M}$, thereby facilitating efficient numerical simulation of the GRF in (2.4). For 2-surfaces in $\mathbb{R}^3$, this is discussed in Section 4.

For symmetric problems like (2.15) multilevel preconditioners afford fast solution methods, see, e.g., Ref. 65. In our setting, an estimate for the condition number of the preconditioned coefficient matrix is required, which is uniform with respect to the parameter ε in (3.4). We thus consider the BPX preconditioner for the parametric family of bounded linear operators $\{\mathcal{A} + \varepsilon^{-2}\mathcal{I}\colon H^1_0(\mathcal{D}) \to H^{-1}(\mathcal{D}),\ \varepsilon > 0\}$, i.e., the elliptic operator $\mathcal{A}$ "shifted" by the factor ε^{-2}.

The finite-dimensional operators $\mathcal{A}_\ell\colon V_\ell \to V_\ell$, $\ell \geq 0$, introduced in Subsection 3.2 are symmetric, positive definite (SPD), and there exists a constant $C > 0$ such that the condition number of $\mathcal{A}_\ell$ satisfies $\varkappa(\mathcal{A}_\ell) \leq C h_\ell^{-2}$ (e.g., [65, Eq. (5.8)]). This limits the applicability of iterative solvers such as conjugate gradient (CG), which converge linearly, with contraction number $\delta_\ell = \big(\sqrt{\varkappa(\mathcal{A}_\ell)}-1\big)/\big(\sqrt{\varkappa(\mathcal{A}_\ell)}+1\big)$. Thus, it may happen that $\delta_\ell \to 1$ as $\ell \to \infty$. Appropriate preconditioning of $\mathcal{A}_\ell$ by $B_\ell : V_\ell \to V_\ell$ can provide linear convergence of preconditioned CG (PCG) with contraction numbers that only depend *mildly* on ℓ. For every discretization level $L \in \mathbb{N}$, the BPX preconditioner, see Ref. 65, on V_L is defined by

$$B_L := \sum_{\ell=0}^{L} R_\ell \Pi_\ell, \tag{3.5}$$

where, for $\ell \in \{0, \dots, L\}$, $R_\ell\colon V_\ell \to V_\ell$ is an SPD operator that is referred to as *smoother* and $\Pi_\ell\colon V_L \to V_\ell$ is the $L_2(\mathcal{D})$-orthogonal projection. For simplicity the dependence on L is not explicitly reflected in the notation of these operators.

We recall that $\dot{H}^0_{\mathcal{A}} = L_2(\mathcal{D})$ and $(\,\cdot\,,\,\cdot\,)_0 = (\,\cdot\,,\,\cdot\,)_{L_2(\mathcal{D})}$, see Lemma 3.2. In what follows, we assume that there exists a constant $C > 0$ such that for every $\ell \geq 0$, for all $v_\ell \in V_\ell$, and for every $\varepsilon \in (0, \infty)$,

$$C^{-1}(R_\ell v_\ell, v_\ell)_0 \leq h_\ell^2 \left(\|\nabla v_\ell\|_0^2 + \varepsilon^{-2} h_\ell^2 \|v_\ell\|_0^2\right) \leq C(R_\ell v_\ell, v_\ell)_0. \tag{3.6}$$

Proposition 3.1. *Suppose that the coefficients A, κ of the differential operator $\mathcal{A}$ in* (3.1) *satisfy Assumption 3.1 with bounds $a_+ \geq a_- > 0$, $\kappa_+ \geq 0$ and that the smoothers $\{R_\ell\}_{\ell \geq 0}$ in* (3.5) *satisfy* (3.6)*. Then, there exists a constant $C > 0$,*

depending on $a_\pm, \kappa_+, \mathcal{D}$ *and independent of* $\varepsilon \in (0,\infty)$*, such that*

$$\varkappa\left(B_L\left(\mathcal{A}_L + \varepsilon^{-2}\mathcal{I}_L\right)\right) \le C\,|\log(h_L)| \quad \forall L \in \mathbb{N}.$$

Proof. For all $\ell \ge 0$, the FE-discretized Dirichlet Laplacian $-\Delta_\ell \colon V_\ell \to V_\ell$ satisfies

$$(-\Delta_\ell v_\ell, w_\ell)_0 := (\nabla v_\ell, \nabla w_\ell)_0 \quad \forall v_\ell, w_\ell \in V_\ell.$$

By Assumption 3.1 and the Poincaré inequality we have, for $\ell \ge 0$ and $v_\ell \in V_\ell$,

$$a_-(-\Delta_\ell v_\ell, v_\ell)_0 \le (\mathcal{A}_\ell v_\ell, v_\ell)_0 \le \left(a_+ + C_{\mathcal{D}}\kappa_+^2\right)(-\Delta_\ell v_\ell, v_\ell)_0, \tag{3.7}$$

where $C_{\mathcal{D}} > 0$ is the constant in the Poincaré inequality, which depends only on $\mathcal{D}$; namely $C_{\mathcal{D}}^{-1}$ is the smallest positive eigenvalue of the Dirichlet Laplacian on $\mathcal{D}$.

Thus, we find with (3.7) and [42, Lem. 6] that, for every $v_L \in V_L$,

$$\begin{aligned}\left(\left(\mathcal{A}_L + \varepsilon^{-2}\mathcal{I}_L\right)v_L, v_L\right)_0 &\le \left(a_+ + C_{\mathcal{D}}\kappa_+^2\right)(-\Delta_L v_L, v_L)_0 + (\varepsilon^{-2}v_L, v_L)_0 \\ &\le C_{A,\kappa,\mathcal{D}}|\log(h_L)|\left(B_L^{-1}v_L, v_L\right)_0,\end{aligned}$$

and the constant $C_{A,\kappa,\mathcal{D}} := \max\left\{1,\, a_+ + C_{\mathcal{D}}\kappa_+^2\right\} > 0$ does not depend on $\varepsilon > 0$. We conclude that $\lambda_{\max}\left(B_L\left(\mathcal{A}_L + \varepsilon^{-2}\mathcal{I}_L\right)\right) \le C_{A,\kappa,\mathcal{D}}|\log(h_L)|$. Here and below, $\lambda_{\max}(\,\cdot\,)$, $\lambda_{\min}(\,\cdot\,)$ denote the largest and smallest eigenvalue of an operator on V_L.

Similarly, we obtain by [42, Lem. 8] and (3.7), for every $v_L \in V_L$,

$$\left(B_L^{-1}v_L, v_L\right)_0 \le \widetilde{C}_{\mathcal{D}}\left(\left(-\Delta_L + \varepsilon^{-2}\mathcal{I}_L\right)v_L, v_L\right)_0 \le C_{A,\mathcal{D}}\left(\left(\mathcal{A}_L + \varepsilon^{-2}\mathcal{I}_L\right)v_L, v_L\right)_0,$$

where again the constant $C_{A,\mathcal{D}} := \widetilde{C}_{\mathcal{D}}\max\left\{1, a_-^{-1}\right\} > 0$ is independent of $\varepsilon > 0$. Thus, $\lambda_{\min}\left(B_L\left(\mathcal{A}_L + \varepsilon^{-2}\mathcal{I}_L\right)\right) \ge C_{A,\mathcal{D}}^{-1}$ and the claim of this proposition follows, since $\varkappa\left(B_L\left(\mathcal{A}_L + \varepsilon^{-2}\mathcal{I}_L\right)\right) = \frac{\lambda_{\max}(B_L(\mathcal{A}_L+\varepsilon^{-2}\mathcal{I}_L))}{\lambda_{\min}(B_L(\mathcal{A}_L+\varepsilon^{-2}\mathcal{I}_L))}$. □

Remark 3.3. Assumption (3.6) is satisfied by Jacobi and Gauss–Seidel smoothers, cf. [42, Thm. 1 and Rem. 2]. Thus, by Proposition 3.1 they may be used with respect to the shifted operator $\mathcal{A}_L + \varepsilon^{-2}\mathcal{I}_L$ for implementations in practice.

Recall the bilinear form a_η with corresponding energy norm $\|\cdot\|_{a,\eta}$, depending on the parameter $\eta > 0$, from (2.13)–(2.14). For given $f_L \in V_L$ and $\varepsilon > 0$, we wish to approximate the solution $u_L \in V_L$ of the parametric linear system

$$a_{\varepsilon^2}(u_L, v_L) = \left(\left(\varepsilon^2\mathcal{A}_L + \mathcal{I}_L\right)u_L, v_L\right)_0 = (f_L, v_L)_0 \quad \forall v_L \in V_L \tag{3.8}$$

by an iterative solver which converges at a rate independent of ε and $L \in \mathbb{N}$.

Proposition 3.2. *Suppose that the coefficients* A, κ *of the differential operator* $\mathcal{A}$ *in* (3.1) *satisfy Assumption 3.1 and that the smoothers* $\{R_\ell\}_{\ell\ge 0}$ *in* (3.5) *satisfy* (3.6)*. Let* $L \in \mathbb{N}$*,* $n_0 \ge 1$ *be given and set the number of iterations to* $n = \left\lceil n_0\sqrt{|\log(h_L)|}\,\right\rceil$*. Let* $U_L^0 \in V_L$ *be an arbitrary initial guess.*

Then, there exist constants $C > 0$*,* $q \in (0,1)$*, independent of* ε*,* n_0*,* h_L *and* U_L^0 *such that the* n*-step BPX preconditioned CG approximation* $U_L^n \in V_L$ *of* u_L *in* (3.8) *satisfies*

$$\|u_L - U_L^n\|_{a,\varepsilon^2} \le Cq^{n_0}\|u_L - U_L^0\|_{a,\varepsilon^2},$$

where $\|\cdot\|_{a,\varepsilon^2}$ *is the norm in* (2.13)*. The computational cost is* $\mathcal{O}(n_0 N_L \sqrt{\log(N_L)})$*, with the constant implicit in* $\mathcal{O}(\cdot)$ *independent of* ε*.*

Proof. For any $U_L^0 \in V_L$ holds (see, e.g., [32, Eq. (11.3.27)])

$$\|u_L - U_L^n\|_{a,\varepsilon^2} \leq 2\left(\frac{\sqrt{\varkappa(B_L(\mathcal{A}_L+\varepsilon^{-2}\mathcal{I}_L))}-1}{\sqrt{\varkappa(B_L(\mathcal{A}_L+\varepsilon^{-2}\mathcal{I}_L))}+1}\right)^n \|u_L - U_L^0\|_{a,\varepsilon^2}. \tag{3.9}$$

By Proposition 3.1, there exists a constant $c>0$, which is independent of ε and h_L, such that the BPX PCG algorithm converges linearly with contraction number $\big(\sqrt{c|\log(h_L)|}-1\big)/\big(\sqrt{c|\log(h_L)|}+1\big)$. The statement

$$\left(\frac{\sqrt{c|\log(h_L)|}-1}{\sqrt{c|\log(h_L)|}+1}\right)^{\sqrt{|\log(h_L)|}} \leq q<1 \tag{3.10}$$

is equivalent to

$$\sqrt{|\log(h_L)|}\log\left(\frac{\sqrt{c|\log(h_L)|}+1}{\sqrt{c|\log(h_L)|}-1}\right) \geq \log(q^{-1}).$$

The facts that $\lim_{x\to\infty} x\log((x+1)/(x-1)) = 2$ and that $x \mapsto x\log((x+1)/(x-1))$ is monotonically decreasing on $(1,\infty)$ imply that (3.10) holds for $q = \exp(-2/\sqrt{c})$. The assertion follows then by the usual convergence estimate of PCG in (3.9). □

Remark 3.4. It is also possible to apply the BPX preconditioner combined with a relaxed Richardson iteration and relaxation parameter in the interval $(0, 2\varkappa^{-1}(B_L(\mathcal{A}_L+\varepsilon^{-2}\mathcal{I}_L)))$ instead of PCG, cf. [65, Prop. 2.3]. Then the statement of Proposition 3.2 holds with number of iterations $n = \lceil n_0\,|\log(h_L)|\,\rceil$ and computational cost $\mathcal{O}(n_0 N_L \log(N_L))$ (with $\mathcal{O}(\cdot)$ independent of $\varepsilon>0$). In the case of a BPX preconditioned relaxed Richardson iteration, the approximate solution depends linearly on the right-hand side, which is generally not the case for PCG. Thus, $\widetilde{\mathcal{Z}}_{h,k}^{\beta}$ in (2.38) generated with BPX preconditioned relaxed Richardson iteration will be a GRF and unbiased.

Remark 3.5. The assertion of Proposition 3.2 verifies Assumption 2.7 for PCG using standard multilevel preconditioning by BPX. Thus, the main result on the consistency error for approximate simulation of the GRF $\mathcal{Z}^{\beta}$ (2.4), Theorem 2.1, holds with Lagrangean FE for GRFs that belong to the class induced by the minimal Assumption 3.1 on the coefficients A and κ. The computational cost for generating one approximate sample $\widetilde{\mathcal{Z}}_{K,h_L}^{\beta}$ of $\mathcal{Z}^{\beta}$ is shown in Table 1, assuming the calibrations $K = \mathcal{O}\big(\log^2(N_L)\big)$ for the number of sinc quadrature nodes, $n_0 = \mathcal{O}(\log(N_L))$ for the number of PCG iterations, and $\widetilde{K} = \mathcal{O}(\log(N_L))$ for the approximation of the matrix square root in (2.23).

In the case that, in addition to Assumption 3.1, $\mathcal{D}$ is convex and the coefficient A is of class $C^{0,1}(\overline{\mathcal{D}})^{n\times n}$, then Assumption 2.5 and, thus, Proposition 2.2 hold with the

rates $\rho_{01} = \rho_{12} = 1/n$ and $\rho_{02} = 2/n$. This follows from elliptic regularity results, see, e.g., [34, Thm. 3.2.1.2], where $\mathcal{P}_{h_\ell}$ may be taken as the L_2-projection Π_ℓ, cf. [59, Eq. (9.27)] and Ref. 12. The corresponding accuracy of the approximation $\widetilde{\mathcal{Z}}^\beta_{K,h_L}$ in $L_2(\Omega; L_2(\mathcal{D}))$ is presented in Table 1 for this case.

Table 1. Computational cost and unbiasedness for simulating $\widetilde{\mathcal{Z}}^\beta_{K,h_L}$ in (2.38) on a Euclidean domain $\mathcal{X} = \overline{\mathcal{D}}$ with BPX. In addition, the $L_2(\Omega; L_2(\mathcal{D}))$-accuracy of this approximation is shown for the case that the operator $\mathcal{A}$ is $H^2(\mathcal{D})$-regular.

Solver	Cost for $\widetilde{\mathcal{Z}}^k_{h_L}$ cf. (2.20)	Total cost for $\widetilde{\mathcal{Z}}^\beta_{K,h_L}$ cf. (2.38)	Unbiased	Accuracy $\lVert \mathcal{Z}^\beta - \widetilde{\mathcal{Z}}^\beta_{K,h_L} \rVert_{L_2(\Omega;L_2)}$
PCG	$\mathcal{O}(N_L \log^{3/2}(N_L))$	$\mathcal{O}(N_L \log^{7/2}(N_L))$	No	$\mathcal{O}\left(N_L^{-1/n(2\beta-n/2-\delta)}\right)$
relaxed Richardson	$\mathcal{O}(N_L \log^2(N_L))$	$\mathcal{O}(N_L \log^4(N_L))$	Yes	

4. GRFs on surfaces

In this section, we consider GRFs on a closed, connected, orientable, smooth, compact 2-surface $\mathcal{M}$ immersed into the Euclidean space $\mathbb{R}^3$ and endowed with the positive surface measure $\nu_\mathcal{M}$ (induced by the first fundamental form). The GRF $\mathcal{Z}^\beta$ with distribution (2.4) is again considered as the solution to the stochastic fractional-order PDE (2.6), where $\mathcal{A}$ is a self-adjoint, second-order, elliptic differential operator in divergence form, i.e., $w \mapsto \mathcal{A}w = -\nabla_\mathcal{M} \cdot (A\nabla_\mathcal{M} w) + \kappa^2 w$ for $w \in \mathscr{D}(\mathcal{A})$, where $\nabla_\mathcal{M} \cdot$ and $\nabla_\mathcal{M}$ denote the surface divergence and gradient. In this case, A is a symmetric tensor field, i.e., $A(x)\colon T_x\mathcal{M} \to T_x\mathcal{M}$ is linear and symmetric for almost all $x \in \mathcal{M}$, where $T_x\mathcal{M}$ denotes the tangent space at x. Below, we collect all assumptions on the coefficients A and κ. In order to use pseudodifferential calculus, we assume that they are smooth. However, we add that the preconditioning results of this section only require A and κ to be positive, bounded, and measurable.

Assumption 4.1. The symmetric tensor field A and $\kappa\colon \mathcal{M} \to \mathbb{R}$ are smooth. Furthermore, there exist constants $a_-, \kappa_- > 0$ such that, for almost all $x \in \mathcal{M}$,

$$|\kappa(x)| \geq \kappa_-, \qquad \forall \xi \in T_x\mathcal{M}: \quad \xi^\top A(x)\xi \geq a_- |\xi|^2 \,. \tag{4.1}$$

Under Assumption 4.1, the Lax–Milgram lemma implies that $\mathcal{A}\colon \dot{H}^1_\mathcal{A} \to \dot{H}^{-1}_\mathcal{A}$ is boundedly invertible, with $\dot{H}^\sigma_\mathcal{A}$ defined as in (2.1). Furthermore, $\mathcal{A} \in OPS^2_{1,0}(\mathcal{M})$ is a self-adjoint, elliptic pseudodifferential operator of second order. By the spectral theorem it admits a countably infinite sequence $\{(\lambda_j, e_j)\}_{j\in\mathbb{N}}$ of eigenpairs with eigenfunctions $\{e_j\}_{j\in\mathbb{N}} \subset C^\infty(\mathcal{M})$ constituting an orthonormal basis of $L_2(\mathcal{M}, \nu_\mathcal{M})$ when normalized in $L_2(\mathcal{M}, \nu_\mathcal{M})$, i.e., $(e_j, e_{j'}) = \delta_{jj'}$, where $\delta_{jj'}$ denotes the Kronecker delta.

It follows from [62, Thms. XII.1.3, XII.2.1] that, for all $\beta \in \mathbb{R}$, the fractional power $\mathcal{A}^\beta$ of $\mathcal{A} \in OPS^2_{1,0}(\mathcal{M})$ is well-defined and (using that $\dim(\mathcal{M}) = 2$) that $\lambda_j(\mathcal{A}^\beta) \sim j^\beta$ as $j \to \infty$. In particular, $\alpha = 1$ in (2.5). Thus, by Lemma 2.2 the SPDE (2.6) admits a unique solution in $L_p(\Omega; L_2(\mathcal{M}, \nu_\mathcal{M}))$ for any $\beta > 1/2$.

Furthermore, as $e_j \in C^\infty(\mathcal{M})$, for every $f \in C^\infty(\mathcal{M})^*$ and every $j \in \mathbb{N}$, we may define $\widehat{f}_j = \langle f, e_j \rangle$ (with $\langle \cdot, \cdot \rangle$ denoting the extension of the $L_2(\mathcal{M}, \nu_\mathcal{M})$ inner product to the $C^\infty(\mathcal{M})^* \times C^\infty(\mathcal{M})$-duality pairing). Then [62, Ex. XII.2.1]

$$\forall \sigma \in \mathbb{R}: \quad f \in H^\sigma(\mathcal{M}) \iff \sum_{j\in\mathbb{N}} \big|\widehat{f}_j\big|^2 j^\sigma < \infty \iff \lim_{N\to\infty} \sum_{j=1}^{N} \widehat{f}_j \varphi_j = f \text{ in } H^\sigma(\mathcal{M}).$$

In particular, we have the isomorphy

$$\forall \sigma \in \mathbb{R}: \quad H^\sigma(\mathcal{M}) \cong \dot{H}^\sigma_\mathcal{A}, \tag{4.2}$$

and the norms are equivalent with constants depending only on σ and $\mathcal{M}$.

As $\mathcal{M}$ was assumed compact, for every $\beta > 0$, $\mathcal{A}^{-\beta} \in OPS^{-2\beta}_{1,0}(\mathcal{M})$ is a compact, self-adjoint operator on $L_2(\mathcal{M}, \nu_\mathcal{M})$, since

$$\mathcal{A}^{-\beta} L_2(\mathcal{M}, \nu_\mathcal{M}) = H^{2\beta}(\mathcal{M}) \hookrightarrow L_2(\mathcal{M}, \nu_\mathcal{M}),$$

with the last injection being compact by Rellich's theorem [62, Ch. I.5, Eq. (I.5.15)].

Remark 4.1. If the assumption that the coefficients A and κ are smooth is dropped, then by the Lax–Milgram lemma $\mathcal{A}\colon \dot{H}^1_\mathcal{A} \to \dot{H}^{-1}_\mathcal{A}$ is still boundedly invertible and the equivalence $H^\sigma(\mathcal{M}) \cong \dot{H}^\sigma_\mathcal{A}$ follows for $\sigma \in [-1, 1]$. Furthermore, also the spectral asymptotics (2.5) with $\alpha = 1$ can be shown by exploiting the min-max principle (which shows that $a_- \lambda_j(-\Delta_\mathcal{M}) + \kappa_-^2 \le \lambda_j(\mathcal{A}) \le a_+ \lambda_j(-\Delta_\mathcal{M}) + \kappa_+^2$) combined with the corresponding asymptotic behavior of the eigenvalues for the negative Laplace–Beltrami operator $-\Delta_\mathcal{M}$ on $\mathcal{M}$.

For the finite element discretization, we suppose available an inscribed polyhedron and denote its surface by $\widetilde{\mathcal{M}}$ with plane faces $\{F_1, \dots, F_J\}$ which is enclosed by $\mathcal{M}$ and which is such that the vertices of $\widetilde{\mathcal{M}}$ are situated on the surface $\mathcal{M}$. The surface measure of $\widetilde{\mathcal{M}}$ is denoted by $\nu_{\widetilde{\mathcal{M}}}$. Moreover, we assume that there exists a bijective, piecewise smooth mapping $\mathcal{F}\colon \widetilde{\mathcal{M}} \to \mathcal{M}$ such that the Jacobian $D\mathcal{F}$ is invertible a.e. The FE spaces on $\mathcal{M}$ will be defined patchwise, where the patches are given by $\mathcal{F}(F_i) \subset \mathcal{M}$, $i = 1, \dots, J$. Let $\widetilde{V}_h$ be a first order FE space on $\widetilde{\mathcal{M}}$ with maximal mesh width h, i.e., let it be the space of piecewise affine functions on a triangulation of $\widetilde{\mathcal{M}}$. A FE space V on $\mathcal{M}$ results by

$$V_h := \big\{ v \in \dot{H}^1_\mathcal{A} : \exists \widetilde{v}_h \in \widetilde{V}_h \text{ s.t. } v_h \circ \mathcal{F} = \widetilde{v}_h \big\}.$$

The bilinear form corresponding to $\mathcal{A}$, upon pullback in local coordinates to

$\widetilde{\mathcal{M}} \subset \mathbb{R}^3$, is, for $v, w \in \dot{H}^1_{\mathcal{A}}$, given by

$$\int_{\mathcal{M}} \left[A\nabla_{\mathcal{M}} v \cdot \nabla_{\mathcal{M}} w + \kappa^2 vw\right] \mathrm{d}\nu_{\mathcal{M}} = \int_{\widetilde{\mathcal{M}}} \left[(D\mathcal{F})^{-\top} \widetilde{A} (D\mathcal{F})^{-1} \nabla_{\widetilde{\mathcal{M}}} \widetilde{v} \cdot \nabla_{\widetilde{\mathcal{M}}} \widetilde{w} + \widetilde{\kappa}^2 \widetilde{v}\widetilde{w}\right] \mathrm{d}\nu^{\mathcal{F}}_{\widetilde{\mathcal{M}}}, \quad (4.3)$$

where $\mathrm{d}\nu^{\mathcal{F}}_{\widetilde{\mathcal{M}}} = \sqrt{\det((D\mathcal{F})^\top (D\mathcal{F}))}\,\mathrm{d}\nu_{\widetilde{\mathcal{M}}}$ and the composition of a function g with $\mathcal{F}$ is denoted by $\widetilde{g}$. In particular, also the $L_2(\mathcal{M}, \nu_{\mathcal{M}})$-inner product may be pulled back and results in a weighted $L_2(\widetilde{\mathcal{M}}, \nu_{\widetilde{\mathcal{M}}})$-inner product with weight $\sqrt{\det((D\mathcal{F})^\top (D\mathcal{F}))}$.

Multilevel preconditioning by BPX can be analyzed via the correspondence in (4.3). Let $\widetilde{V}_\ell$, $\ell \geq 0$ be a nested sequence of FE spaces that result by uniformly refining an initial triangulation on $\widetilde{\mathcal{M}}$ with maximal mesh width h_ℓ. The corresponding nested FE spaces on $\mathcal{M}$ are denoted by V_ℓ, $\ell \geq 0$. Denote by $\mathcal{A}_\ell$, $\ell \geq 0$, the FE-discretizations of $\mathcal{A}$. Recall that $\dot{H}^0_{\mathcal{A}} = L_2(\mathcal{M}, \nu_{\mathcal{M}})$ with norm $\|\cdot\|_0 = \|\cdot\|_{L_2(\mathcal{M},\nu_{\mathcal{M}})}$ and consider the BPX preconditioner with SPD smoothers R_ℓ, $\ell = 0, \dots, L$, which satisfy for some constant $C > 0$

$$C^{-1}(R_\ell v_\ell, v_\ell)_0 \leq h_\ell^2 \left(\|\nabla_{\mathcal{M}} v_\ell\|_0^2 + \varepsilon^{-2} h_\ell^2 \|v_\ell\|_0^2\right) \leq C(R_\ell v_\ell, v_\ell)_0. \quad (4.4)$$

This property is satisfied, e.g., for Jacobi smoothers, see Subsection 3.3 for details.

Proposition 4.1. *Suppose that the coefficients A and κ are bounded, measurable, and satisfy* (4.1)*. Further, suppose that the smoothers $\{R_\ell\}_{\ell \geq 0}$ satisfy* (4.4)*. There exists a constant $C > 0$, independent of $\varepsilon \in (0, \infty)$, such that*

$$\varkappa(B_L(\mathcal{A}_L + \varepsilon^{-2}\mathcal{I})) \leq C|\log(h_L)| \quad \forall L \in \mathbb{N}.$$

Proof. We begin by verifying the statements of [42, Lem. 6 and 8] in the case of the operator $-\Delta_{\widetilde{\mathcal{M}}} + \varepsilon^{-1}\mathcal{I}$ on $\widetilde{\mathcal{M}}$, where $\Delta_{\widetilde{\mathcal{M}}}$ denotes the Laplace–Beltrami on $\widetilde{\mathcal{M}}$. Denote by $\Delta_{\widetilde{\mathcal{M}},L}$ the FE discretized Laplace–Beltrami operator on level L.

The proof of [42, Lem. 6] is based on a *strengthened Cauchy–Schwarz* inequality, cf. [65, Lem. 6.1]. Since the strengthened Cauchy–Schwarz inequality is proven in Ref. 65 element-wise without requiring boundary conditions, the same proof is also applicable for boundaries of polyhedra. Thus, as in [42, Lem. 6], there exists a constant $C > 0$, which does not depend on h_L, such that for every $\widetilde{v} \in \widetilde{V}_L$

$$\left((-\Delta_{\widetilde{\mathcal{M}},L} + \varepsilon^{-2}\mathcal{I})\widetilde{v}, \widetilde{v}\right)_0 \leq C|\log(h_L)| \left(\widetilde{B}_L^{-1}\widetilde{v}, \widetilde{v}\right)_0. \quad (4.5)$$

Here, $\widetilde{B}_L$ denotes the corresponding linear operator on the FE space $\widetilde{V}_L$.

The proof of [42, Lem. 8] is based on an identity from the abstract theory of *parallel subspace correction*, cf. [66, Lem. 2.4], which implies with the assumption on B_L in (4.4) a version of [42, Eq. (4.7)]. For all $\widetilde{v} \in \widetilde{V}_L$, we have

$$\begin{aligned} C^{-1}(\widetilde{B}_L^{-1}\widetilde{v}, \widetilde{v})_0 &\leq \inf_{\sum_{\ell=0}^L \widetilde{v}_\ell = \widetilde{v}} \left\{((-\Delta_{\widetilde{\mathcal{M}},L} + \varepsilon^{-2}\mathcal{I})\widetilde{v}_0, \widetilde{v}_0)_0 + \sum_{\ell=1}^L h_\ell^{-2} \|\widetilde{v}_\ell\|_0^2\right\} \\ &\leq C(\widetilde{B}_L^{-1}\widetilde{v}, \widetilde{v})_0, \end{aligned} \quad (4.6)$$

where $C > 0$ is a constant that does not depend on L. For $\ell \in \{0, \dots, L\}$, let $\widetilde{\Pi}_\ell$ denote L_2-projection onto $\widetilde{V}_\ell$. The next step is to verify that there exists a constant $C > 0$ which does not depend on h_ℓ, $\ell = 0, \dots, L$, such that, for every $\widetilde{v} \in \widetilde{V}_L$,

$$\|\nabla_{\widetilde{\mathcal{M}}} \widetilde{\Pi}_0 \widetilde{v}\|_0^2 + \sum_{\ell=1}^{L} h_\ell^{-2} \|(\widetilde{\Pi}_\ell - \widetilde{\Pi}_{\ell-1})\widetilde{v}\|_0^2 \leq C \|\nabla_{\widetilde{\mathcal{M}}} \widetilde{v}\|_0^2. \tag{4.7}$$

The estimate (4.7) has been verified in the Euclidean case in [11, Eq. (4.5)]. The particular ingredients of the proof are estimates [11, Eq. (4.6) and (4.7)]. A first order approximation property of the L_2-projection $\widetilde{\Pi}_\ell$ and stability in the $H^1(\widetilde{\mathcal{M}})$-norm results in [11, Eq. (4.6)]. In our case, for the approximation property of Π_ℓ see, e.g., [59, Thm. 10.2] and for the stability in the $H^1(\widetilde{\mathcal{M}})$-norm see, e.g., [59, Lem. 9.13] (and also Ref. 12) for the Euclidean case, which can be applied patchwise. Furthermore, since $-\Delta_{\widetilde{\mathcal{M}}}+1$ is boundedly invertible from $H^{1+\varepsilon'}(\widetilde{\mathcal{M}}) \to H^{-1+\varepsilon'}(\widetilde{\mathcal{M}})$ for some $\varepsilon' > 0$ and since the FE spaces $\widetilde{V}_\ell$ satisfy an approximation property (see, e.g., [59, Lem. 10.8]) the estimate [11, Eq. (4.7)] holds with $\alpha = \varepsilon'$ in the notation of Ref. 11. As a result, the proof of [11, Eq. (4.5)] is applicable line by line and we conclude (4.7). Inserting the decomposition $\widetilde{v} = \widetilde{\Pi}_0 + \sum_{\ell=1}^{L} (\widetilde{\Pi}_\ell - \widetilde{\Pi}_{\ell-1})\widetilde{v}$ into (4.6), we find with (4.7) the following version of [42, Lem. 8]: There exists a constant $C > 0$, which does not depend on L and $\varepsilon \in (0, \infty)$, such that, for all $\widetilde{v} \in \widetilde{V}_L$,

$$\big(\widetilde{B}_L^{-1} \widetilde{v}, \widetilde{v}\big)_0 \leq C \big((-\Delta_{\widetilde{\mathcal{M}},L} + \varepsilon^{-2}\mathcal{I})\widetilde{v}, \widetilde{v}\big)_0. \tag{4.8}$$

Recall that the proof of Proposition 3.1 relied on the estimates from [42, Lem. 6 and 8], which we have derived for the setting here as (4.5) and (4.8). By (4.3) the FE operator on V_ℓ corresponding to $\mathcal{A} + \varepsilon^{-2}\mathcal{I}$ can be identified with a symmetric, elliptic FE operator with variable coefficients on $\widetilde{V}_\ell$, given by

$$\widetilde{V}_\ell \times \widetilde{V}_\ell \ni (\widetilde{v}, \widetilde{w}) \mapsto \int_{\widetilde{\mathcal{M}}} \Big[(D\mathcal{F})^{-\top} \widetilde{A} (D\mathcal{F})^{-1} \nabla_{\widetilde{\mathcal{M}}} \widetilde{v} \cdot \nabla_{\widetilde{\mathcal{M}}} \widetilde{w} + (\widetilde{\kappa}^2 + \varepsilon^{-2}) \widetilde{v} \widetilde{w} \Big] \, \mathrm{d}\nu_{\widetilde{\mathcal{M}}}^{\mathcal{F}}.$$

Thus, the proof of Proposition 3.1 is applicable, where we point out that here also variable coefficients appear in the shifted term ε^{-2}; that is due to the transformation of the measure. This constitutes a simple extension of Proposition 3.1 and we omit the details. □

Remark 4.2. The error bounds of Proposition 3.2 for PCG with BPX hold also in the setting of this section. Thus, Assumption 2.7 is satisfied and the unified consistency error bound of Theorem 2.1 applies also for the approximate simulation of the GRF $\mathcal{Z}^\beta$ with distribution (2.4) on a closed, connected, orientable, smooth, compact 2-surface $\mathcal{M}$.

Remark 4.3. The sinc-FE algorithm discussed in this section can also be considered on piecewise smooth Lipschitz surfaces. The presented analysis may be reconciled. We refer to [22, p. 2259] for gaining an impression of such an extension which would partly be beyond classical theory of pseudodifferential operators on smooth surfaces as, e.g., presented in Ref. 62.

Remark 4.4. For the particular choice $\mathcal{F}(x) := x - \mathrm{d}(x)\nabla\mathrm{d}(x)$, where d is the signed distance function with a negative sign inside $\mathcal{M}$, explicit formulae have been derived in [24, Eq. (2.10) and (2.12)] that allow for an implementation of the algorithm provided d is accessible. The signed distance function admits a closed form expression in the case of $\mathcal{M} = \mathbb{S}^2$, i.e., the 2-sphere. FE approximations have been computationally realized in [37, Sec. 6]. Thus, fast sampling of *non-isotropic* GRFs on $\mathbb{S}^2$ is admissible by the algorithms presented in this work. The simulation of isotropic GRFs, whose covariance structures are invariant under rotations, was previously treated with fast Fourier techniques in Refs. 19, 43. In addition, the presently developed methods extend in a unified way to any smooth surface that is homeomorphic to the 2-sphere.

5. Wavelet preconditioning

The BPX preconditioning techniques of the previous Sections 3–4 allowed under Assumption 3.1 (respectively, under (4.1) as part of Assumption 4.1, see also Remark 4.1) for an iterative solution of the linear systems (2.21) at accuracy $\mathcal{O}(N_h^{-\rho})$ for a prescribed consistency order $\rho > 0$ (limited by the regularity of the GRF $\mathcal{Z}^\beta$ and the order afforded by the Galerkin discretization), cf. Proposition 2.2 and Theorem 2.1. The incurred computational costs in work and memory are essentially $\mathcal{O}(N_h)$, see Table 1. As we have explained, BPX preconditioning applies to unstructured regular triangulations of $\mathcal{X}$ in both cases, for a Euclidean domain $\mathcal{X} = \overline{\mathcal{D}}$ and for a 2-surface $\mathcal{X} = \mathcal{M}$. In the following, we present an alternative approach which is based on the (assumed) availability of a *multiresolution analysis* (MRA) in the spaces $L_2(\mathcal{X}, \nu_\mathcal{X})$ and $\dot{H}^1_\mathcal{A}$, see (2.1). To streamline the discussion, we provide all statements and proofs in a unified fashion, with $\mathcal{X} \in \{\overline{\mathcal{D}}, \mathcal{M}\}$, where $\mathcal{D} \subset \mathbb{R}^n$ is as in Section 3, and where $\mathcal{M}$ is as in Section 4. Throughout this section, we work under Assumptions 3.1 and 4.1, respectively.

We recapitulate, from Refs. 22, 52, 61, notation and basic results on wavelet Galerkin discretizations of operator equations. We limit the presentation to Riesz bases, and norm equivalences required for the multilevel preconditioning of the parametric linear systems (2.21). We remark that having at hand a wavelet basis $\mathbf{\Psi}$ of the spaces $\dot{H}^s_\mathcal{A}$ which is stable in a range of Sobolev orders $s \in (-\tilde{\gamma}, \gamma)$ for some $\tilde{\gamma} > 0$, $\gamma > 1$ allows *compressive numerical approximation* of samples of the GRF $\mathcal{Z}^\beta$ in (2.4). Wavelet-based approximation methods can, therefore, in principle also exploit more regularity of the GRF $\mathcal{Z}^\beta$ that is afforded, for example, by a higher fractional exponent $\beta > 0$, see [18, Lem. 4.1].

Wavelet preconditioning (see, e.g., Refs. 21, 61 and the references there) is based on the observation that *wavelet FEM provide Riesz bases for $\dot{H}^\sigma_\mathcal{A}$ for a range of σ around $\sigma = 0$*. As is customary (see, e.g. Ref. 22), we let $\mathbf{\Psi} = \{\psi_\lambda : \lambda \in \mathcal{J}\}$ denote a countable collection of wavelets. Here, $\mathcal{J}$ is a countable index set of multi-indices $\lambda = (\ell, k, e)$, with $\ell(\lambda) = |\lambda| \in \mathbb{N}_0$ denoting the *scale of resolution*, $k(\lambda)$ denoting the *spatial location* of wavelet ψ_λ in $\mathcal{X}$ and (optionally) $e(\lambda)$ is the *type of wavelet*.

In particular, we view $\mathbf{\Psi}$ as *infinite collection* of wavelets; Galerkin approximations will of course only use finite sections of these which correspond to a finite set $\Lambda \subset \mathcal{J}$ of "active indices" $\lambda \in \Lambda$.

We require $\mathbf{\Psi}$ to be a Riesz basis of $L_2(\mathcal{X}, \nu_\mathcal{X})$, which is to say that for every $v \in L_2(\mathcal{X}, \nu_\mathcal{X})$, there exists a unique vector $\mathbf{v} \in \ell_2(\mathcal{J})$ such that

$$v = \sum_{\lambda \in \mathcal{J}} \mathbf{v}_\lambda \psi_\lambda = \mathbf{v}^\top \mathbf{\Psi} \tag{5.1}$$

holds, with equality in $L_2(\mathcal{X}, \nu_\mathcal{X})$. We shall require systems $\mathbf{\Psi}$ which are localized, i.e.,

$$\mathcal{X}_\lambda := \operatorname{supp}(\psi_\lambda) \subset \mathcal{X}, \quad \operatorname{diam}(\mathcal{X}_\lambda) \simeq 2^{-|\lambda|}, \quad \lambda \in \mathcal{J}. \tag{5.2}$$

For a function $v \in L_2(\mathcal{X}, \nu_\mathcal{X})$, the vector $\mathbf{v} \in \ell_2(\mathcal{J})$ of coefficients with respect to the multiresolution system $\mathbf{\Psi}$ characterizes the regularity of v on the scale of spaces $\{\dot{H}^\sigma_\mathcal{A} : \sigma \in (-\widetilde{\gamma}, \gamma)\}$ for some constants $\gamma, \tilde{\gamma} > 0$ (which depend on $\mathbf{\Psi}$). To formally state this property, we introduce bi-infinite *diagonal scaling matrices* $\mathbf{D} = \operatorname{diag}\{w_\lambda : \lambda \in \mathcal{J}\}$ with diagonal elements $w_\lambda > 0$. More generally, for $s \in \mathbb{R}$, we write $\mathbf{D}^s = \operatorname{diag}\big\{2^{|\lambda| s} : \lambda \in \mathcal{J}\big\}$. The key observation for preconditioning is then the following *norm equivalence* on $\dot{H}^\sigma_{-\Delta_\mathcal{X}+\mathcal{I}}$. For $\sigma \in (-\widetilde{\gamma}, \gamma)$, there exist constants $0 < c_\sigma \leq C_\sigma < \infty$ such that the coefficient vectors in (5.1) satisfy

$$c_\sigma \|\mathbf{v}\|_{\ell_2} \leq \|\mathbf{v}^\top \mathbf{D}^{-\sigma} \mathbf{\Psi}\|_{\dot{H}^\sigma_{-\Delta_\mathcal{X}+\mathcal{I}}} \leq C_\sigma \|\mathbf{v}\|_{\ell_2} \quad \forall \mathbf{v} \in \ell_2(\mathcal{J}). \tag{5.3}$$

In order not to clutter the exposition, we restrict ourselves here to compactly supported wavelets and remark that certain wavelets with sufficient decay also allow norm equivalences of type (5.3), see [51, Section 6.10]; the unbounded support entails significant technical repercussions.

Equation (5.3) implies in particular that *the scaled basis* $\mathbf{D}^{-\sigma}\mathbf{\Psi}$ *is a Riesz basis for* $\dot{H}^\sigma_\mathcal{A}$ in (2.1), i.e., every $v \in \dot{H}^\sigma_\mathcal{A}$ admits a representation $v = \mathbf{v}^\top \mathbf{D}^{-\sigma}\mathbf{\Psi}$ with a unique $\mathbf{v} \in \ell_2(\mathcal{J})$ such that (5.3) holds.

An immediate consequence of (5.3) is the following result on *robust wavelet preconditioning*.

Proposition 5.1. *Let* $\mathbf{\Psi}$ *be a Riesz basis for* $L_2(\mathcal{X}, \nu_\mathcal{X})$ *which satisfies* (5.3) *for some* $\widetilde{\gamma} > 0$ *and* $\gamma > 1$. *For* $\varepsilon > 0$, *define on* $\dot{H}^1_\mathcal{A}$ *the parametric norm* $\|\cdot\|_{H_\varepsilon(\mathcal{X})}$ *by*

$$\|v\|^2_{H_\varepsilon(\mathcal{X})} := \varepsilon^2 \big[(\nabla_\mathcal{X} v, \nabla_\mathcal{X} v)_{L_2(\mathcal{X}, \nu_\mathcal{X})} + (v, v)_{L_2(\mathcal{X}, \nu_\mathcal{X})}\big] + (v, v)_{L_2(\mathcal{X}, \nu_\mathcal{X})}, \tag{5.4}$$

where $\nabla_\mathcal{X}$ *denotes the Euclidean gradient if* $\mathcal{X} = \overline{\mathcal{D}}$ *and the surface gradient if* $\mathcal{X} = \mathcal{M}$. *Then, for every* $\varepsilon > 0$, *with the constants* $c_\sigma, C_\sigma > 0$ *from* (5.3), *and with the* bi-infinite diagonal preconditioning matrix

$$\mathbf{D}_\varepsilon := \operatorname{diag}\big\{\big(1 + \varepsilon 2^{|\lambda|}\big) : \lambda \in \mathcal{J}\big\}, \tag{5.5}$$

we have, for every $\mathbf{v} \in \ell_2(\mathcal{J})$ *and for all* $\varepsilon > 0$,

$$\big(2\left(c_0^{-2} + c_1^{-2}\right)\big)^{-1/2} \|\mathbf{v}\|_{\ell_2} \leq \|\mathbf{v}^\top \mathbf{D}_\varepsilon^{-1} \mathbf{\Psi}\|_{H_\varepsilon(\mathcal{X})} \leq \big(C_0^2 + C_1^2\big)^{1/2} \|\mathbf{v}\|_{\ell_2}. \tag{5.6}$$

If, in addition, Assumption 3.1 (or Assumption 4.1) is satisfied, then there exist constants $0 < \check{c} \leq \hat{c} < \infty$ *(depending on* $a_\pm$*,* $\kappa_\pm$ *in Assumption 3.1/4.1, and on* $c_i, C_i > 0$*,* $i = 0, 1$*, in the norm equivalence* (5.3)*) such that for the parametric bilinear form* $a_\eta(\,\cdot\,,\,\cdot\,)$ *in* (2.13) *we have, for any* $\varepsilon > 0$*,*

$$\forall \mathbf{v} \in \ell_2(\mathcal{J}): \quad \check{c}^2 \|\mathbf{v}\|_{\ell_2}^2 \leq a_{\varepsilon^2}\big(\mathbf{v}^\top \mathbf{D}_\varepsilon^{-1}\boldsymbol{\Psi}, \mathbf{v}^\top \mathbf{D}_\varepsilon^{-1}\boldsymbol{\Psi}\big) \leq \hat{c}^2 \|\mathbf{v}\|_{\ell_2}^2 \,. \tag{5.7}$$

In particular, there exists a constant $\bar{c}$ *such that, for every* $\varepsilon > 0$*,*

$$\varkappa\left(\mathbf{D}_\varepsilon^{-1}\left(\mathbf{M}^{\boldsymbol{\Psi}} + \varepsilon^2 \mathbf{A}^{\boldsymbol{\Psi}}\right)\mathbf{D}_\varepsilon^{-1}\right) \leq \bar{c}, \tag{5.8}$$

where $\mathbf{M}^{\boldsymbol{\Psi}}$ *and* $\mathbf{A}^{\boldsymbol{\Psi}}$ *are the Gramian and operator matrix in* (2.22) *with respect to the wavelet basis* $\boldsymbol{\Psi}$*.*

Proof. Statement (5.6) can be proven as in [21, Rem. 1] by replacing the H^1 semi-norm by the H^1-norm in the proof of [21, Rem. 1], which is justified by our assumptions on $\boldsymbol{\Psi}$ in (5.3).

To show (5.7)–(5.8), we observe that under Assumption 3.1/4.1 the parametric bilinear form $a_{\varepsilon^2}(\,\cdot\,,\,\cdot\,) = \varepsilon^2(\nabla\,\cdot\,, A\nabla\,\cdot\,)_0 + (\,\cdot\,, (1+\varepsilon^2\kappa^2)\,\cdot\,)_0$ corresponding to (3.4) is uniformly equivalent to the parametric norm $\|\,\cdot\,\|_{H_\varepsilon(\mathcal{X})}$ in (5.4) with respect to $\varepsilon > 0$. We will show this first for $\mathcal{X} = \mathcal{M}$, i.e., in the setting of Assumption 4.1. Specifically,

$$\forall v \in \dot{H}^1_{\mathcal{A}}: \quad \min\{a_-, \kappa_-^2, 1\}\|v\|^2_{H_\varepsilon(\mathcal{X})} \leq a_{\varepsilon^2}(v,v) \leq \max\{a_+, \kappa_+^2, 1\}\|v\|^2_{H_\varepsilon(\mathcal{X})}.$$

In the case $\mathcal{X} = \overline{\mathcal{D}}$, Assumption 3.1 permits $\operatorname{ess\,inf}_{x\in\mathcal{D}}\{\kappa^2(x)\} = \kappa_-^2 = 0$, since the boundary conditions (homogeneous Dirichlet) facilitate a Poincaré inequality with constant $C_{\mathcal{D}} > 0$, i.e., $\|v\|_{L_2(\mathcal{D})} \leq C_{\mathcal{D}}\|\nabla v\|_{L_2(\mathcal{D})}$ for every $v \in \dot{H}^1_{\mathcal{A}}$. As a consequence, the lower bound of $a_{\varepsilon^2}(v,v)$ uniformly with respect to $\varepsilon > 0$ follows,

$$\forall v \in \dot{H}^1_{\mathcal{A}}: \quad \min\left\{\tfrac{a_-}{2}(1 + C_{\mathcal{D}}^{-2}), 1\right\}\|v\|^2_{H_\varepsilon(\mathcal{X})} \leq a_{\varepsilon^2}(v,v) \leq \max\{a_+, \kappa_+^2, 1\}\|v\|^2_{H_\varepsilon(\mathcal{X})}.$$

Thus, assertions (5.7) and (5.8) follow from (5.6). □

Proposition 5.1 and its proof do not require the smoothness assumption on A and κ^2 in Assumption 4.1. We add some further comments on Proposition 5.1.

Remark 5.1. For computational purposes, indices of "active" wavelets need to be restricted to finitely many indices contained in some index set $\Lambda \subset \mathcal{J}$. Suitable choices of Λ can accommodate, for example, local mesh refinement in bounded domains that account for corner singularities (see, e.g., Ref. 39 and references there).

The norm equivalences (5.6)–(5.7) of Proposition 5.1 are valid for every infinite coefficient sequence $\mathbf{v} \in \ell_2(\mathcal{J})$. They hold in particular for Λ-*"supported" subsequences*, i.e., for $\mathbf{v}_\Lambda = \{v_\lambda : \lambda \in \Lambda\}$. We furthermore note that the constants in (5.6)–(5.7) and also in the condition number bound (5.8) are independent of Λ. Thus, for $L \in \mathbb{N}$ and $\Lambda_L := \{\lambda \in \mathcal{J} : |\lambda| \leq L\}$ with $\#(\Lambda_L) = \mathcal{O}\big(2^{Ln}\big)$, the norm equivalences (5.6)–(5.7) hold on $\ell_2(\Lambda_L)$ with constants $\check{c}$, $\hat{c}$ and $\bar{c}$ which are independent of the refinement level L and of $\varepsilon > 0$.

Furthermore, by Proposition 5.1 the singularly perturbed, truncated Galerkin projected operators in wavelets $(\mathbf{M}^{\mathbf{\Psi}} + \varepsilon^2 \mathbf{A}^{\mathbf{\Psi}})_\Lambda$ allow for *uniformly optimal diagonal preconditioning* by $\left(\mathbf{D}_\varepsilon^{-1}\right)_\Lambda$, i.e., with condition bounded uniformly in $\varepsilon > 0$. In particular, the occurring condition number bounds hold without log-factors. This therefore improves the BPX results of Propositions 3.1 and 4.1 (which, however, might be an artifact of the method of proof). As a consequence, Assumption 2.7 is satisfied with n_0 denoting the iteration numbers of either PCG or a relaxed Richardson iteration. The system matrices that are Λ_L-sections of $\mathbf{M}^{\mathbf{\Psi}}$ and $\mathbf{A}^{\mathbf{\Psi}}$ have $\mathcal{O}\big(2^{nL}L\big)$ non-zero entries. However, by compression techniques from Ref. 22 exploiting certain *vanishing moment* properties of the MRA, a consistent *approximate* matrix vector product can be realized that only uses $\mathcal{O}\big(2^{nL}\big)$ matrix entries. Thus, the overall computational cost for the sinc quadrature combined with MRA is in the setting of Table 1 reduced to $\mathcal{O}(N_L \log(N_L))$ for $\widetilde{\mathcal{Z}}^k_{h_L}$ and to $\mathcal{O}\big(N_L \log^3(N_L)\big)$ for $\widetilde{\mathcal{Z}}^\beta_{K,h_L}$ at the same accuracy.

Remark 5.2. The preconditioning result requires only the Riesz basis property of $\mathbf{\Psi}$; it *does not* suppose specific properties on the dual wavelet system $\widetilde{\mathbf{\Psi}}$ (see, e.g., Refs. 21, 61 for a definition). Such assumptions usually appear in the analysis and synthesis operations in signal processing. This affords extra flexibility in the construction of explicit systems $\mathbf{\Psi}$ which can be exploited for minimizing condition numbers of Gramian matrices and support sizes.

Remark 5.3. We refer to Ref. 56 and the references there for the construction of a continuous, piecewise quadratic spline wavelet basis on regular, simplicial triangulations of bounded polygons $\mathcal{D} \subset \mathbb{R}^2$ with straight sides. This construction affords the stability (5.3) with $\widetilde{\gamma} = \gamma = 3/2$, and also extends to Lipschitz polytopes in $\mathbb{R}^3$.

6. Numerical experiments

In this section we study the numerical approximation of the GRF $\mathcal{Z}^\beta$ in (2.4) on a convex polygon $\mathcal{D}_1 \subset \mathbb{R}^2$, on a non-convex polygon $\mathcal{D}_2 \subset \mathbb{R}^2$, and on the 2-sphere $\mathbb{S}^2 \subset \mathbb{R}^3$. To this end, we consider the operator $\mathcal{A} = -\Delta_\mathcal{X} + \kappa^2$. Here, $\Delta_\mathcal{X}$ denotes the Laplace operator with homogeneous Dirichlet boundary conditions if $\mathcal{X} \in \{\overline{\mathcal{D}}_1, \overline{\mathcal{D}}_2\}$, and, for $\mathcal{X} = \mathcal{M}$, $-\Delta_\mathcal{M}$ is the Laplace–Beltrami operator on $\mathcal{M} = \mathbb{S}^2$. The coefficient function κ^2 will be piecewise constant. We suppose given a sequence of nested FE spaces V_ℓ, $\ell \geq 0$, which results by an uniform refinement of an initial regular triangulation of $\overline{\mathcal{D}}$ in affine triangles or of $\mathbb{S}^2$ in curved triangles as discussed in Section 4. In particular, we realize an *exact* triangulation of $\mathbb{S}^2$, i.e., absence of geometry approximation errors.

6.1. *Approximation of the square root of the mass matrix*

We numerically test the performance of the algorithm discussed in Section 2.2.4 to approximate the square root of the mass matrix. Specifically, we require the action

of this approximation on a vector of i.i.d. $\mathsf{N}(0,1)$-distributed coefficients. In formula (2.23) the required solutions to linear systems are performed with CG. Since the condition number of the mass matrix does not depend on the mesh width for quasi-uniform refinement, CG converges linearly with a convergence rate that is independent of the dimension of the FE space. Also note that the condition number of the sum of two SPD matrices is upper bounded by the sum of the individual condition numbers. The algorithm in (2.23) requires values for $\widehat{M}$ and $\widehat{m}$ that upper and lower bound the spectrum of the mass matrix. These are obtained by multiplying, respectively dividing, the result of 20 iterations of the power method, respectively of the inverse iteration, by the value 1.1. In Figure 1, the relative error is visualized for the approximation of $\sqrt{\mathbf{M}}\mathbf{y}$ in the Euclidean norm, where $\mathbf{y}$ has i.i.d. $\mathsf{N}(0,1)$-distributed entries. In the case $\dim(V_\ell) = 1985$, the approximation can be compared to the exact vector $\sqrt{\mathbf{M}}\mathbf{y}$. For $\dim(V_\ell) = 130561$, we use a reference vector obtained with $\widetilde{K} = 20$ in (2.23). The linear systems in (2.23) were approximately solved by CG with error tolerance 10^{-12}, which also limits the approximation quality. This is visible in Figure 1 in the case $\dim(V_\ell) = 1985$, where the approximation is compared to the exact vector $\sqrt{\mathbf{M}}\mathbf{y}$. These tests were performed for $\mathcal{D}_1 = (0,1)^2$ with homogeneous Dirichlet boundary conditions.

6.2. *Simulation of non-stationary GRFs in bounded Euclidean domains*

We consider two bounded polygons $\mathcal{D}_1, \mathcal{D}_2$ and impose homogeneous Dirichlet boundary conditions. For $\beta \in (0,1)$, the sinc-FE approximation $\mathcal{Z}^\beta_{K,h}$ in (2.18) of $\mathcal{Z}^\beta$ in (2.4) depends on the FE mesh width h and on the number K, where the sinc quadrature has $2K+1$ terms. Since $\mathcal{D}_1 = (0,1)^2$ is convex, the inverse of the shifted Dirichlet Laplacian $(-\Delta + \kappa^2(x))^{-1}$ is bounded from $L_2(\mathcal{D})$ to $H^2(\mathcal{D})$. Consequently, Assumption 2.5 holds with $\rho_{01} = \rho_{12} = 1/2$, $\rho_{02} = 1$, and the FE error is asymptotically (as $N_\ell \to \infty$) $\mathcal{O}\big(N_\ell^{-\beta+1/2+\delta}\big)$, see Proposition 2.2 and Table 1. Balancing the bounds for the sinc quadrature error $\mathcal{O}\big(e^{-c\sqrt{K}}\big)$ for $c = 2\min\{\beta, 1-\beta\}$ yields $K_\ell^\star = ((\beta - 1/2)\log(N_\ell)/c)^2$. We use $K_\ell = 4K_\ell^\star$ in our simulations. In Figure 2, we plot the CPU time for $\beta = 0.75$ as an average of 40 runs, where we apply PCG with BPX and Jacobi smoothers. PCG is stopped with an error tolerance of 10^{-12} and we use $\widetilde{K} = 10$ on all levels. As a comparison, $N_\ell \log^2(N_\ell)$ is also shown, which would be the computational cost if one PDE solve in the sinc formula were of computational cost $\mathcal{O}(N_\ell)$. The slopes are fitted with least squares in log scale. The function $\kappa^2(x_1, x_2)$ is chosen to be κ_1^2 if $x_1 < 1/2$ and κ_2^2 otherwise for $\kappa_1^2, \kappa_2^2 \geq 0$. In Figure 2, $\kappa_1^2 = 20, \kappa_2^2 = 200$.

The relative $L_2(\Omega; L_2(\mathcal{D}_i))$-error, $i = 1, 2$, is studied in Figures 3–4. To this end, the $L_2(\Omega; L_2(\mathcal{D}_i))$-error is approximated with $R = 60$ Monte Carlo samples and with a reference solution, which is obtained for Figure 3 on a FE mesh with two additional levels of refinement with 2095105 degrees of freedom using a sparse direct solver `SuperLU`, see, e.g., Ref. 25, and $K_{\text{ref}} = K_{L+2} + 10$ for the sinc quadrature. The

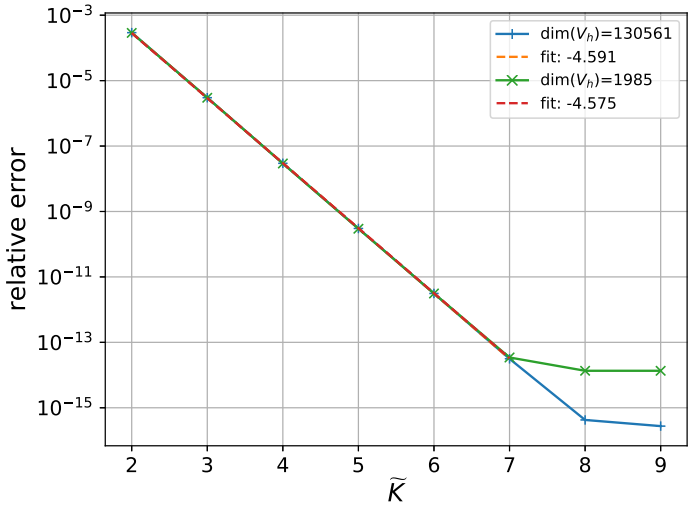

Figure 1. Relative approximation error in Euclidean norm of $\sqrt{\mathbf{M}}\mathbf{y}$

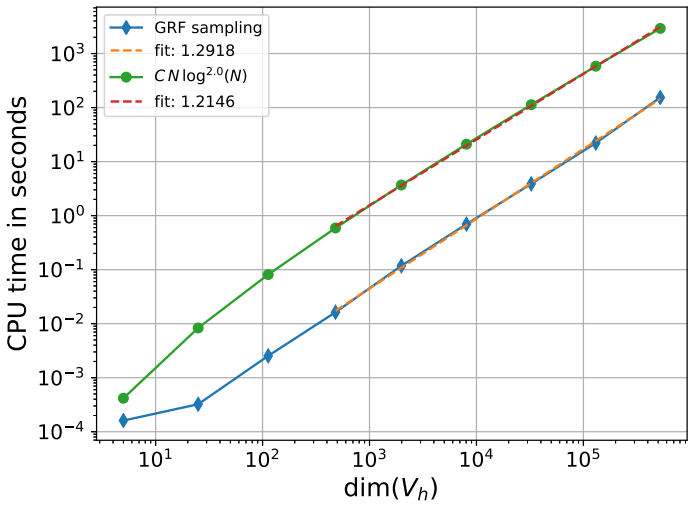

Figure 2. CPU time per sample versus $N_h = \dim(V_h)$ for $\widetilde{\mathcal{Z}}^{\beta}_{K,h}$ for $\beta = 0.75$

random forcing is approximated with $\widetilde{K} = 20$ on the reference discretization level and subsequently projected using the $L_2(\mathcal{D}_i)$ projection onto coarser discretization levels. This enables a comparison of the approximate GRFs on different levels of accuracy in the error plots. In Figure 3, we choose again $\mathcal{D}_1 = (0,1)^2$ and $\beta = 0.75$.

The convergence rate expected from the presented theory is $\approx 1/4$ with respect to the degrees of freedom N_ℓ of the FE space. We observe a pre-asymptotic regime, whose size seems to depend on the function κ^2, which determines the spatial correlation length of the GRF. More precisely, the particular form $\mathcal{A}_\kappa = \mathcal{A}_0 + \kappa^2$ with $\mathcal{A}_0 = -\Delta$ entails a (β-independent) pre-asymptotic regime in the onset of FE convergence which, for $\kappa \equiv$ const., corresponds to $h^{-1} \lesssim \kappa$. This is due to the lack of scale-resolution in the FE eigenfunction approximation in this regime. This effect is also visible in the present numerical experiment until the FE mesh is sufficiently fine to resolve the correlation length scale of the GRF $\mathcal{Z}^\beta$ in $\mathcal{D}_1$.

We furthermore investigate in Figure 4 the effect of a non-convex polygon on the asymptotic FE convergence rate. There, the convergence behavior for the square $\mathcal{D}_1 = (0,1)^2$ and for the polygon $\mathcal{D}_2 = (0,1)^2 \backslash \triangle$ are compared, where the triangle $\triangle$ has corners $(1,0)^\top$, $(1,1/10)^\top$, and $(1/2,1/2)^\top$ (the largest interior angle at $(1/2,1/2)^\top$ is $\theta \approx 1.96\pi$). Here, we set $\beta = 1.5$ and follow Section 2.2.5, where K_ℓ is chosen as above with $c = 1/2$. The reference solution is obtained as before with $\widetilde{K} = 20$, but with one further level of refinement with 523265 degrees of freedom using a sparse direct solver. The expected convergence rate is ≈ 1 for the convex polygon $\mathcal{D}_1$, see Table 1. For the non-convex polygon $\mathcal{D}_2$, we observe a reduced convergence rate with uniform refinement. This is explained by the fact that the inverse of the Dirichlet Laplacian is bounded from $L_2(\mathcal{D}_2)$ to $H^{1+t}(\mathcal{D}_2)$ for every $t \in (0, \min\{\pi/\theta, 1\}) \approx (0, 0.51)$. The empirical convergence rates in Figures 3 and 4 are computed by least squares taking into account the five data pairs corresponding to the five finest resolutions. The observed convergence rate for $\mathcal{D}_2$ is in accordance with the recently derived results on convergence of Whittle–Matérn fields if $\mathcal{A}$ is only $H^{1+t}(\mathcal{D}_2)$-regular for some $t \in (0,1)$, see [18, Def. 6.20 and Thm. 6.23]. The tests in Figure 4 were performed with $\kappa_1^2 = \kappa_2^2 = 10$.

6.3. *Simulation of non-isotropic GRFs on the sphere*

We consider the 2-sphere $\mathbb{S}^2 = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$ as a special case of a 2-dimensional closed, connected, orientable, smooth, compact surface. For *isotropic* GRFs on $\mathbb{S}^2$ (with isotropy taking the role of stationarity in the Euclidean case), fast simulation on equispaced partitions via FFT are possible. Here, we address a more general, non-isotropic setting where the function $\kappa^2(x)$ is given by $\kappa^2(x_1,x_2,x_3) = \kappa_1$ if $x_1 < 0$ and by $\kappa^2(x_1,x_2,x_3) = \kappa_2$ otherwise for some $\kappa_1, \kappa_2 > 0$. The resulting GRF $\mathcal{Z}^\beta$ is non-isotropic, i.e., the two point covariance is not given by a scalar function evaluated at the geodesic (here great circle) distance, since its covariance operator is not diagonalized by spherical harmonics, a necessary condition for isotropy of $\mathcal{Z}^\beta$, see [50, Thm. 5.13].

Finite Elements with nested spaces have been computationally realized on $\mathbb{S}^2$ for example in [37, Sec. 6]. We use again PCG with BPX and Jacobi smoothers. The stopping criterion for the iterative linear system solvers is as above the error tolerance 10^{-12}. The steering parameters of the sinc quadrature are likewise chosen

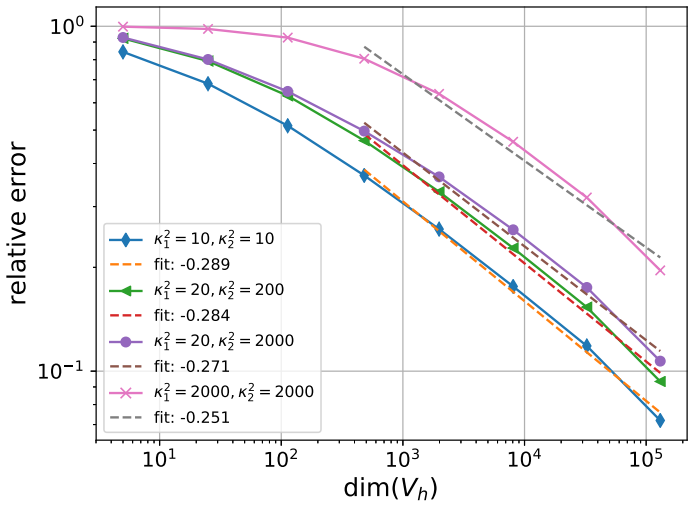

Figure 3. Relative $L_2(\Omega; L_2(\mathcal{D}_1))$-error of $\widetilde{\mathcal{Z}}^{\beta}_{K,h}$. $\mathcal{D}_1 = (0,1)^2$, $\beta = 0.75$ various values of κ_i^2, $i = 1, 2$. Onset of asymptotic convergence is delayed for large values of κ (corresponding to small spatial correlation length).

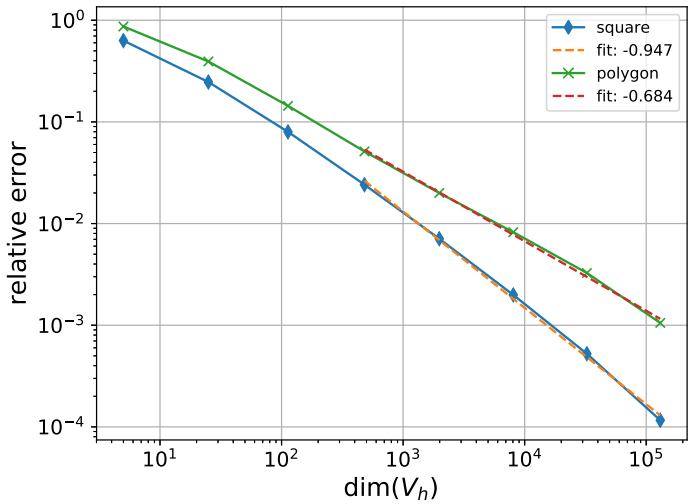

Figure 4. Relative $L_2(\Omega; L_2(\mathcal{D}_i))$-error, $i = 1, 2$, of $\widetilde{\mathcal{Z}}^{\beta}_{K,h}$. $\mathcal{D}_1 = (0,1)^2$ vs. $\mathcal{D}_2 = (0,1)^2 \backslash \triangle$, $\beta = 1.5$. Reduced convergence rate for uniform refinement due to low path regularity caused by corner singularities in polygonal domain.

as above; the domain is again 2-dimensional. In Figure 5 the CPU time is plotted as the average of 40 runs using $\kappa_1^2 = 20$ and $\kappa_2^2 = 200$. In Figure 6, the $L_2(\Omega; L_2(\mathbb{S}^2))$-error of the sinc-FE approximation for the non-isotropic GRF $\mathcal{Z}^\beta$ is shown, where the rate is again computed via a least squares fit. The $L_2(\Omega; L_2(\mathbb{S}^2))$-norm is approximated as the sample average of 60 Monte Carlo samples. The random forcing is as above approximated with $\widetilde{K} = 20$ on the discretization level of the reference solution and projected onto coarser levels with the $L_2(\mathbb{S}^2)$ projection. The reference solution is as above the numerical solution taken on two additional levels of mesh refinement with 4194306 degrees of freedom using a sparse direct solver.

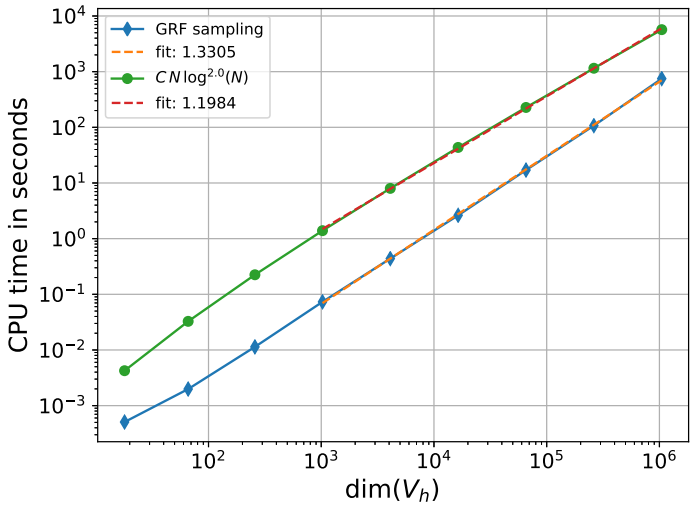

Figure 5. Simulation of non-isotropic GRFs on the sphere. CPU time per sample versus $N_h = \dim(V_h)$ for $\widetilde{\mathcal{Z}}^\beta_{K,h}$ with $\beta = 0.75$ on $\mathbb{S}^2$

The implementation for all numerical tests is based on the C++ library `BETL`, cf. Ref. 41. The CPU time in Figures 2 and 5 has been computed sequentially using the CPU `Intel(R) Xeon(R) CPU E5-2697 v2 @ 2.70GHz`.

7. Conclusions

We have developed the mathematical analysis and implementation of a class of Gaussian random fields $\mathcal{Z}^\beta$ on certain compact metric spaces $\mathcal{X}$, particularly on closures of bounded Euclidean subdomains $\mathcal{D} \subset \mathbb{R}^d$ and on compact 2-surfaces without boundary. The fields $\mathcal{Z}^\beta$ are obtained as solutions of the fractional-order equation $\mathcal{A}^\beta \mathcal{Z}^\beta = \mathcal{W}$, where $\mathcal{W}$ denotes white noise and where $\mathcal{A}$ denotes a linear, self-adjoint, second-order, elliptic reaction-diffusion operator on $\mathcal{X}$. The numerical simulation of the GRF $\mathcal{Z}^\beta$ has been realized by approximating $\mathcal{A}^{-\beta}$ via a sinc

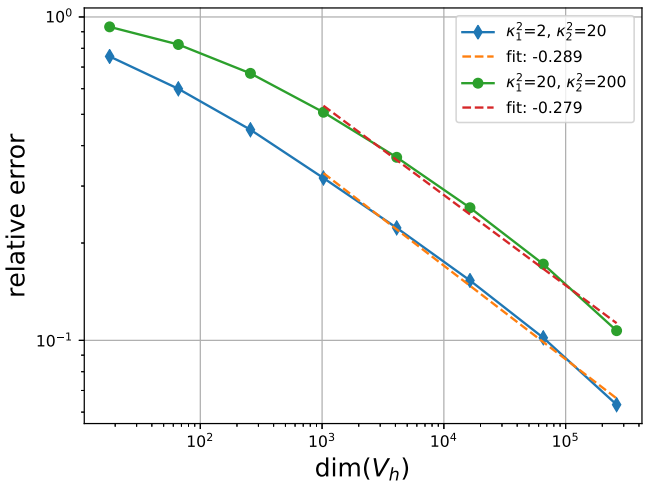

Figure 6. Simulation of non-isotropic GRFs on the sphere. Relative $L_2(\Omega; L_2(\mathbb{S}^2))$-error of $\widetilde{\mathcal{Z}}^\beta_{K,h}$ for $\beta = 0.75$, $\kappa_1^2 \in \{2, 20\}$, $\kappa_2^2 \in \{20, 200\}$

quadrature combined with a multilevel FE discretization of $\mathcal{A}$. Under the minimal Assumption 3.1 on the coefficients of $\mathcal{A}$, we have provided two computational strategies to numerically generate samples of $\mathcal{Z}^\beta$ on simplicial triangulations $\mathcal{T}$ of $\mathcal{X}$ in essentially linear complexity $\mathcal{O}(\#(\mathcal{T}))$ (up to log-terms) per sample. Here, $\mathcal{X}$ may be the closure of a bounded Euclidean domain or a manifold satisfying certain conditions. We note that these two strategies can furthermore be combined with approaches, where the fractional inverse $\mathcal{A}^{-\beta}$ is treated numerically via a rational (Padé) approximation, such as in Ref. 6, since they typically also require solving singularly perturbed elliptic problems.

In Section 5 we introduced MRA's on $\mathcal{X} \in \{\overline{\mathcal{D}}, \mathcal{M}\}$ for the purpose of multilevel preconditioning. Having at hand a FE wavelet basis $\mathbf{\Psi}$, which is stable in the spaces $\dot{H}^\sigma_{\mathcal{A}}$ in a range of Sobolev orders $\sigma \in (-\tilde{\gamma}, \gamma)$ for $\tilde{\gamma} > 0$ and $\gamma > 1$, allows for a *compressive numerical simulation* of the GRF $\mathcal{Z}^\beta$. The localization of supports of elements in such bases is also the key in linear scaling quasi-Monte Carlo algorithms.[30,38–40] For this reason, wavelet-based approximations of the equation $\mathcal{A}^\beta \mathcal{Z}^\beta = \mathcal{W}$ can, in principle, exploit higher $\mathbb{P}$-a.s. Sobolev regularity of the GRF $\mathcal{Z}$ that is afforded, for example, by a greater value of $\beta > 0$, see [18, Lem. 4.1] for the case of convex Euclidean domains. We remark that the assumption of global smoothness of the surface $\mathcal{M}$ which we imposed in Sections 4–5 can be relaxed to piecewise smooth surfaces which admit a global Lipschitz atlas, see Remark 4.3 and Ref. 22.

The presently developed multilevel approaches allow the fast simulation of GRFs also in a number applications which are currently of interest in computational UQ

and in data assimilation, e.g., when $\mathcal{X}$ is a high-dimensional data graph $\mathcal{G}$.[16,17,36] The need for fast (multilevel) simulation of GRFs arises, e.g., in (multilevel) Markov chain Monte Carlo methods in Bayesian UQ, see Ref. 5 and the references there. They also extend to *spatiotemporal GRFs* that are of interest in data assimilation and spatial statistics, see, e.g., Ref. 13.

Acknowledgment

The authors acknowledge David Bolin for drawing their attention to Ref. 35.

Bibliography

1. M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series* (For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964).
2. A. V. Balakrishnan, Fractional powers of closed operators and the semigroups generated by them, *Pacific J. Math.* **10** (1960) 419–437.
3. A. V. Balakrishnan, *Applied functional analysis*, volume 3 of *Applications of Mathematics* (Springer-Verlag, New York-Berlin, 1981), second edition.
4. M. M. Beaky, R. J. Scherrer and J. V. Villumsen, Topology of large-scale structure in seeded hot dark matter models, *Astrophys. J., Part 1* **387** (1992) 443–448.
5. A. L. Bertozzi, X. Luo, A. M. Stuart and K. C. Zygalakis, Uncertainty quantification in graph-based classification of high dimensional data, *SIAM/ASA J. Uncertain. Quantif.* **6** (2018) 568–595.
6. D. Bolin and K. Kirchner, The rational SPDE approach for Gaussian random fields with general smoothness, *J. Comput. Graph. Statist.* Electronic.
7. D. Bolin, K. Kirchner and M. Kovács, Numerical solution of fractional elliptic stochastic PDEs with spatial white noise, *IMA J. Numer. Anal.* Electronic.
8. D. Bolin, K. Kirchner and M. Kovács, Weak convergence of Galerkin approximations for fractional elliptic stochastic PDEs with spatial white noise, *BIT* **58** (2018) 881–906.
9. D. Bolin and F. Lindgren, Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping, *Ann. Appl. Stat.* **5** (2011) 523–550.
10. A. Bonito and J. E. Pasciak, Numerical approximation of fractional powers of elliptic operators, *Math. Comp.* **84** (2015) 2083–2110.
11. J. H. Bramble and J. E. Pasciak, New estimates for multilevel algorithms including the V-cycle, *Math. Comp.* **60** (1993) 447–471.
12. J. H. Bramble, J. E. Pasciak and O. Steinbach, On the stability of the L^2 projection in $H^1(\Omega)$, *Math. Comp.* **71** (2002) 147–156.
13. M. Cameletti, F. Lindgren, D. Simpson and H. Rue, Spatio-temporal modeling of particulate matter concentration through the SPDE approach, *AStA Adv. Stat. Anal.* **97** (2013) 109–131.
14. G. Chan and A. T. A. Wood, Algorithm AS 312: An algorithm for simulating stationary Gaussian random fields, *J. Roy. Statist. Soc. Ser. C* **46** (1997) 171–181.
15. J. Chen and M. L. Stein, Linear-cost covariance functions for Gaussian random fields, 2017, preprint, arXiv:1711.05895.
16. R. R. Coifman, I. G. Kevrekidis, S. Lafon, M. Maggioni and B. Nadler, Diffusion

maps, reduction coordinates, and low dimensional representation of stochastic systems, *Multiscale Model. Simul.* **7** (2008) 842–864.
17. R. R. Coifman and M. Maggioni, Diffusion wavelets, *Appl. Comput. Harmon. Anal.* **21** (2006) 53–94.
18. S. G. Cox and K. Kirchner, Regularity and convergence analysis in Sobolev and Hölder spaces for generalized Whittle–Matérn fields, Technical Report 2019-23, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2019.
19. P. E. Creasey and A. Lang, Fast generation of isotropic Gaussian random fields on the sphere, *Monte Carlo Methods Appl.* **24** (2018) 1–11.
20. M. Croci, M. B. Giles, M. E. Rognes and P. E. Farrell, Efficient white noise sampling and coupling for multilevel Monte Carlo with nonnested meshes, *SIAM/ASA J. Uncertain. Quantif.* **6** (2018) 1630–1655.
21. W. Dahmen, Multiscale and wavelet methods for operator equations, in *Multiscale problems and methods in numerical simulations* (Springer, Berlin, 2003), volume 1825 of *Lecture Notes in Math.*, pp. 31–96.
22. W. Dahmen, H. Harbrecht and R. Schneider, Compression techniques for boundary integral equations—asymptotically optimal complexity estimates, *SIAM J. Numer. Anal.* **43** (2006) 2251–2271.
23. E. B. Davies, *Spectral theory and differential operators*, volume 42 of *Cambridge Studies in Advanced Mathematics* (Cambridge University Press, Cambridge, 1995).
24. A. Demlow, Higher-order finite element methods and pointwise error estimates for elliptic problems on surfaces, *SIAM J. Numer. Anal.* **47** (2009) 805–827.
25. J. W. Demmel, S. C. Eisenstat, J. R. Gilbert, X. S. Li and J. W. H. Liu, A supernodal approach to sparse partial pivoting, *SIAM J. Matrix Anal. Appl.* **20** (1999) 720–755.
26. C. R. Dietrich and G. N. Newsam, Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix, *SIAM J. Sci. Comput.* **18** (1997) 1088–1107.
27. T. J. Dodwell, C. Ketelsen, R. Scheichl and A. L. Teckentrup, A hierarchical multilevel Markov chain Monte Carlo algorithm with applications to uncertainty quantification in subsurface flow, *SIAM/ASA J. Uncertain. Quantif.* **3** (2015) 1075–1108.
28. J. Dölz, H. Harbrecht and Ch. Schwab, Covariance regularity and $\mathcal{H}$-matrix approximation for rough random fields, *Numer. Math.* **135** (2017) 1045–1071.
29. M. Feischl, F. Y. Kuo and I. H. Sloan, Fast random field generation with H-matrices, *Numer. Math.* **140** (2018) 639–676.
30. R. N. Gantner, L. Herrmann and Ch. Schwab, Quasi-Monte Carlo integration for affine-parametric, elliptic PDEs: local supports and product weights, *SIAM J. Numer. Anal.* **56** (2018) 111–135.
31. D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Classics in Mathematics (Springer-Verlag, Berlin, 2001), reprint of the 1998 edition.
32. G. H. Golub and C. F. Van Loan, *Matrix computations*, Johns Hopkins Studies in the Mathematical Sciences (Johns Hopkins University Press, Baltimore, MD, 2013), fourth edition.
33. I. G. Graham, F. Y. Kuo, D. Nuyens, R. Scheichl and I. H. Sloan, Circulant embedding with QMC: analysis for elliptic PDE with lognormal coefficients, *Numer. Math.* **140** (2018) 479–511.
34. P. Grisvard, *Elliptic problems in nonsmooth domains*, volume 69 of *Classics in Applied Mathematics* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011).
35. N. Hale, N. J. Higham and L. N. Trefethen, Computing $\mathbf{A}^{\alpha}$, $\log(\mathbf{A})$, and related

matrix functions by contour integrals, *SIAM J. Numer. Anal.* **46** (2008) 2505–2523.
36. D. K. Hammond, P. Vandergheynst and R. Gribonval, Wavelets on graphs via spectral graph theory, *Appl. Comput. Harmon. Anal.* **30** (2011) 129–150.
37. L. Herrmann, A. Lang and Ch. Schwab, Numerical analysis of lognormal diffusions on the sphere, *Stoch. Partial Differ. Equ. Anal. Comput.* **6** (2018) 1–44.
38. L. Herrmann and Ch. Schwab, Multilevel quasi-Monte Carlo integration with product weights for elliptic PDEs with lognormal coefficients, *ESAIM Math. Model. Numer. Anal.* **53** (2019) 1507–1552.
39. L. Herrmann and Ch. Schwab, Multilevel quasi-Monte Carlo uncertainty quantification for advection-diffusion-reaction, Technical Report 2019-06, Seminar for Applied Mathematics, ETH Zürich, 2019, to appear in Monte Carlo and Quasi-Monte Carlo Methods – MCQMC, Rennes, France, July 2018.
40. L. Herrmann and Ch. Schwab, QMC integration for lognormal-parametric, elliptic PDEs: local supports and product weights, *Numer. Math.* **141** (2019) 63–102.
41. R. Hiptmair and L. Kielhorn, BETL — A generic boundary element template library, Technical Report 2012-36, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2012.
42. T. V. Kolev, J. Xu and Y. Zhu, Multilevel preconditioners for reaction-diffusion problems with discontinuous coefficients, *J. Sci. Comput.* **67** (2016) 324–350.
43. A. Lang and Ch. Schwab, Isotropic Gaussian random fields on the sphere: regularity, fast simulation and stochastic partial differential equations, *Ann. Appl. Probab.* **25** (2015) 3047–3094.
44. J. Latz, M. Eisenberger and E. Ullmann, Fast sampling of parameterised Gaussian random fields, *Comput. Methods Appl. Mech. Engrg.* **348** (2019) 978–1012.
45. M. Ledoux and M. Talagrand, *Probability in Banach spaces*, Classics in Mathematics (Springer-Verlag, Berlin, 2011), isoperimetry and processes, Reprint of the 1991 edition.
46. F. Lindgren, H. v. Rue and J. Lindström, An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach, *J. R. Stat. Soc. Ser. B Stat. Methodol.* **73** (2011) 423–498, with discussion.
47. J.-L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications. Vol. I* (Springer-Verlag, New York-Heidelberg, 1972), translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.
48. A. Lunardi, *Interpolation theory*, volume 16 of *Appunti. Scuola Normale Superiore di Pisa (Nuova Serie)* (Edizioni della Normale, Pisa, 2018).
49. J. Lund and K. L. Bowers, *Sinc methods for quadrature and differential equations* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992).
50. D. Marinucci and G. Peccati, *Random fields on the sphere*, volume 389 of *London Mathematical Society Lecture Note Series* (Cambridge University Press, Cambridge, 2011), representation, limit theorems and cosmological applications.
51. Y. Meyer, *Wavelets and operators*, volume 37 of *Cambridge Studies in Advanced Mathematics* (Cambridge University Press, Cambridge, 1992), translated from the 1990 French original by D. H. Salinger.
52. H. Nguyen and R. Stevenson, Finite element wavelets with improved quantitative properties, *J. Comput. Appl. Math.* **230** (2009) 706–727.
53. S. Osborn, P. S. Vassilevski and U. Villa, A multilevel, hierarchical sampling technique for spatially correlated random fields, *SIAM J. Sci. Comput.* **39** (2017) S543–S562.
54. S. Osborn, P. Zulian, T. Benson, U. Villa, R. Krause and P. S. Vassilevski, Scalable hierarchical PDE sampler for generating spatially correlated random fields using

nonmatching meshes, *Numer. Linear Algebra Appl.* **25** (2018) e2146, 23.
55. W. D. Penny, N. J. Trujillo-Barreto and K. J. Friston, Bayesian fMRI time series analysis with spatial priors, *NeuroImage* **24** (2005) 350–362.
56. N. Rekatsinas and R. Stevenson, A quadratic finite element wavelet Riesz basis, *Int. J. Wavelets Multiresolut. Inf. Process.* **16** (2018) 1850033, 17.
57. S. R. Sain, R. Furrer and N. Cressie, A spatial analysis of multivariate output from regional climate models, *Ann. Appl. Stat.* **5** (2011) 150–175.
58. R. Scheichl, A. M. Stuart and A. L. Teckentrup, Quasi-Monte Carlo and multilevel Monte Carlo methods for computing posterior expectations in elliptic inverse problems, *SIAM/ASA J. Uncertain. Quantif.* **5** (2017) 493–518.
59. O. Steinbach, *Numerical approximation methods for elliptic boundary value problems* (Springer, New York, 2008), finite and boundary elements, Translated from the 2003 German original.
60. F. Stenger, *Numerical methods based on sinc and analytic functions*, volume 20 of *Springer Series in Computational Mathematics* (Springer-Verlag, New York, 1993).
61. R. Stevenson, Adaptive wavelet methods for solving operator equations: an overview, in *Multiscale, nonlinear and adaptive approximation* (Springer, Berlin, 2009), pp. 543–597.
62. M. E. Taylor, *Pseudodifferential operators*, volume 34 of *Princeton Mathematical Series* (Princeton University Press, Princeton, N.J., 1981).
63. S. Torquato, *Random heterogeneous materials*, volume 16 of *Interdisciplinary Applied Mathematics* (Springer-Verlag, New York, 2002), microstructure and macroscopic properties.
64. P. Whittle, Stochastic processes in several dimensions, *Bull. Inst. Internat. Statist.* **40** (1963) 974–994.
65. J. Xu, Iterative methods by space decomposition and subspace correction, *SIAM Rev.* **34** (1992) 581–613.
66. J. Xu and L. Zikatanov, The method of alternating projections and the method of subspace corrections in Hilbert space, *J. Amer. Math. Soc.* **15** (2002) 573–597.