CrossMark

Representations of Gaussian Random Fields and Approximation of Elliptic PDEs with Lognormal Coefficients

Markus Bachmayr[1] · **Albert Cohen**[2] ·
Giovanni Migliorati[2]

Abstract Approximation of elliptic PDEs with random diffusion coefficients typically requires a representation of the diffusion field in terms of a sequence $y = (y_j)_{j \geq 1}$ of scalar random variables. One may then apply high-dimensional approximation methods to the solution map $y \mapsto u(y)$. Although Karhunen–Loève representations are commonly used, it was recently shown, in the relevant case of lognormal diffusion fields, that multilevel-type expansions may yield better approximation rates. Motivated by these results, we construct wavelet-type representations of stationary Gaussian random fields defined on arbitrary bounded domains. The size and localization properties of these wavelets are studied, and used to obtain polynomial approximation results for the related elliptic PDE which outperform those achievable when using Karhunen–Loève representations. Our construction is based on a periodic extension of the stationary random field, and the expansion on the domain is then obtained by simple restriction. This makes the approach easily applicable even when the computational domain of the PDE has a complicated geometry. In particular, we apply this construction to the class of Gaussian processes defined by the family of Matérn covari-

---

✉ Markus Bachmayr
  bachmayr@ins.uni-bonn.de

  Albert Cohen
  cohen@ljll.math.upmc.fr

  Giovanni Migliorati
  migliorati@ljll.math.upmc.fr

[1] Hausdorff Center for Mathematics, Institute for Numerical Simulation, Endenicher Allee 60, 53115 Bonn, Germany

[2] Laboratoire Jacques-Louis Lions, Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 7598, 4 place Jussieu, 75005 Paris, France

🐦 Birkhäuser

ances. The proposed periodic continuation technique has other relevant applications such as fast simulation of trajectories. It can be regarded as a continuous analog of circulant embedding techniques introduced for Toeplitz matrices. One of its specific features is that the rate of decay of the eigenvalues of the covariance operator of the periodized process provably matches that of the Fourier transform of the covariance function of the original process.

Keywords Gaussian random fields · Lognormal diffusion problems · Periodic continuation · Karhunen–Loève expansions · Wavelets

Mathematics Subject Classification 60G15 · 42C40 · 65T40 · 65T60

1 Introduction

Stochastic PDEs are commonly used to model uncertain physical phenomena. One such model that is used, for instance, in groundwater modeling is the diffusion equation

$$- \operatorname{div}(a \nabla u) = f \text{ in } D, \quad u|_{\partial D} = 0, \tag{1}$$

with lognormally distributed coefficient, that is, $a = \exp(b)$, where b is a centered Gaussian random field defined on the computational domain $D \subset \mathbb{R}^d$ (where typically $d = 2$ or $d = 3$).

Uncertainty quantification aims at describing the statistical properties of the resulting solution u, with various computational objectives: evaluating the mean field $\bar{u} = \mathbb{E}(u)$, estimating a plausible a based on some measurement data of the solution, describing the law of a scalar quantity of interest $Q(u)$.

For such objectives, it is most convenient to represent b in the form of an expansion

$$b = \sum_{j \geq 1} y_j \psi_j, \tag{2}$$

where y_j are i.i.d. $\mathcal{N}(0, 1)$, that is, independent scalar standard Gaussian random variables, and ψ_j are suitable functions on D. Once such an expansion is given, one may introduce approximations to the solution map

$$y \mapsto u(y), \quad y = (y_j)_{j \geq 1}, \tag{3}$$

for example by multivariate polynomials in the variables y_j. Such approximations provide a fast way to evaluate $u(y)$ for any choice of y up to some prescribed accuracy, which is of crucial help for the above mentioned tasks. Expansions of the form (2) can also be of practical use for the fast generation of trajectories, provided that the ψ_j have simple analytic expressions or are easy to compute numerically.

The centered Gaussian field b is characterized by its covariance function

$$(x, x') \mapsto K(x, x') = \mathbb{E}(b(x) b(x')), \quad x, x' \in D. \tag{4}$$

A standard way of obtaining a representation (2) is by using the Karhunen–Loève (KL) basis, that is, the $L^2(D)$-orthonormal eigenfunctions $(\varphi_j)_{j \geq 1}$ of the integral operator

$$T : v \mapsto Tv = \int_D K(\cdot, z)\, v(z)\, dz, \tag{5}$$

with corresponding eigenvalues $\lambda_j \geq 0$ arranged in decreasing order. One then obtains (2) by setting

$$\psi_j := \sqrt{\lambda_j} \varphi_j. \tag{6}$$

The distinguishing feature of this particular choice is that in addition to the statistical orthogonality $\mathbb{E}(y_j y_k) = \delta_{j,k}$, the functions ψ_j are orthogonal in $L^2(D)$. However, other expansions of the form (2) may also be considered if one does not impose the $L^2(D)$-orthogonality.

As shown in [23], general expansions of the form (2) with i.i.d. $\mathcal{N}(0, 1)$ coefficients are characterized by the fact that the ψ_j form a *tight frame* of the reproducing kernel Hilbert space (RKHS) $\mathcal{H}$ induced by K. Recall that $\mathcal{H}$ is the closure of the finite linear combinations of the functions $K_z := K(\cdot, z)$ for $z \in D$, with respect to the norm induced by the inner product $\langle K_x, K_z \rangle_{\mathcal{H}} := K(x, z)$, see [1] for a general treatment. Recall also that a tight frame of a Hilbert space $\mathcal{H}$ is a complete system that satisfies the identity

$$\sum_{j \geq 1} |\langle g, \psi_j \rangle_{\mathcal{H}}|^2 = \|g\|_{\mathcal{H}}^2, \quad g \in \mathcal{H}. \tag{7}$$

We refer to [11] for classical examples of time-frequency or time-scale frames. In contrast to orthonormal bases, such systems may be redundant. The possible redundancy in (2) can be illustrated by the following trivial example: if y_1 and y_2 are i.i.d. $\mathcal{N}(0, 1)$ and ψ is a given function, then $y_1 \psi + y_2 \psi = z(\sqrt{2}\psi)$ with $z = (y_1 + y_2)/\sqrt{2}$ also $\mathcal{N}(0, 1)$.

As an elementary yet useful example of the different possibilities for expanding a Gaussian process, consider the Brownian bridge on $D = [0, 1]$ whose covariance is given by $K(x, x') = \min\{x, x'\} - xx'$. On the one hand, the representation based on the KL expansion is given by the trigonometric functions

$$\psi_j(x) := \frac{\sqrt{2}}{\pi j} \sin(\pi j x), \quad j \geq 1. \tag{8}$$

On the other hand, another classical representation of this process is given by the Schauder basis, which consists of the hat functions

$$\psi_j(x) := 2^{-\ell/2} \sigma(2^\ell x - k), \quad \ell \geq 0, \quad k = 0, \ldots, 2^\ell - 1, \quad j = 2^\ell + k, \tag{9}$$

where

$$\sigma(x) := \max\{1 - |2x - 1|, 0\}. \tag{10}$$

Both systems are orthonormal bases (and thus tight frames) of the RKHS, which in this case is $\mathcal{H} = H_0^1(D)$ endowed with norm $\|v\|_{\mathcal{H}} := \|v'\|_{L^2(D)}$.

Lax–Milgram theory ensures that for each individual $y = (y_j)_{j \geq 1} \in U := \mathbb{R}^{\mathbb{N}}$ such that $\sum_{j \geq 1} y_j \psi_j$ converges in $L^\infty(D)$, the corresponding solution $u(y)$ is well defined in $V := H_0^1(D)$, with a priori bound

$$\|u(y)\|_{H_0^1} \leq C \exp\left(\left\|\sum_{j \geq 1} y_j \psi_j\right\|_{L^\infty}\right), \quad C := \|f\|_{V'}. \tag{11}$$

Sufficient conditions have been established, either in terms of the covariance function K or of the size properties of the ψ_j, which ensure that the solution map $y \mapsto u(y)$ belongs to $L^k(U, V, \gamma)$ for all $k < \infty$, see [3,6,10,17,21]. Here $L^k(U, V, \gamma)$ is the usual Bochner space where γ denotes the countable tensor product of the univariate standard Gaussian measure.

In turn, approximation can be obtained in $L^2(U, V, \gamma)$, corresponding to mean-square convergence, by truncation of the tensor product Hermite expansion

$$u(y) = \sum_{\nu \in \mathcal{F}} u_\nu H_\nu(y), \quad u_\nu = \int_U u(y) H_\nu(y) d\gamma(y) \in V, \quad H_\nu(y) := \prod_{j \geq 1} H_{\nu_j}(y_j). \tag{12}$$

Here we have denoted by $(H_n)_{n \geq 0}$ the sequence of univariate Hermite polynomials with normalization in $L^2(\mathbb{R}, g(t)dt)$ where $g(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$, and by $\mathcal{F}$ the set of finitely supported sequences $\nu = (\nu_j)_{j \geq 1}$ of non-negative integers.

The best error for a given number n of terms retained in the above expansion is attained by the so-called best n-term Hermite approximation $y \mapsto u_n(y)$ obtained by retaining the indices ν corresponding to the n largest $\|u_\nu\|_V$. By combining Parseval's identity with Stechkin's lemma [13], the resulting error can be quantified in terms of the ℓ^p-summability of the sequence $(\|u_\nu\|_V)_{\nu \in \mathcal{F}}$ for $p < 2$, namely

$$\|u - u_n\|_{L^2(U, V, \gamma)} \leq Cn^{-s}, \quad s := \frac{1}{p} - \frac{1}{2}, \quad C := \|(\|u_\nu\|_V)_{\nu \in \mathcal{F}}\|_{\ell^p(\mathcal{F})}. \tag{13}$$

As shown in [3], this extra summability may depend strongly on the particular representation (2) that is used, or in other words, on the choice of coordinates $(y_j)_{j \geq 1}$. Whereas previous results establishing summability properties of $(\|u_\nu\|_V)_{\nu \in \mathcal{F}}$ are based only on the summability of $(\|\psi_j\|_{L^\infty(D)})_{j \geq 1}$, that is, on the absolute sizes of the ψ_j, the results in [2,3] also take into account the localization properties of ψ_j. More specifically, the following is shown in [3], assuming that D is a bounded Lipschitz domain.

Theorem 1.1 Let $0 < p < 2$ and let $q = q(p) := \frac{2p}{2-p}$. Assume that there exists a positive sequence $(\rho_j)_{j \geq 1}$ such that

$$\sup_{x \in D} \sum_{j \geq 1} \rho_j |\psi_j(x)| < \infty \tag{14}$$

and

$$(\rho_j^{-1})_{j \geq 1} \in \ell^q(\mathbb{N}). \tag{15}$$

Then the solution map $y \mapsto u(y)$ *belongs to* $L^k(U, V, \gamma)$ *for all* $0 \leq k < \infty$. *Moreover,* $(\|u_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$. *In particular, best n-term Hermite approximations converge in* $L^2(U, V, \gamma)$ *with rate* n^{-s} *for* $s = \frac{1}{p} - \frac{1}{2} = \frac{1}{q}$.

Let us mention that although this theorem refers to the particular elliptic equation (1), inspection of its proof shows that a similar result holds for other types of equation. This includes in particular fourth-order elliptic equations or parabolic equations, again with lognormal coefficients (see the discussion in the end of Sect. 2 of [2] in the case of affine coefficients, which also applies to the lognormal case).

The above result draws an important distinction between representations using globally or locally supported functions ψ_j. If nothing is assumed on the supports of ψ_j, we may only apply the following immediate consequence of Theorem 1.1.

Corollary 1.2 *If* $(\psi_j)_{j \geq 1}$ *is a family of functions with arbitrary support such that* $(\|\psi_j\|_{L^\infty})_{j \geq 1}$ *belongs to* $\ell^r(\mathbb{N})$ *for some* $r < 1$, *then best n-term Hermite approximations converge in* $L^2(U, V, \gamma)$ *with rate* n^{-s} *for* $s = \frac{1}{r} - 1$.

Consider for example the above mentioned Brownian bridge. On the one hand, the KL functions given by (8) are globally supported with size of order j^{-1}. In turn, for any $q < \infty$, there exists no sequence $(\rho_j)_{j \geq 1}$ satisfying (14) and (15), and therefore the best *n*-term truncation cannot be proved to converge with any algebraic rate. On the other hand, the Schauder functions given by (9) have local support properties which allow us to fulfill (14) with $\rho_j = j^s$ for any $s < \frac{1}{2}$. Hence best *n*-term Hermite truncation based on the Schauder representation is ensured to converge with rate n^{-s} for any $s < \frac{1}{2}$.

The same analysis applies to a general process on a bounded domain $D \subset \mathbb{R}^d$ if it admits an expansion (2) where the ψ_j have wavelet-like localization properties. In this case, it is convenient to index the basis functions ψ_j by a scale-space index λ, with $|\lambda|$ denoting the corresponding scale level. We denote by $\mathcal{I}$ the set of these indices, where $\#\{\lambda \in \mathcal{I} : |\lambda| = \ell\} \sim 2^{d\ell}$ for $\ell \geq 0$.

Corollary 1.3 *Let* $(\psi_j)_{j \geq 1} = (\psi_\lambda)_{\lambda \in \mathcal{I}}$ *be a wavelet basis such that for some* $\alpha > 0$,

$$\sup_{x \in D} \sum_{|\lambda| = \ell} |\psi_\lambda(x)| \leq C2^{-\alpha \ell}, \quad \ell \geq 0. \tag{16}$$

If $(\|\psi_\lambda\|_{L^\infty})_{\lambda \in \mathcal{I}}$ *belongs to* $\ell^q(\mathcal{I})$, *which holds for* $q > \frac{d}{\alpha}$, *then the best n-term Hermite approximations converge in* $L^2(U, V, \gamma)$ *with rate* n^{-s} *for all* $s < \frac{1}{q} < \frac{\alpha}{d}$.

This leads to the question whether suitable wavelet-type systems $(\psi_j)_{j \geq 1}$ forming orthonormal bases or tight frames of $\mathcal{H}$ can also be found for more general b with given covariance kernel K. In view of the above corollary, we are interested in the decay exponent α that can be ensured in (16). For practical purposes, we also need these systems to either have exact analytic expressions or be computable by simple and efficient numerical procedures.

We shall focus on *stationary* random fields, that is, covariances of the form

$$K(x, x') = k(x - x'), \quad x, x' \in D, \tag{17}$$

where k is an even function defined over $\mathbb{R}^d$ which is the inverse Fourier transform of a positive measure. One typical class of examples is given by the family of *Matérn covariances*

$$k(x) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}|x|}{\lambda} \right)^{\nu} K_{\nu} \left(\frac{\sqrt{2\nu}|x|}{\lambda} \right), \tag{18}$$

where $\nu, \lambda > 0$ and K_ν is the modified Bessel function of the second kind, with Fourier transform given by

$$\hat{k}(\omega) = c_{\nu,\lambda} \left(\frac{2\nu}{\lambda^2} + |\omega|^2 \right)^{-(\nu+d/2)}, \quad c_{\nu,\lambda} := \frac{2^d \pi^{d/2} \Gamma(\nu + d/2)(2\nu)^{\nu}}{\Gamma(\nu)\lambda^{2\nu}}. \tag{19}$$

Here, for the Fourier transform, we use the convention

$$\hat{f}(\omega) = \int_{\mathbb{R}^d} f(x) \, e^{-ix\cdot\omega} \, dx. \tag{20}$$

The parameters ν and λ quantify the smoothness and correlation length, respectively, of the process.

The idea of using wavelet-type systems for the representation of Gaussian processes was put forward in the pioneering works of Ciesielski [7], motivated by the problem of analyzing the Hölder smoothness of univariate Gaussian processes. This program was pursued in [8] where Sobolev–Besov smoothness was investigated, in particular for the fractional Brownian motion. These papers were based on representing the processes of interest in the Schauder basis, with resulting components that are generally not independent and therefore not of the form (2).

A general approach was proposed in [4] to obtain wavelet-type representations of Gaussian processes defined on $\mathbb{R}^d$, with independent components. This approach requires that the covariance function is the Green's function of an operator of pseudo-differential type, and it includes Matérn covariances, see also [9]. A related approach was proposed in [15] for stationary Gaussian processes, motivated by fast methods for generating trajectories. Both approaches strongly rely on the Fourier transform over $\mathbb{R}^d$, and do not carry over in a simple manner to the case of a bounded domain $D \subset \mathbb{R}^d$.

A general construction of wavelet expansions of the form (2) for Gaussian processes was recently proposed in [22], in the general framework of Dirichlet spaces. Here the needed assumptions are that the covariance operator commutes with the operator that defines the Dirichlet structure. The latter does not have a simple explicit form in the case of Matérn covariances on a domain D, which makes this approach difficult to analyze and implement in our setting. Let us also mention the construction in [18], where an orthonormal basis for the RKHS is built by a direct Gram-Schmidt process, however, with generally no size and localization bounds on the resulting basis functions.

In the present paper, we propose an approach where the complications that may arise due to the geometry of D are circumvented by performing a periodic continuation of the random field b on a larger torus $\mathbb{T}$. The existence of such a continuation is discussed in Sect. 2. This leads to a simple construction of KL-type and wavelet-type expansions, by restrictions to D of similar expansions defined on $\mathbb{T}$. While the systems introduced

on $\mathbb{T}$ are bases, their restrictions to D are redundant frames of the RKHS, the amount of redundancy being essentially reflected by the ratio $|\mathbb{T}|/|D|$. Our construction thus achieves numerical simplicity at the price of redundancy. Here, D is a general bounded domain with no particular regularity assumption.

The periodic continuation has some parallels to *circulant embedding*, proposed independently in [14] and [28] as an algebraic technique for evaluating a stationary Gaussian random field, given by k, at the points of a uniform grid. Here, the (block) Toeplitz matrix formed by the grid values of k is embedded into a (block) circulant matrix, enabling a factorization by fast Fourier transform. Although it was shown in [12] that a positive definite Toeplitz matrix can always be embedded into a sufficiently large positive definite circulant matrix, its required size depends on the particular k under consideration, similarly to the size of $\mathbb{T}$ in our construction. Under more restrictive assumptions on k, simple procedures produce an embedding into a matrix of size proportional to the original one [14]. This strategy has also been applied in the numerical treatment of lognormal diffusion problems by quasi-Monte Carlo (QMC) methods [19].

Also in the setting of sampling-based methods such as QMC, our results have several potentially advantageous features. They yield a periodically extended random field on all of D, rather than on a uniform spatial grid. Any further approximation (e.g. on unstructured finite element meshes) can thus be adjusted independently of the periodic extension and of a chosen expansion of the random field. Furthermore, the decay properties of the KL eigenvalues of the periodic process are directly controlled by the decay of the Fourier transform of k.

In the case of KL-type expansions, which are studied in Sect. 3, the functions ψ_j that we obtain are simply the restrictions to D of trigonometric functions. This is an advantage in term of numerical simplicity compared to the $L^2(D)$-orthogonal KL functions of b which may not be easy to compute accurately, and in addition may not satisfy uniform L^∞ bounds. Wavelet-type expansions are defined in Sect. 4, and we establish their localization and size properties. In the case of Matérn covariances, they correspond to the value $\alpha = \nu$ in (16), where ν is the smoothness parameter in (19). Therefore, corresponding best n-term Hermite approximations converge with algebraic rate n^{-s} for any $0 < s < \nu/d$. Finally, in Sect. 5 details are given on numerical procedures which may be applied to define the periodic continuation and to construct the resulting wavelets, in the case of Matérn covariance, depending on the parameters λ, ν.

Remark 1.4 Our efforts in this paper are directed at improving the choice of the basis $(\psi_j)_{j \geq 1}$ in the representation of a Gaussian process, in the sense of achieving the property

$$\sup_{x \in D} \sum_{j \geq 1} \rho_j |\psi_j(x)| < \infty \tag{21}$$

with $(\rho_j^{-1}) \in \ell^q$ for the smallest possible value of q. While the wavelet bases that we construct are superior to KL bases in that particular sense, in the case of Matérn process, an open question is whether they are optimal among all possible representations. Closely related is the question of the choice of a representation that optimizes the

convergence of the truncated random series $b_n = \sum_{j=1}^n y_j \psi_j$ in the L^∞ norm towards the random field b. While KL bases are known to be optimal in the sense of minimizing the mean square error $E(\|b - b_n\|_{L^2}^2)$, they are generally suboptimal when replacing L^2 by L^∞ in the norm measuring the truncation error. The condition (21) can be used to estimate the convergence $\|b - b_n\|_{L^\infty}$ in various ways, for example in the almost sure sense. Indeed, we may write

$$\|b - b_n\|_{L^\infty} \le C \sup_{j>n} \rho_j^{-1} |y_j|, \quad C := \sup_{x \in D} \sum_{j \ge 1} \rho_j |\psi_j(x)|. \tag{22}$$

Then, observing that if ω_j is a slowly increasing sequence such that $\sum_{j \ge 1} \exp(-\omega_j^2) < \infty$ (for example, take $\omega_j = \sqrt{2 \ln j + 1}$), it is easily checked that $\sup_{j>n} \omega_j^{-1} |y_j|$ is finite with probability one. This allows us to derive that, almost surely,

$$\sup_{n \ge 1} \alpha_n^{-1} \|b - b_n\|_{L^\infty} < \infty, \quad \alpha_n := \sup_{j>n} \rho_j^{-1} \omega_j. \tag{23}$$

If $(\rho_j^{-1}) \in \ell^q$, and assuming without loss of generality that the ρ_j are in increasing order, then it follows that $\rho_j^{-1} \le C n^{-1/q}$ and therefore we obtain the almost sure convergence rate n^{-s} in L^∞ for any $0 < s < \frac{1}{q}$. In the case of Matérn covariances, using the constructed wavelet bases, we obtain such convergence rates for all values of $\nu > 0$, which does not seem to be achievable when using KL bases. It is again an open question to understand whether wavelet bases provide the best possible convergence rate.

Notation: Throughout the paper, as in the above introduction, we use C to denote a constant that may change between occurences (even in a chain of inequalities), and when necessary we indicate its value or the parameters on which it depends.

2 Periodic Continuation of a Stationary Process

Let $(b(x))_{x \in \mathbb{R}^d}$ be a real-valued, stationary and centered Gaussian process defined on $\mathbb{R}^d$, whose covariance is of the form

$$\mathbb{E}(b(x)b(x')) = k(x - x'), \tag{24}$$

where k is a real-valued and even function which is the inverse Fourier transform of a non-negative function $\hat{k}$. We work under the assumption that $\hat{k}$ is such that

$$0 \le \hat{k}(\omega) \le C(1 + |\omega|^2)^{-r}, \quad \omega \in \mathbb{R}^d, \tag{25}$$

for some $r > d/2$ and $C > 0$. Obviously, the Matérn covariances (18) satisfy this assumption with $r = \nu + d/2$ and C depending on (d, λ, ν). We consider the restricted process $(b(x))_{x \in D}$ defined on the bounded domain D of interest.

We aim for representations of the general form (2) where the y_j are i.i.d. $\mathcal{N}(0,1)$ and the $(\psi_j)_{j\geq 1}$ are a given sequence of functions defined on D. As explained in the introduction, one natural choice is $\psi_j = \sqrt{\lambda_j}\varphi_j$, where (φ_j, λ_j) are the eigenfunctions and eigenvalues of the covariance operator. However, this renormalized KL representation may not meet our requirements due to the possibly global support of the ψ_j and due to the slow decay of their L^∞ norms, while other representations could be more appropriate.

Our strategy for deriving better representations of the process over D is to view it as the restriction to D of a periodic stationary Gaussian process b_{p} defined on a suitable larger torus $\mathbb{T}$. As a consequence, any representation

$$b_{\mathrm{p}} = \sum_{j\geq 1} y_j \tilde{\psi}_j, \tag{26}$$

with y_j i.i.d. $\mathcal{N}(0,1)$ and $(\tilde{\psi}_j)_{j\geq 1}$ a given system of functions, yields a representation

$$b = \sum_{j\geq 1} y_j \psi_j, \quad \psi_j := \tilde{\psi}_j|_D. \tag{27}$$

The construction of b_{p} requires additional assumptions on the covariance function k.

Let $\delta := \mathrm{diam}(D)$, so that in a suitable coordinate system, D can be embedded into the box $[-\frac{\delta}{2}, \frac{\delta}{2}]^d$. We want to construct a periodic process $(b_{\mathrm{p}}(x))_{x\in\mathbb{T}}$ on a torus $\mathbb{T} = [-\gamma, \gamma]^d$ with $\gamma > \delta$, whose restriction $b_{\mathrm{p}}|_D$ on D is such that $b_{\mathrm{p}}|_D \sim b$, that is, $b_{\mathrm{p}}|_D$ and b share the same law. This is feasible provided that we can find an even and $\mathbb{T}$-periodic function k_{p} which agrees with k over $[-\delta, \delta]^d$ and such that the Fourier coefficients

$$c_n(k_{\mathrm{p}}) := \int_{\mathbb{T}} k_{\mathrm{p}}(z)\, e^{-i\frac{\pi}{\gamma}n\cdot z}\, dz, \quad n \in \mathbb{Z}^d, \tag{28}$$

are non-negative. In addition we would like that these Fourier coefficients have a similar rate of decay as the function $\hat{k}$, that is,

$$0 \leq c_n(k_{\mathrm{p}}) \leq C(1 + |n|^2)^{-r}, \quad n \in \mathbb{Z}^d, \tag{29}$$

for some $C > 0$. Note that k_{p} generally differs from the periodization $\sum_{n\in\mathbb{Z}^d} k(\cdot + 2\gamma n)$, which corresponds to a periodic process that does not agree with b on D.

One natural way of constructing the function k_{p} is by truncation and periodization: first we choose a sufficiently smooth and even cutoff function $\phi \colon \mathbb{R}^d \to \mathbb{R}$, to be specified further, such that $\phi|_{[-\delta,\delta]^d} = 1$ and $\phi(x) = 0$ for $x \notin [-\kappa, \kappa]^d$ where $\kappa := 2\gamma - \delta$, and define the truncation

$$k_{\mathrm{t}}(z) := k(z)\, \phi(z). \tag{30}$$

We now define k_{p} as the periodization of k_{t}, that is,

$$k_{\mathrm{p}}(z) = \sum_{n\in\mathbb{Z}^d} k_{\mathrm{t}}(z + 2\gamma n). \tag{31}$$

Obviously, k_p agrees with k over $[-\delta, \delta]^d$, and

$$c_n(k_p) = \hat{k}_t\left(\frac{\pi}{\gamma}n\right). \tag{32}$$

Therefore (29) follows if we can establish

$$0 \le \hat{k}_t(\omega) \le C(1 + |\omega|^2)^{-r}, \quad \omega \in \mathbb{R}^d, \tag{33}$$

for some $C > 0$. Since we have

$$\hat{k}_t = (2\pi)^{-d}\,\hat{k} * \hat{\phi}, \tag{34}$$

it is easily seen that the upper inequality in (33) follows from the upper inequality in (25), provided that ϕ is chosen sufficiently smooth such that

$$|\hat{\phi}(\omega)| \le C(1 + |\omega|^2)^{-r}, \tag{35}$$

for some $C > 0$. Indeed, combining (34) with (25) and (35), we obtain

$$
\begin{aligned}
(2\pi)^d |\hat{k}_t(\omega)| &\le \left|\int_{|\xi| \le |\omega|/2} \hat{k}(\xi)\,\hat{\phi}(\omega - \xi)\,d\xi\right| + \left|\int_{|\xi| \ge |\omega|/2} \hat{k}(\xi)\,\hat{\phi}(\omega - \xi)\,d\xi\right| \\
&\le \|\hat{k}\|_{L^1} \max_{|\xi| \ge |\omega|/2} |\hat{\phi}(\xi)| + \|\hat{\phi}\|_{L^1} \max_{|\xi| \ge |\omega|/2} |\hat{k}(\xi)| \\
&\le C(1 + |\omega|^2)^{-r}.
\end{aligned}
\tag{36}
$$

The main problem is to guarantee the lower inequality in (33), that is, the non-negativity of $\hat{k}_t$. Note that $\hat{\phi}$ cannot be non-negative: since $1 = \phi(0) = (2\pi)^{-d}\int_{\mathbb{R}^d} \hat{\phi}(\omega)d\omega$, the non-negativity of $\hat{\phi}$ would imply that

$$|\phi(x)| = (2\pi)^{-d}\left|\int_{\mathbb{R}^d} \hat{\phi}(\omega)\,e^{ix\cdot\omega}\,d\omega\right| < 1, \quad x \ne 0, \tag{37}$$

therefore contradicting the assumption $\phi|_{[-\delta,\delta]^d} = 1$.

It follows that for any such ϕ, the convolution operator

$$v \mapsto v * \hat{\phi}, \tag{38}$$

does not preserve positivity for all functions v. Here, we are only interested in preserving positivity for the particular function $\hat{k}$. However, the following result shows that this is in general not feasible only under the assumption (25).

Theorem 2.1 *For any $r > d/2$, there exists an even function k that satisfies (25) and such that for any ϕ satisfying $|\hat{\phi}(\omega)| \le C(1 + |\omega|^2)^{-s}$ for some $s > r$, $\phi|_{[-\delta,\delta]^d} = 1$, $\phi(x) = 0$ for $x \notin [-\kappa, \kappa]^d$ for some $\kappa > \delta$, the function $\hat{k}_t = (2\pi)^{-d}\,\hat{k} * \hat{\phi}$ is not non-negative.*

Proof Let h be a non-negative, smooth, even function on $\mathbb{R}^d$ with $h(0) = 1$ and support contained in the unit ball. For $\ell \in \mathbb{N}$, we choose arbitrary but fixed $\omega_\ell \in \mathbb{R}^d$ such that $|\omega_\ell| = 2^\ell$. We now define k by its Fourier transform as

$$\hat{k}(\omega) := \sum_{\ell \geq 1} 2^{-2r\ell}\Big(h\big(\ell(\omega - \omega_\ell)\big) + h\big(\ell(\omega + \omega_\ell)\big)\Big). \qquad (39)$$

Then clearly, (25) is satisfied. As demonstrated above, there exists $\omega^* \in \mathbb{R}^d$ such that $\hat{\phi}(\omega^*) < 0$. For $\ell > 1$, consider

$$\hat{k}_t(\omega^* + \omega_\ell) = (2\pi)^{-d}\int_{\mathbb{R}^d} \hat{\phi}(\omega^* - \xi)\,\hat{k}(\xi + \omega_\ell)\,d\xi = (2\pi)^{-d}\big(I_1(\ell) + I_2(\ell)\big), \quad (40)$$

where

$$I_1(\ell) := \int_{\mathbb{R}^d} \hat{\phi}(\omega^* - \xi)2^{-2r\ell}h(\ell\xi)\,d\xi = \ell^{-d}2^{-2r\ell}\int_{\mathbb{R}^d} \hat{\phi}(\omega^* - \ell^{-1}\xi)\,h(\xi)\,d\xi,$$

$$I_2(\ell) := \int_{\mathbb{R}^d} \hat{\phi}(\omega^* - \xi)\big(\hat{k}(\xi + \omega_\ell) - 2^{-2r\ell}h(\ell\xi)\big)\,d\xi.$$

On the one hand,

$$\lim_{\ell \to \infty} \ell^d 2^{2r\ell}I_1(\ell) = \hat{\phi}(\omega^*)\int_{\mathbb{R}^d} h(\xi)\,d\xi < 0. \qquad (41)$$

On the other hand, $\hat{k} \in L^1(\mathbb{R}^d)$ and $\hat{k}(\xi + \omega_\ell) - 2^{-2r\ell}h(\ell\xi)$ vanishes for $|\xi| \leq 2^{\ell-2}$. As a consequence,

$$|I_2(\ell)| \leq \|\hat{k}\|_{L^1(\mathbb{R}^d)} \max_{|\xi| \geq 2^{\ell-2}}|\hat{\phi}(\omega^* - \xi)|. \qquad (42)$$

For ℓ such that $2^{\ell-2} > 2|\omega^*|$, we thus have

$$|I_2(\ell)| \leq C\|\hat{k}\|_{L^1(\mathbb{R}^d)}(1 + 2^{2\ell-6})^{-s}. \qquad (43)$$

Therefore

$$\lim_{\ell \to \infty} \ell^d 2^{2r\ell}|I_2(\ell)| = 0. \qquad (44)$$

As a consequence, $\hat{k}_t(\omega^* + \omega_\ell) < 0$ for sufficiently large ℓ. □

The above counterexample reveals that further assumptions are needed on the covariance function k. Specifically, we work under the stronger assumptions

$$c(1 + |\omega|^2)^{-s} \leq \hat{k}(\omega) \leq C(1 + |\omega|^2)^{-r}, \qquad (45)$$

for some $s \geq r > d/2$ and $0 < c \leq C$, and

$$\lim_{R \to \infty} \int_{|x|>R} |\partial^\alpha k(x)| \, dx = 0, \quad |\alpha| \leq 2\lceil s \rceil, \tag{46}$$

where $\lceil s \rceil$ is the smallest integer greater or equal to s.

Remark 2.2 In the case of the Matérn covariance (18) with parameters $\nu, \lambda > 0$, the assumption (45) holds with $s = r = \nu + d/2$ as a consequence of (19). The assumption (46) actually holds for all derivation orders α, as a consequence of the exponential decay of the modified Bessel functions of the second kind K_ν and of their derivatives. This exponential decay can be seen, for example, from the integral representation

$$K_\nu(x) = \int_0^\infty e^{-x \cosh t} \cosh \nu t \, dt, \tag{47}$$

see [26, p. 181]. While in the present paper we focus on the example of Matérn covariances, there are of course other relevant processes that satisfy (45), in particular anisotropic processes. As an example consider a d-dimensional stationary process with anisotropic covariance given by

$$k(x) = k_{\nu_1, \lambda_1}(x_1) \ldots k_{\nu_d, \lambda_d}(x_d), \tag{48}$$

where $k_{\nu, \lambda}$ stands for the one-dimensional Matérn covariance as given in (18). Taking the Fourier transform, it is easily seen that (45) again holds, however with values $s > r$ in contrast to the isotropic case.

Theorem 2.3 *Let k be an even function such that (45) and (46) hold. Then for $\kappa > \delta$ sufficiently large, there exists ϕ satisfying (35), $\phi|_{[-\delta, \delta]^d} = 1$, and $\phi(x) = 0$ for $x \notin [-\kappa, \kappa]^d$ such that $\hat{k}_t = (2\pi)^{-d} \hat{k} * \hat{\phi}$ is positive.*

Proof We first choose a function $\phi_{2\delta} \in C^{2p}(\mathbb{R}^d)$ with $p := \lceil s \rceil$, supported on $[-2\delta, 2\delta]^d$ and such that $\phi_{2\delta}|_{[-\delta, \delta]^d} = 1$. Then for each $\kappa \geq 2\delta$, we define

$$\phi_\kappa(x) = \phi_{2\delta}(2\delta x/\kappa). \tag{49}$$

We thus have, for all $\kappa \geq 2\delta$,

$$\max_{|\alpha| \leq 2p} \|\partial^\alpha \phi_\kappa\|_{L^\infty} \leq \max_{|\alpha| \leq 2p} \|\partial^\alpha \phi_{2\delta}\|_{L^\infty} =: M < \infty. \tag{50}$$

Note that ϕ_κ satisfies (35) with a constant C that depends on κ. For a value of κ to be fixed further, we take $\phi := \phi_\kappa$, and let $\theta := 1 - \phi$. Then

$$\hat{k}_t(\omega) = \hat{k}(\omega) - \widehat{k\theta}(\omega) \geq c(1 + |\omega|^2)^{-s} - \widehat{k\theta}(\omega), \tag{51}$$

and

$$\left|\widehat{k\theta}(\omega)\right| \le (1+|\omega|^2)^{-p} \int_{\mathbb{R}^d} |(I-\Delta)^p(k\theta)|\, dx \le (1+|\omega|^2)^{-s} \int_{\mathbb{R}^d} |(I-\Delta)^p(k\theta)|\, dx. \tag{52}$$

By repeated application of Leibniz' rule and separately bounding each term, one finds

$$\int_{\mathbb{R}^d} |(I-\Delta)^p(k\theta)|\, dx \le C(\kappa) := (1+4d)^p M \max_{|\alpha|\le 2p} \int_{|x|>\kappa/2} |\partial^\alpha k(x)|\, dx. \tag{53}$$

Since $\lim_{\kappa\to\infty} C(\kappa) = 0$ by (46), it follows that $\hat{k}_t$ is positive for κ chosen large enough such that $C(\kappa) \le c$. □

Remark 2.4 From the proof of Theorem 2.3, for a given k and a choice of $\phi_{2\delta}$ one may also extract an upper bound for the required size of κ (and hence γ) from the decay of the right side of (53); note that here, p and M also depend on k via s. For the case of Matérn covariances, the resulting requirements on γ are illustrated numerically in Fig. 1 of Sect. 5.1.

3 Karhunen–Loève Representations

Let us recall that the standard Karhunen–Loève (KL) decomposition of the stationary Gaussian process b has the form (2) with

$$\psi_j = \psi_j^{\mathrm{KL}} := \sqrt{\lambda_j}\, \varphi_j, \tag{54}$$

where $(\lambda_j)_{j\ge 1}$ is the sequence of positive eigenvalues of the covariance operator

$$T : v \mapsto Tv = \int_D k(\cdot - x)\, v(x)\, dx, \tag{55}$$

arranged in decreasing order, and $(\varphi_j)_{j\ge 1}$ is the associated $L^2(D)$-orthonormal basis of eigenfunctions.

Working under assumptions (45) and (46), the periodized construction described in the previous section provides us with an alternative decomposition based on the covariance operator associated to the periodized process b_{p}, that is,

$$T_{\mathrm{p}} : v \mapsto T_{\mathrm{p}}v = \int_{\mathbb{T}} k_{\mathrm{p}}(\cdot - x)\, v(x)\, dx, \tag{56}$$

with $\mathbb{T} = [-\gamma, \gamma]^d$. The $L^2(\mathbb{T})$-orthonormal eigenfunctions of this operator are explicitly given by the trigonometric functions

$$\theta_n(z) := t_{n_1}(z_1)\cdots t_{n_d}(z_d), \quad n = (n_1, \ldots, n_d) \in \mathbb{N}_0^d, \tag{57}$$

where $t_0(z) = (2\gamma)^{-1/2}$ and

$$t_{2m}(z) = \gamma^{-1/2} \cos\left(\frac{m\pi z}{\gamma}\right) \quad \text{and} \quad t_{2m-1}(z) = \gamma^{-1/2} \sin\left(\frac{m\pi z}{\gamma}\right), \quad m \geq 1. \tag{58}$$

The eigenvalues are related to the Fourier coefficients $c_n(k_p)$ defined in (28). For the above eigenfunction θ_n, with $n = (n_1, \ldots, n_d)$ and each n_i being either of the form $2m_i$ or $2m_i - 1$, the corresponding eigenvalue is

$$c_m(k_p), \quad m = (m_1, \ldots, m_d). \tag{59}$$

We denote by $(\lambda_{p,j})_{j\geq 1}$ a decreasing rearrangement of these eigenvalues, with corresponding eigenfunctions $(\varphi_{p,j})_{j\geq 1}$. We may thus write

$$b_p = \sum_{j \geq 1} y_j \psi_{p,j}, \tag{60}$$

where the y_j are i.i.d. $\mathcal{N}(0, 1)$ and

$$\psi_{p,j} := \sqrt{\lambda_{p,j}}\, \varphi_{p,j}. \tag{61}$$

Since $b_p \sim b$ on D, this yields a decomposition of b by restriction, that is, taking

$$\psi_j = \psi_j^R := \sqrt{\lambda_{p,j}}\, \varphi_{p,j}|_D. \tag{62}$$

From (29), we obtain that, for some $C > 0$,

$$\#\{n \in \mathbb{Z}^d \ : \ |c_n(k_p)| \geq \eta\} \leq C\eta^{-2r/d}, \quad \eta > 0. \tag{63}$$

Thus, by decreasing rearrangement, we obtain a decay estimate of the form

$$\lambda_{p,j} \leq Cj^{-\frac{2r}{d}}, \quad j \geq 1, \tag{64}$$

for some $C > 0$. One first observation is that a similar decay estimate holds for the original eigenvalues λ_j.

Theorem 3.1 *If the covariance function satisfies (45) and (46), then we have*

$$\lambda_j \leq Cj^{-\frac{2r}{d}}, \quad j \geq 1, \tag{65}$$

for some $C > 0$.

Proof We denote by

$$V_n = \mathrm{span}\{\varphi_1, \ldots, \varphi_n\}, \tag{66}$$

the spaces generated by the Karhunen–Loève functions. These spaces satisfy the optimality property

$$\sum_{j>n} \lambda_j = \mathbb{E}\big(\|b - P_{V_n} b\|^2_{L^2(D)}\big) = \min_{\dim(V)=n} \mathbb{E}\big(\|b - P_V b\|^2_{L^2(D)}\big), \qquad (67)$$

where P_V is the $L^2(D)$-orthogonal projector. With

$$V_{\mathrm{p},n} := \mathrm{span}\{\varphi_{\mathrm{p},1}, \ldots, \varphi_{\mathrm{p},n}\}, \qquad (68)$$

the spaces generated by the KL functions of b_p, we denote by

$$W_n := \mathrm{span}\{\varphi_{\mathrm{p},1}|_D, \ldots, \varphi_{\mathrm{p},n}|_D\} \qquad (69)$$

their restriction to D. We thus have

$$\sum_{j>n} \lambda_j \leq \mathbb{E}\big(\|b - P_{W_n} b\|^2_{L^2(D)}\big), \qquad (70)$$

and since b and b_p agree on D, it follows that

$$\sum_{j>n} \lambda_j \leq \mathbb{E}\big(\|b_\mathrm{p} - P_{V_{\mathrm{p},n}} b_\mathrm{p}\|^2_{L^2(\mathbb{T})}\big), \qquad (71)$$

where $P_{V_{\mathrm{p},n}}$ is the $L^2(\mathbb{T})$-orthogonal projector. Therefore

$$\sum_{j>n} \lambda_j \leq \sum_{j>n} \lambda_{\mathrm{p},j} \leq Cn^{1-\frac{2r}{d}}, \quad n \geq 1, \qquad (72)$$

where the second inequality follows from (64). Since the λ_j are positive non-increasing, we obtain

$$\lambda_j \leq \frac{2}{j} \sum_{\lfloor \frac{j}{2} \rfloor < l \leq j} \lambda_l \leq \frac{2}{j} \sum_{l > \lfloor \frac{j}{2} \rfloor} \lambda_l \leq Cj^{-\frac{2r}{d}}, \qquad (73)$$

which is (65).   □

Remark 3.2 In the case of the Matérn covariance, in view of (19), we therefore obtain (65) with the value $r := \nu + d/2$. This estimate was derived in [20] by a different approach, using the theory developed by Widom for the eigenvalues of convolution-type operators. This theory makes the assumption that $\hat{k}$ is unimodal in each variable, see [27, p. 290], which holds for Matérn covariances, but is not needed in the above construction based on assumptions (45) and (46).

One interest of using the representation based on the functions ψ_j^R defined by restriction according to (62) is that the functions $\varphi_{p,j}$ are explicitly given by tensorized trigonometric functions. In particular, they are uniformly bounded. It follows that

$$\|\psi_j^R\|_{L^\infty} \le C\lambda_{p,j}^{1/2}, \quad j \ge 1, \tag{74}$$

and therefore, by (64),

$$\|\psi_j^R\|_{L^\infty} \le Cj^{-\frac{r}{d}}, \quad j \ge 1. \tag{75}$$

In contrast, such uniform bounds for the $L^\infty(D)$ norms are generally not available for the KL functions φ_j, which are in addition not easily computable in the case of a general multivariate domain.

In the particular case of Matérn covariances, the L^∞ norms of these functions have been estimated in [20], by arguments which use their natural connections with Hilbertian Sobolev spaces. We give below a similar argument which improves on the estimates established in [20]. Fixing an s such that $\frac{d}{2} < s < r = \nu + d/2$, and assuming that the domain D satisfies the uniform cone condition, we may use Sobolev embedding to obtain

$$\|\varphi_j\|_{L^\infty(D)} \le C\|\varphi_j\|_{H^s(D)}. \tag{76}$$

We then find by interpolation inequalities between Sobolev spaces that

$$\|\varphi_j\|_{L^\infty(D)} \le C\|\varphi_j\|_{L^2(D)}^{1-s/r} \|\varphi_j\|_{H^r(D)}^{s/r} = C\|\varphi_j\|_{H^r(D)}^{s/r}. \tag{77}$$

In order to estimate the $H^r(D)$ norms of the functions φ_j, we use the following bound for the covariance operator T: for any $v \in L^2(D)$, denoting by w its extension by zero to $\mathbb{R}^d$, we have

$$\begin{aligned}
\|Tv\|_{H^r(D)}^2 = \|(k * w)|_D\|_{H^r(D)}^2 &\le \|k * w\|_{H^r(\mathbb{R}^d)}^2 \\
&= C\int_{\mathbb{R}^d} (1 + |\omega|^2)^r |\hat{k}(\omega)\hat{w}(\omega)|^2 d\omega \\
&\le C\int_{\mathbb{R}^d} (1 + |\omega|^2)^{-r} |\hat{w}(\omega)|^2 d\omega \\
&\le C\int_{\mathbb{R}^d} \hat{k}(\omega)|\hat{w}(\omega)|^2 d\omega \\
&= C\langle k * w, w\rangle_{L^2(\mathbb{R}^d)} \le C\|Tv\|_{L^2(D)}\|v\|_{L^2(D)},
\end{aligned}$$

where we have used the characterization of Hilbertian Sobolev spaces by Fourier transforms and the particular form of the Matérn covariance (recall that by convention the constant C is allowed to change value in the above chain of inequalities). Taking $v = \varphi_j$ and using $T\varphi_j = \lambda_j \varphi_j$, we thus obtain

$$\|\varphi_j\|_{H^r(D)} \le C\lambda_j^{-1/2}. \tag{78}$$

In summary we have obtained the non-uniform bound

$$\|\varphi_j\|_{L^\infty(D)} \le C\lambda_j^{-\frac{s}{2r}} \tag{79}$$

and therefore, by (65),

$$\|\psi_j^{\mathrm{KL}}\|_{L^\infty(D)} \le Cj^{-\frac{r-s}{d}}. \tag{80}$$

In particular, we may take $s = \frac{d}{2} + \varepsilon$ for any sufficiently small $\varepsilon > 0$ to obtain

$$\|\psi_j^{\mathrm{KL}}\|_{L^\infty(D)} \le Cj^{-\frac{r}{d}+\frac{1}{2}+\varepsilon}. \tag{81}$$

This needs to be compared with (64), which in the present case of the Matérn covariance yields

$$\|\psi_j^{\mathrm{R}}\|_{L^\infty(D)} \le Cj^{-\frac{r}{d}}, \tag{82}$$

since $\|\varphi_{\mathrm{p},j}\|_{L^\infty(D)} \le 1$.

Let us mention that in the particular univariate case $d = 1$, numerical experiment seem to indicate that $\|\varphi_j\|_{L^\infty(D)}$ stays bounded independently of j, and therefore that the upper bound (79) is not always sharp. On the other hand, one can also exhibit examples of stationary processes such that the corresponding KL functions on the domain D are not uniformly bounded. Take for example the case $D = [-1, 1]$ and k such that $\hat{k} = \chi_{[-F,F]}$ for which the KL functions φ_j coincide with the univariate prolate spheroidal functions introduced in [25]. It is known that these functions are uniformly close to the Legendre polynomials L_j as $j \to \infty$, which shows that $\|\varphi_j\|_{L^\infty(D)} \sim j^{1/2}$, see [5].

In summary, we have obtained substantially better bounds on the decay of $\|\psi_j^{\mathrm{R}}\|_{L^\infty(D)}$ than available for $\|\psi_j^{\mathrm{KL}}\|_{L^\infty(D)}$, and in addition the ψ_j^{R} are easily computed numerically while this is generally not the case for the ψ_j^{KL}. However, (82) still leads to rather severe restrictions on the values of r for which Theorem 1.1 is applicable via Corollary 1.2, due to the global supports of the functions ψ_j^{R}. In the following section we consider an alternative wavelet-type construction for which Theorem 1.1, with Corollary 1.3, yields an approximation rate for corresponding solutions of u for any $r > \frac{d}{2}$.

4 Wavelet Representations

Our starting point is an $L^2(\mathbb{R})$-orthonormal wavelet basis, that is a basis of the form

$$\{\varphi(\cdot - n)\}_{n\in\mathbb{Z}} \cup \{2^{\ell/2}\psi(2^\ell \cdot -n)\}_{\ell\ge 0, n\in\mathbb{Z}} \tag{83}$$

where φ and ψ are the scaling function and mother wavelet, respectively. For simplicity we use the Meyer wavelets, whose construction is detailed in [11,24], and for which

$$\mathrm{supp}(\hat{\varphi}) = \left[-\frac{4\pi}{3}, \frac{4\pi}{3}\right] \quad \text{and} \quad \mathrm{supp}(\hat{\psi}) = \left[-\frac{8\pi}{3}, -\frac{2\pi}{3}\right] \cup \left[\frac{2\pi}{3}, \frac{8\pi}{3}\right]. \tag{84}$$

The functions $\hat{\varphi}$ and $\hat{\psi}$ may be chosen to be smooth, but for our purpose it will be enough to assume that they have M uniformly bounded derivatives with an integer $M \geq d + 1$.

Denoting $\psi_0 := \varphi$ and $\psi_1 := \psi$, the multivariate scaling function and wavelets are defined by

$$\Phi(x) := \varphi(x_1) \cdots \varphi(x_d), \qquad \Psi_\varepsilon(x) := \psi_{\varepsilon_1}(x_1) \cdots \psi_{\varepsilon_d}(x_d), \quad \varepsilon \in \mathcal{C}, \tag{85}$$

where $\mathcal{C} := \{0, 1\}^d \setminus \{(0, \ldots, 0)\}$. Then

$$\{\Phi(\cdot - n) : n \in \mathbb{Z}^d\} \cup \{\Psi_{\varepsilon, n, \ell} : n \in \mathbb{Z}^d, \ell \geq 0, \varepsilon \in \mathcal{C}\}, \tag{86}$$

is an orthonormal basis of $L^2(\mathbb{R}^d)$, where we have used the notation

$$\Psi_{\varepsilon, n, \ell} := 2^{d\ell/2} \Psi_\varepsilon(2^\ell \cdot - n). \tag{87}$$

We obtain an orthonormal basis of $L^2(\mathbb{T})$ by rescaling and periodization. This basis consists of the constant scaling function

$$\Phi^{\mathrm{p}}(x) := \sum_{m \in \mathbb{Z}^d} (2\gamma)^{-d/2} \Phi\big((2\gamma)^{-1} x - m\big) = (2\gamma)^{-d/2}, \tag{88}$$

and the $\mathbb{T}$-periodic wavelets

$$\Psi^{\mathrm{p}}_{\varepsilon, \ell, n}(x) := \sum_{m \in \mathbb{Z}^d} (2\gamma)^{-d/2} \Psi_{\varepsilon, n, \ell}\big((2\gamma)^{-1} x - m\big), \tag{89}$$

for $n \in \{0, \ldots, 2^\ell - 1\}^d$, $\ell \geq 0$, $\varepsilon \in \mathcal{C}$. From the Poisson summation formula, and the support properties of $\hat{\varphi}$ and $\hat{\psi}$, it is easily seen that the above wavelets, at a given scale level ℓ, are finite linear combinations of the Fourier exponentials e_n with $\|n\|_\infty \leq 2^{\ell+2}$. In other words, they are trigonometric polynomials of degree at most $2^{\ell+2}$ in each variable.

We now make the following general observation: let $(g_j)_{j \geq 1}$ be any orthonormal basis of $L^2(\mathbb{T})$, with each basis function having the Fourier expansion

$$g_j = (2\gamma)^{-d/2} \sum_{n \in \mathbb{Z}^d} c_n(g_j) e_n, \tag{90}$$

where

$$e_n(z) := (2\gamma)^{-d/2} e^{i \frac{\pi}{\gamma} n \cdot z}, \quad n \in \mathbb{Z}^d. \tag{91}$$

Then, defining the functions $(\bar{g}_j)_{j \geq 1}$ by $\bar{g}_j := S g_j$, where S is the filtering operator

$$v \mapsto Sv := (2\gamma)^{-d/2} \sum_{n \in \mathbb{Z}^d} \sqrt{c_n(k_{\mathrm{p}})} \, c_n(v) e_n, \tag{92}$$

we obtain a decomposition

$$b_{\mathrm{p}} = \sum_{j \geq 1} y_j \bar{g}_j, \tag{93}$$

where the y_j are i.i.d. $\mathcal{N}(0, 1)$, and therefore $b = \sum_{j \geq 1} y_j \bar{g}_j|_D$.

We apply this procedure to the above described periodic wavelet basis, therefore obtaining new periodic functions

$$\bar{\Phi}^{\mathrm{p}} = S\Phi^{\mathrm{p}} \quad \text{and} \quad \bar{\Psi}^{\mathrm{p}}_{\varepsilon,\ell,n} := S\Psi^{\mathrm{p}}_{\varepsilon,\ell,n}, \tag{94}$$

which are adapted to the decomposition of b_{p}. By construction $\bar{\Phi}^{\mathrm{p}}$ is again a constant function with value $(2\gamma)^{-d/2}\sqrt{c_0(k_{\mathrm{p}})}$, and the functions $\bar{\Psi}^{\mathrm{p}}_{\varepsilon,\ell,n}$ are trigonometric polynomials of degree at most $2^{\ell+2}$ in each variable. We next study in more detail the size and localization properties of these functions and show that they essentially behave like a wavelet basis with normalization $2^{\ell(d/2-r)}$ in place of $2^{\ell d/2}$. This geometric decay of the L^∞ norms, combined with the localization properties, will allow us to apply Theorem 1.1 and Corollary 1.3 for any value of $r > \frac{d}{2}$.

Note that since k_{p} has been obtained by periodizing the truncated function $k_{\mathrm{t}} = k\phi$, an equivalent construction of the functions $\bar{\Psi}^{\mathrm{p}}_{\varepsilon,\ell,n}$ is obtained by first defining over $\mathbb{R}^d$ the rescaled and filtered wavelets $\bar{\Psi}_{\varepsilon,\ell,n}$ according to

$$\widehat{\bar{\Psi}}_{\varepsilon,\ell,n}(\omega) = \hat{k}_{\mathrm{t}}^{\frac{1}{2}}(\omega)\,(2\gamma)^{d/2}\widehat{\Psi}_{\varepsilon,\ell,n}(2\gamma\omega), \tag{95}$$

and then applying $\mathbb{T}$-periodization, that is,

$$\bar{\Psi}^{\mathrm{p}}_{\varepsilon,\ell,n}(x) = \sum_{m \in \mathbb{Z}^d} \bar{\Psi}_{\varepsilon,\ell,n}(x + 2\gamma m). \tag{96}$$

We thus focus our attention on the size and localization properties of the functions $\bar{\Psi}_{\varepsilon,\ell,n}$.

Note that these functions inherit from the wavelet basis the translation invariance structure, since, at any given scale level ℓ, the functions $\bar{\Psi}_{\varepsilon,\ell,n}$ are translates by $2\gamma 2^{-\ell}n$ of $\bar{\Psi}_{\varepsilon,\ell,0}$, that is,

$$\bar{\Psi}_{\varepsilon,\ell,n} = \bar{\Psi}_{\varepsilon,\ell,0}(\cdot - 2\gamma 2^{-\ell}n). \tag{97}$$

However, they do not inherit the dilation invariance structure, since the functions $\bar{\Psi}_{\varepsilon,\ell,0}$ are not obtained by a simple rescaling of $\bar{\Psi}_{\varepsilon,0,0}$. We introduce rescaled functions $F_{\varepsilon,\ell}$ such that

$$\bar{\Psi}_{\varepsilon,\ell,0}(x) = 2^{\ell(d/2-r)}F_{\varepsilon,\ell}(2^\ell x), \tag{98}$$

that is, $F_{\varepsilon,\ell}$ is defined by

$$F_{\varepsilon,\ell}(x) = 2^{-\ell(d/2-r)}\bar{\Psi}_{\varepsilon,\ell,0}(2^{-\ell}x). \tag{99}$$

Our objective is now to show that the functions $F_{\varepsilon,\ell}$ satisfy a uniform localization estimate, independently of ε and ℓ.

For this purpose, we require some additional assumptions on the covariance function, namely that (45) holds with $s = r$, and that the partial derivatives of $\hat{k}$ satisfy improved decay estimates

$$|\partial^\alpha \hat{k}(\omega)| \le C(1 + |\omega|^2)^{-(r+|\alpha|/2)}, \quad |\alpha| \le M, \tag{100}$$

for an $M \ge d + 1$. It is easily seen that Matérn covariances satisfy such estimates, by straightforward differentiation of (19).

Lemma 4.1 *Let k satisfy (45), (46) with $s = r$ and (100) with an integer $M \ge d + 1$. Then ϕ in Theorem 2.3 can be chosen such that*

$$|\partial^\alpha \hat{k}_t^{\frac{1}{2}}(\omega)| \le C(1 + |\omega|)^{-(r+|\alpha|)}, \quad 0 \le |\alpha| \le M. \tag{101}$$

Proof We proceed exactly as in the proof of Theorem 2.3 to obtain a family of functions $\phi_\kappa \in C^{2p}(\mathbb{R}^d)$ for $\kappa \ge 2\delta$, supported on $[-2\delta, 2\delta]^d$ and such that $\phi_{2\delta}|_{[-\delta,\delta]^d} = 1$, and satisfying

$$\max_{|\alpha| \le 2p} \|\partial^\alpha \phi_\kappa\|_{L^\infty} \le D < \infty \tag{102}$$

for some $D > 0$, but here with $p := \lceil r + M/2 \rceil$.

As in the proof of Theorem 2.3, we obtain that for $\phi := \phi_\kappa$ with κ sufficiently large, there exist $\tilde{c}, \tilde{C} > 0$ such that

$$\tilde{c}(1 + |\omega|^2)^{-r} \le \hat{k}_t(\omega) \le \tilde{C}(1 + |\omega|^2)^{-r}. \tag{103}$$

Note that ϕ obtained in this manner satisfies

$$|\hat{\phi}(\omega)| \le C(1 + |\omega|^2)^{-p}. \tag{104}$$

Since $\hat{k}_t = (2\pi)^{-d} \hat{k} * \hat{\phi}$, we have $\partial^\alpha \hat{k}_t = (2\pi)^{-d} \partial^\alpha \hat{k} * \hat{\phi}$ and therefore

$$|\partial^\alpha \hat{k}_t(\omega)| = (2\pi)^{-d} |(\partial^\alpha \hat{k} * \hat{\phi})(\omega)| \le C(1 + |\omega|^2)^{-(r+|\alpha|/2)}, \quad |\alpha| \le M, \tag{105}$$

by the same argument used for the upper inequality in (33).

We now turn to (101). For $\alpha = 0$, this simply follows from (100) combined with $(1 + |\omega|^2) \le (1 + |\omega|)^2$. For $\alpha \ne 0$ such that $|\alpha| \le M$, we obtain by induction that $\partial^\alpha \hat{k}_t^{\frac{1}{2}}$ is of the form

$$\partial^\alpha \hat{k}_t^{\frac{1}{2}} = \sum_{m=1}^{|\alpha|} \sum_{\beta_1 + \cdots + \beta_m = \alpha} C_{\beta_1,\ldots,\beta_m} \hat{k}_t^{\frac{1}{2}-m} \prod_{\ell=1}^m \partial^{\beta_\ell} \hat{k}_t \tag{106}$$

for certain $C_{\beta_1,\ldots,\beta_m} \in \mathbb{R}$. Note that we are allowed to divide by $\hat{k}_t$ since it is positive by (103). Since for $\beta_1, \ldots, \beta_m$ such that $\beta_1 + \cdots + \beta_m = \alpha$, we have

$$\hat{k}_t^{\frac{1}{2}-m}(\omega) \prod_{\ell=1}^{m} \partial^{\beta_\ell} \hat{k}_t(\omega) \leq C(1+|\omega|^2)^{-\frac{r}{2}+mr} \prod_{\ell=1}^{m} (1+|\omega|^2)^{-(r+|\beta_\ell|/2)}, \quad (107)$$

and thus

$$\hat{k}_t^{\frac{1}{2}-m}(\omega) \prod_{\ell=1}^{m} \partial^{\beta_\ell} \hat{k}_t(\omega) \leq C(1+|\omega|)^{-(r+|\alpha|)}, \quad (108)$$

we arrive at (101). $\qquad\square$

Theorem 4.2 *Let k satisfy (45), (46) with $s = r$ and (100) with an integer $M \geq d+1$, and let ϕ be chosen as in Lemma 4.1. Then the functions $F_{\varepsilon,\ell}$ satisfy*

$$|F_{\varepsilon,\ell}(x)| \leq C(1+|x|)^{-M}, \quad (109)$$

for some $C > 0$ that is independent of ℓ and ε.

Proof From its definition (99) we have

$$\widehat{F}_{\varepsilon,\ell}(\omega) = 2^{\ell(d/2+r)} \widehat{\Psi}_{\varepsilon,\ell,0}(2^\ell \omega), \quad (110)$$

and therefore by (95),

$$\widehat{F}_{\varepsilon,\ell}(\omega) = (2\gamma)^{d/2} 2^{\ell(d/2+r)} \hat{k}_t^{\frac{1}{2}}(2^\ell \omega) \, \widehat{\Psi}_{\varepsilon,\ell,0}(2\gamma 2^\ell \omega) = (2\gamma)^{d/2} 2^{\ell r} \hat{k}_t^{\frac{1}{2}}(2^\ell \omega) \, \widehat{\Psi}_\varepsilon(2\gamma\omega), \quad (111)$$

where we have used the scaling relation between $\Psi_{\varepsilon,\ell,0}$ and Ψ_ε. The functions $\widehat{F}_{\varepsilon,\ell}$ are uniformly compactly supported since

$$|\omega|_\infty \geq \frac{8\pi}{6\gamma} \quad \Longrightarrow \quad \widehat{\Psi}_\varepsilon(2\gamma\omega) = 0. \quad (112)$$

Applying partial differentiation for any α such that $|\alpha| \leq M$, and using the multivariate Leibniz formula, we find that

$$\partial^\alpha \widehat{F}_{\varepsilon,\ell}(\omega) = (2\gamma)^{d/2} 2^{\ell r} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \left(2^{\ell|\beta|} \partial^\beta \hat{k}_t^{\frac{1}{2}}(2^\ell \omega) \right) \left((2\gamma)^{|\alpha|-|\beta|} \partial^{\alpha-\beta} \widehat{\Psi}_\varepsilon(2\gamma\omega) \right). \quad (113)$$

The second factor $(2\gamma)^{|\alpha|-|\beta|} \partial^{\alpha-\beta} \widehat{\Psi}_\varepsilon(2\gamma\omega)$ in each term is uniformly bounded independently of ω and β, in view of the smoothness assumption that we have imposed on $\hat{\psi}$. As to the first factor, since we only consider $\frac{2\pi}{6\gamma} \leq |\omega| \leq \sqrt{d}\frac{8\pi}{6\gamma}$, we may use (101) to conclude that

$$|2^{\ell|\beta|} \partial^\beta \hat{k}_t^{\frac{1}{2}}(2^\ell \omega)| \leq C 2^{-r\ell}. \quad (114)$$

It follows that the derivatives $\partial^\alpha \widehat{F}_{\varepsilon,\ell}(\omega)$ are uniformly bounded, independently of ℓ and ε, for all $|\alpha| \le M$, which implies (109) since they are in addition uniformly compactly supported. $\qquad\square$

As we shall show next, Theorem 4.2 implies that Corollary 1.3 can be applied to the wavelet basis defined by (96), that is, with the basis in the corollary chosen as the scaling function $\bar{\Phi}^p$ and the wavelets $\bar{\Psi}^p_{\varepsilon,\ell,n}$ ordered by increasing scale level.

Corollary 4.3 *Under the assumptions of Theorem 4.2, for each $\ell \ge 0$ one has*

$$\sup_{x\in\mathbb{T}} \sum_{\varepsilon\in\mathcal{C}} \sum_{n\in\{0,\dots,2^\ell-1\}^d} |\bar{\Psi}^p_{\varepsilon,\ell,n}(x)| \le C 2^{-\alpha\ell}, \tag{115}$$

with $\alpha := r - \frac{d}{2}$ and C independent of ℓ.

Proof For the summation over n in (115), by (96) and (98) we obtain

$$\sum_{n\in\{0,\dots,2^\ell-1\}^d} |\bar{\Psi}^p_{\varepsilon,\ell,n}(x)| \le 2^{\ell(d/2-r)} \sum_{m\in\mathbb{Z}^d} \sum_{n\in\{0,\dots,2^\ell-1\}^d} |F_{\varepsilon,\ell}(2^\ell x + 2\gamma(2^\ell m - n))|$$

$$= 2^{\ell(d/2-r)} \sum_{k\in\mathbb{Z}^d} |F_{\varepsilon,\ell}(2^\ell x + 2\gamma k)|.$$

By Theorem 4.2,

$$\sum_{k\in\mathbb{Z}^d} |F_{\varepsilon,\ell}(2^\ell x + 2\gamma k)| \le C \sum_{k\in\mathbb{Z}^d} (1+|2^\ell x+2\gamma k|)^{-M} \le C \max_{z\in\mathbb{T}} \sum_{k\in\mathbb{Z}^d} (1+|z+2\gamma k|)^{-M}, \tag{116}$$

and the expression on the right is bounded since $M \ge d + 1$. $\qquad\square$

Remark 4.4 In the estimate (115), the precise value of M enters only into the constant C. For numerical purposes, however, larger values of M corresponding to stronger spatial localization of the functions $\bar{\Psi}^p_{\varepsilon,\ell,n}$ can be advantageous. Note that in the case of the Matérn covariance, if $\phi \in C^\infty$, then Lemma 4.1 can be applied for any integer $M \ge d + 1$. If in addition $\hat{\varphi}, \hat{\psi} \in C^\infty$ holds for the functions generating the Meyer wavelets, then Theorem 4.2 can be applied for any such M as well, and the resulting spatial decay of $\bar{\Psi}^p_{\varepsilon,\ell,n}$ is faster than any polynomial order.

5 Numerical Aspects

We now discuss in more concrete detail the periodic continuation and the resulting KL and wavelet representations in the case of the Matérn covariances (18). In particular we discuss how these representations can be efficiently computed by using FFT and show some numerical examples which reveal the effect of the parameters λ and ν of the Matérn covariance. While these computational principles apply in any dimension, we work in the univariate setting, both for the sake of notational simplicity and for visualization purposes. Our computational domain is thus an interval $\left]-\frac{\delta}{2}, \frac{\delta}{2}\right[$ for some

$\delta > 0$. Note however that varying the correlation length parameter λ in (18) amounts to rescaling the interval. Therefore we fix $\delta = 1$, that is

$$D := \left]-\frac{1}{2}, \frac{1}{2}\right[, \tag{117}$$

and study the effect of varying λ.

5.1 Truncation and Positivity

We need to first choose a ϕ satisfying the conditions in Theorem 2.3 such that $k_t = k\phi$ has nonnegative Fourier transform. One option, which yields $\phi \in C^\infty(\mathbb{R})$, is based on the function θ defined by

$$\theta(x) := \begin{cases} \exp(-x^{-1}), & x > 0, \\ 0, & x \leq 0. \end{cases} \tag{118}$$

We then set

$$\phi(x) := \frac{\theta\left(\frac{\kappa - |x|}{\kappa - \delta}\right)}{\theta\left(\frac{\kappa - |x|}{\kappa - \delta}\right) + \theta\left(\frac{|x| - \delta}{\kappa - \delta}\right)}. \tag{119}$$

Recall that $\kappa = 2\gamma - \delta > \delta$. In view of Theorem 2.3, in order to ensure $\hat{k}_t \geq 0$, it then suffices to take γ sufficiently large.

In order to illustrate the dependence of the required value of γ on the parameters ν, λ of the Matérn covariance, we now describe, for any chosen value of γ, a simple scheme for approximating $\hat{k}_t$ based on the discrete Fourier transform. Let $L > \kappa$ (so that $\mathrm{supp}\, k_t \subset [-L, L]$), and let $N = 2^J$ for some $J > 0$. We consider the approximation of $\hat{k}_t$ by the trapezoidal rule,

$$\hat{k}_t(\omega) = \int_{-L}^{L} k_t(x)\, e^{-i\omega x}\, dx \approx h \sum_{n=-N/2}^{N/2-1} k_t(x_n) e^{-i\omega x_n} =: D_{L,N}(\omega), \tag{120}$$

where $h := 2L/N$ and $x_n := nh$. The sum on the right side can be evaluated by the FFT to obtain the values $D_{L,N}(\omega_k)$ with

$$\omega_k := \frac{\pi k}{L}, \quad k = -\frac{N}{2}, \ldots, \frac{N}{2} - 1. \tag{121}$$

Thus, by making N and L large we may approximately compute $\hat{k}_t$ on an arbitrarily large range of ω and with arbitrarily fine sampling rate.

Since k_t agrees on $[-L, L]$ with its $2L$-periodic extension, we have

$$k_t(x_n) = \sum_{\ell \in \mathbb{Z}} \left(\frac{1}{2L} \int_{-L}^{L} k_t(y)\, e^{-\pi i y \ell/L}\, dy\right) e^{\pi i x_n \ell/L}, \tag{122}$$

and consequently

$$D_{L,N}(\omega_k) = h \sum_{n=-N/2}^{N/2-1} \left(\sum_{\ell \in \mathbb{Z}} \frac{1}{2L} \hat{k}_t(\omega_\ell) \, e^{\pi i x_n \ell / L} \right) e^{-\pi i x_n k / L} = \sum_{m \in \mathbb{Z}} \hat{k}_t(\omega_{k+mN}).$$
(123)

We thus have the error representation

$$|\hat{k}_t(\omega_k) - D_{L,N}(\omega_k)| = \left| \sum_{\substack{m \in \mathbb{Z} \\ m \neq 0}} \hat{k}_t(\omega_{k+mN}) \right|.$$
(124)

In view of (19) and (36), in our present setting we obtain

$$|\hat{k}_t(\omega_k) - D_{L,N}(\omega_k)| \leq C N^{-(2\nu+1)}, \quad k = -\frac{N}{2}, \dots, \frac{N}{2} - 1,$$
(125)

where $C > 0$ is independent of k, L and N.

Based on this approximation, we can check positivity of the obtained approximate values $D_{L,N}(\omega_k)$ for each γ and combine this with a simple bisection scheme to obtain an estimate of the minimum required value $\gamma_{\min}$ for which $\hat{k}_t$ remains non-negative on the chosen grid. As illustrated in Fig. 1, we observe that $\gamma_{\min}$ remains close to its lower bound $\delta = 1$ for $\nu, \lambda < 1$, and shows approximately bilinear growth for larger ν, λ. In other words, the continuation process requires a significantly larger domain as smoothness or correlation length increase. This also implies that the KL or wavelet frames of the RKHS obtained in Sects. 3 and 4 become more redundant as these parameters increase.
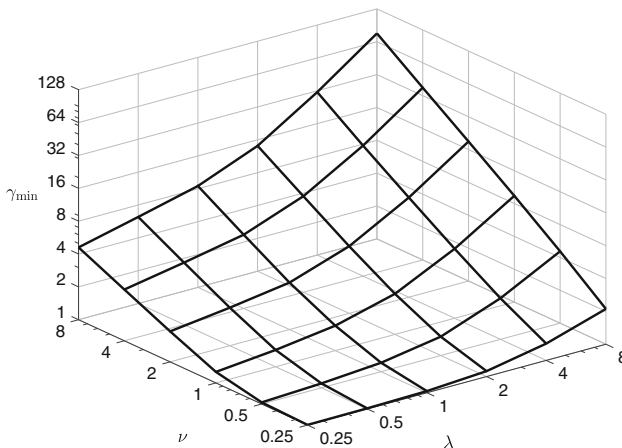
Fig. 1 Numerically observed minimum value of γ required for positivity of $\hat{k}_t$, with k as in (18), in dependence on the Matérn parameters λ, ν

5.2 Matérn Wavelets

The Meyer scaling function and wavelet can be defined by first taking $\hat{\varphi}$ to be a non-negative function such that

$$|\hat{\varphi}(\omega)|^2 = \beta(\omega), \tag{126}$$

where $\beta(\omega)$ is a smooth and even function supported in $[-\frac{4\pi}{3}, \frac{4\pi}{3}]$, such that

$$\beta(\omega) = 1, \quad \omega \in \left[-\frac{2\pi}{3}, \frac{2\pi}{3}\right], \tag{127}$$

and

$$\beta(\pi - \omega) + \beta(\pi + \omega) = 1, \quad \omega \in \left[0, \frac{\pi}{3}\right]. \tag{128}$$

Then, one defines $\hat{\psi}$ by

$$\hat{\psi}(\omega) := \left(\beta(\omega/2) - \beta(\omega)\right)^{1/2} e^{i\omega/2}. \tag{129}$$

One simple example with explicit expressions of $\hat{\varphi}$ and $\hat{\psi}$, following the construction given in [11], is

$$\hat{\varphi}(\omega) := \begin{cases} 1, & |\omega| \leq \frac{2\pi}{3}, \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{3|\omega|}{2\pi} - 1\right)\right), & \frac{2\pi}{3} < |\omega| < \frac{4\pi}{3}, \\ 0, & \text{otherwise}, \end{cases} \tag{130}$$

and

$$\hat{\psi}(\omega) := \begin{cases} \sin\left(\frac{\pi}{2}\nu\left(\frac{3|\omega|}{2\pi} - 1\right)\right)e^{i\omega/2}, & \frac{2\pi}{3} < |\omega| \leq \frac{4\pi}{3}, \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{3|\omega|}{4\pi} - 1\right)\right)e^{i\omega/2}, & \frac{4\pi}{3} < |\omega| \leq \frac{8\pi}{3}, \\ 0, & \text{otherwise}, \end{cases} \tag{131}$$

where we take

$$\nu(x) := \frac{\theta(x)}{\theta(x) + \theta(1 - x)} \tag{132}$$

with θ given by (118).

From these we now construct the one-dimensional versions of $\bar{\Phi}^p$ and $\bar{\Psi}^p$ described in Sect. 4. Recall that for the scaling function, $\bar{\Phi}^p = (2\gamma)^{-1/2}\sqrt{\hat{k}_t(0)}$, which we can directly approximate using (120). For the wavelets, it suffices to consider $\bar{\Psi}^p_{\varepsilon,\ell,0}$ for each wavelet type ε and scale level ℓ, since all further wavelets are obtained as translates of these functions. In the present univariate case, there is only a single wavelet type $\varepsilon = 1$, and we omit the corresponding subscript in what follows.

By (92) and (32), we have

$$\bar{\Psi}^{\mathrm{p}}_{\ell,0}(x) = \frac{1}{2\gamma} \sum_{n \in \mathbb{Z}} \sqrt{\hat{k}_{\mathrm{t}}\left(\frac{\pi}{\gamma}n\right)} \left(\sqrt{2^{1-\ell}\gamma}\,\hat{\psi}(2^{-\ell+1}\pi n)\right) e^{i\frac{\pi}{\gamma}nx}. \qquad (133)$$

We now choose L in (120) as $L = 2\gamma$, that is, $\omega_k = \frac{\pi k}{2\gamma}$. We assume that γ and $N = 2^J$, with $J > 1$, are sufficiently large to ensure $D_{L,N}(\omega_k) \geq 0$ for the range of k in (121). This allows us to approximate the above expression by

$$\tilde{\Psi}^{\mathrm{p}}_{\ell,0}(x) := \frac{1}{\sqrt{2^{\ell+1}\gamma}} \sum_{n=-N/4}^{N/4-1} \sqrt{D_{L,N}(\omega_{2n})}\,\hat{\psi}(2^{-\ell+1}\pi n)\, e^{i\frac{\pi}{\gamma}nx}. \qquad (134)$$

Using the compact support of $\hat{\psi}$ and that $|\hat{\psi}| \leq 1$, we obtain

$$|\bar{\Psi}^{\mathrm{p}}_{\ell,0}(x) - \tilde{\Psi}^{\mathrm{p}}_{\ell,0}(x)| \leq C2^{-\ell/2}\Bigg\{ \sum_{\substack{|n|\geq N/4 \\ \frac{1}{3}2^\ell < |n| < \frac{4}{3}2^\ell}} \sqrt{\hat{k}_{\mathrm{t}}(\pi\gamma^{-1}n)}$$
$$+ \sum_{\substack{n\in\{-N/4,\ldots,N/4-1\} \\ \frac{1}{3}2^\ell < |n| < \frac{4}{3}2^\ell}} \left|\sqrt{\hat{k}_{\mathrm{t}}(\pi\gamma^{-1}n)} - \sqrt{D_{L,N}(\omega_{2n})}\right|\Bigg\}. \qquad (135)$$

Recall by (103) we have $c(1+|\omega|)^{-2\nu-1} \leq \hat{k}_{\mathrm{t}}(\omega) \leq C(1+|\omega|)^{-2\nu-1}$. It follows that the first sum can be bounded according to
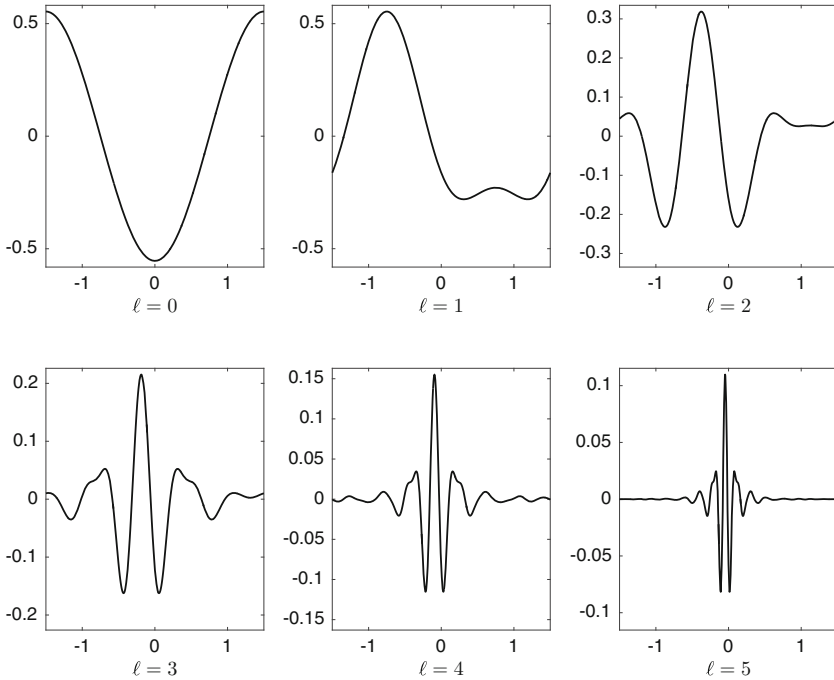
$$\sum_{\substack{|n|\geq N/4 \\ \frac{1}{3}2^\ell < |n| < \frac{4}{3}2^\ell}} \sqrt{\hat{k}_{\mathrm{t}}(\pi\gamma^{-1}n)} \leq C \min\{N^{-\nu+\frac{1}{2}}, 2^\ell N^{-\nu-\frac{1}{2}}\}. \qquad (136)$$

For the second sum, we combine (125) with

$$\left|\sqrt{\hat{k}_{\mathrm{t}}(\pi\gamma^{-1}n)} - \sqrt{D_{L,N}(\omega_{2n})}\right| \leq \frac{\left|\hat{k}_{\mathrm{t}}(\pi\gamma^{-1}n) - D_{L,N}(\omega_{2n})\right|}{\sqrt{\hat{k}_{\mathrm{t}}(\pi\gamma^{-1}n)}}, \qquad (137)$$

to obtain

$$\sum_{\substack{n\in\{-N/4,\ldots,N/4-1\} \\ \frac{1}{3}2^\ell < |n| < \frac{4}{3}2^\ell}} \left|\sqrt{\hat{k}_{\mathrm{t}}(\pi\gamma^{-1}n)} - \sqrt{D_{L,N}(\omega_{2n})}\right| \leq CN^{-2\nu-1} \min\{N^{\nu+\frac{3}{2}}, 2^\ell N^{\nu+\frac{1}{2}}\}. \qquad (138)$$

Fig. 2 Wavelets $\bar{\Psi}^{\mathrm{p}}_{\ell,0}$ obtained with $\lambda = 1$, $\nu = \frac{1}{2}$, and $\gamma = \frac{3}{2}$, for $\ell = 0, \ldots, 5$

This yields

$$|\bar{\Psi}^{\mathrm{p}}_{\ell,0}(x) - \tilde{\Psi}^{\mathrm{p}}_{\ell,0}(x)| \leq C 2^{-\ell/2} \min\{N^{-\nu+\frac{1}{2}}, 2^{\ell} N^{-\nu-\frac{1}{2}}\} = C 2^{-\nu J - \frac{1}{2}|J-\ell|}, \quad (139)$$

where $C > 0$ depends on ν, λ, γ, but not on J or ℓ.

The sum in (134) can be evaluated by FFT simultaneously for the 2^{J-1} arguments

$$x = \frac{4\gamma k}{N}, \quad k = -\frac{N}{4}, \ldots, \frac{N}{4} - 1, \quad (140)$$

at cost of order $J 2^J$. In other words, if we prescribe a grid size $h \sim 2^{-J}$, as a consequence of (139) we can determine the values of the wavelets at any level ℓ at the grid points up to an error of order h^ν, using $h^{-1}|\log h|$ operations in total. Since the wavelets are trigonometric polynomials, their values between grid points can be approximated with similar order of accuracy by local polynomial interpolation of neighboring grid values.

As an illustration, we display in Figs. 2 and 3 the obtained wavelets at scales $\ell = 0, \ldots, 5$ for $\lambda = 1$ and $(\nu, \gamma) = (1/2, 3/2)$, $(\nu, \gamma) = (4, 5)$. Note that these wavelets behave asymptotically similarly to standard wavelets in terms of scale invariance. As expected the size decay in scale depends on ν.
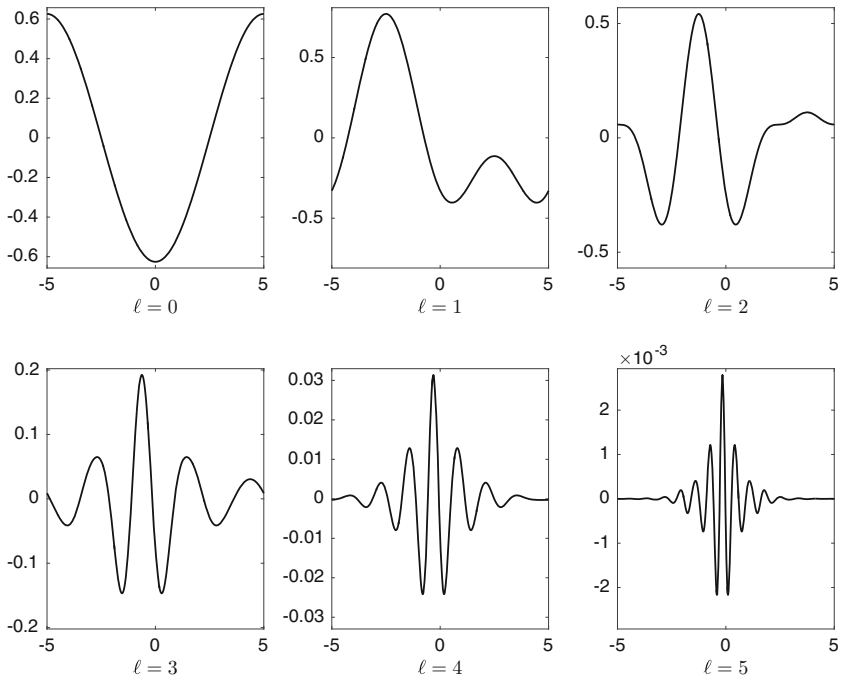
Fig. 3 Wavelets $\bar{\Psi}^{p}_{\ell,0}$ obtained with $\lambda = 1$, $\nu = 4$, and $\gamma = 5$, for $\ell = 0, \ldots, 5$

Remark 5.1 In typical applications, the exact wavelet expansion of the random field b is truncated to a finite number of terms, up to some prescribed error. The estimate (139) shows that in addition we can control the approximation of each wavelet in the uniform sense, with arbirarily high precision governed by J. Combining such estimates allows us to control the resulting error after truncation of the expansion and approximation of each wavelet.

References

1. Aronszajn, N.: Theory of reproducing kernels. Trans. Am. Math. Soc. **68**, 337–404 (1950)
2. Bachmayr, M., Cohen, A., Migliorati, G.: Sparse polynomial approximation of parametric elliptic PDEs. Part I: affine coefficients, ESAIM. Math. Model. Numer. Anal. **51**(1), 321–339 (2017)
3. Bachmayr, M., Cohen, A., DeVore, R., Migliorati, G.: Sparse polynomial approximation of parametric elliptic PDEs. Part II: lognormal coefficients, ESAIM. Math. Model. Numer. Anal. **51**(1), 341–363 (2017)
4. Benassi, A., Jaffard, S., Roux, D.: Elliptic gaussian random processes. Revista Mathemática Iberoamericana **13**, 19–90 (1997)
5. Bonami, A., Karoui, A.: Uniform approximation and explicit estimates for the prolate spheroidal wave functions. Constr. Approx. **43**, 15–45 (2016)

6. Charrier, J.: Strong and weak error estimates for elliptic partial differential equations with random coefficients. SIAM J. Numer. Anal. **50**(1), 216–246 (2012)
7. Ciesielski, Z.: Hölder conditions for realizations of Gaussian processes. Trans. Am. Math. Soc. **99**, 403–413 (1961)
8. Ciesielski, Z., Kerkyacharian, G., Roynette, B.: Quelques espaces fonctionnels associés à des processus gaussiens. Studia Mathematica **107**, 171–204 (1993)
9. Cohen, S., Istas, J.: Fractional Fields and Applications. Springer, New York (2013)
10. Dashti, M., Stuart, A.M.: The Bayesian approach to inverse problems. In: Ghanem, R., Higdon, D., Owhadi, H. (eds.) Handbook of Uncertainty Quantification. Springer, Berlin (2015)
11. Daubechies, I.: Ten lectures on wavelets. In: CBMS-NSF Regional Conference Series in Applied Mathematics, 61, SIAM Philadelphia (1992)
12. Dembo, A., Mallows, C.L., Shepp, L.A.: Embedding nonnegative definite Toeplitz matrices in nonnegative definite circulant matrices, with application to covariance estimation. IEEE Trans. Inf. Theory **35**, 1206–1212 (1989)
13. DeVore, R.: Nonlinear approximation. Acta Numerica **7**, 51–150 (1998)
14. Dietrich, C.R., Newsam, G.N.: Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. SIAM J. Sci. Comput. **18**(4), 1088–1107 (1997)
15. Elliott Jr., F., Horntrop, D., Majda, A.J.: A Fourier-Wavelet Monte Carlo method for fractal random fields. J. Comput. Phys. **132**, 384–408 (1997)
16. Galvis, J., Sarkis, M.: Approximating infinity-dimensional stochastic Darcy's equations without uniform ellipticity. SIAM J. Numer. Anal. **47**(5), 3624–3651 (2009)
17. Gittelson, C.: Stochastic Galerkin discretization of the log-normal isotropic diffusion problem. Math. Models Methods Appl. Sci. **20**(2), 237–263 (2010)
18. Gittelson, C.: Representation of Gaussian fields in series with independent coefficients. IMA J. Numer. Anal. **32**, 294–319 (2012)
19. Graham, I.G., Kuo, F.Y., Nuyens, D., Scheichl, R., Sloan, I.H.: Quasi-Monte Carlo methods for for elliptic PDEs with random coefficients and applications. J. Comput. Phys. **230**(10), 3668–3694 (2011)
20. Graham, I.G., Kuo, F.Y., Nichols, J.A., Scheichl, R., Schwab, Ch., Sloan, I.H.: Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. Numer. Math. **131**, 329–368 (2015)
21. Hoang, V.H., Schwab, C.: N-term Galerkin Wiener chaos approximation rates for elliptic PDEs with lognormal Gaussian random inputs. Math. Models Methods Appl. Sci. **24**, 797–826 (2014)
22. Kerkyacharian, G., Ogawa, S., Petrushev, P., Picard, D.: Regularity of Gaussian processes on Dirichlet spaces. arXiv:1508.00822 (2015)
23. Luschgy, H., Pagès, G.: Expansions for Gaussian processes and Parseval frames. Electron. J. Prob. **14**, 1198–1221 (2009). Paper no. 42
24. Meyer, Y.: Wavelets and Operators. Cambridge Studies in Advanced Mathematics, vol. 37. Cambridge University Press, Cambridge (1992)
25. Slepian, D., Pollak, H.O.: Prolate spheroidal wave functions, Fourier analysis and uncertainty—I. Bell Syst. Tech. J. **40**, 43–64 (1961)
26. Watson, G.N.: A Treatise on the Theory of Bessel Functions, 2nd edn. Cambridge University Press, Cambridge (1995)
27. Widom, H.: Asymptotic behaviour of the eigenvalues of certain integral operators. Trans. Am. Math. Soc. **109**, 278–295 (1963)
28. Wood, A.T.A., Chan, G.: Simulation of stationary Gaussian processes in $[0, 1]^d$. J. Comput. Graph. Stat. **3**(4), 409–432 (1994)