A Computational Introduction to Random Functions

Emilio Porcu and Robert Schaback

October 8, 2025

Abstract

For readers working in applications, this paper serves as a basic primer for the calculation of (pseudo-) random functions. Keeping probabilistic arguments at a minimum in favour of computation, random functions are studied by randomizing coefficients of expansions. This reveals a natural connection to Reproducing Kernel Hilbert spaces. Randomizing the coefficients of orthonormal Newton bases of Reproducing Kernel Hilbert spaces allows a comprehensive, constructive, and computational presentation, providing regularity results and error bounds in terms of variances. The approach is unexpectedly general, because paths of most random fields are shown to arise this way. A natural extension allows to use expansions in L_2 spaces, in particular wavelets and finite elements. Smoothing operators are introduced for the transition from white noise to random functions in higher-order Sobolev spaces, without solving stochastic partial differential equations. Numerical examples serve for illustration.

1 Introduction

It is well-known how to calculate (pseudo-) random numbers. All useful programming languages have routines for this. But how to calculate (pseudo-) random functions? These are useful for testing the behaviour of numerical algorithms to see how the results vary for varying inputs with specified properties [5, 22]. Therefore this paper primarily addresses readers working on computational applications, and only a minimum of probabilistic or stochastic background is assumed. The focus is on functions as programmable mappings with prescribed smoothness properties, not finite sets of random values. Because Sobolev spaces are central to numerical algorithms, a goal is to calculate (pseudo-) random functions from Sobolev spaces.

For random functions, hardly any literature is on the market. Even random sequences are widely avoided topics. In particular, *"Axiomatic probability theory deliberately avoids a definition of a random sequence"* [2, p. 44] and *"The Bourbaki school considered the statement* "let us consider a random sequence" *an abuse of language* [36, p.166], cited after the 2025 wikipedia entry for "Random Sequence". We do not discuss the reasons here. Nevertheless, Spatial Statistics consider *random fields* and use their

paths acting as random functions. This uses heavy probabilistic machinery to circumvent the aforementioned problem. We ignore most of the vast literature in this direction, taking a computational shortcut, being strongly influenced by [12–14, 19, 31, 34], suggested for further reading.

The easiest way to generate random functions on a domain Ω is to start from finite linear combinations

$$f_{W,V}(x) := \sum_{k=1}^{n} w_k(x) v_k,\ x \in \Omega \tag{1}$$

of given basis functions w_k, $1 \le k \le n$, using independently sampled random scalar coefficients v_k with mean zero and unit variance. Looking at the properties of such random functions, Section 2.1 shows that *kernels*

$$C_W(x,y) = \sum_{k=1}^{n} w_k(x) w_k(y),\ x, y \in \Omega \tag{2}$$

now arise necessarily as *covariance functions* that connect the covariance of values at different points. The pointwise values of such random functions are events of classical random variables, and these define a random field. In contrast to the standard literature on random fields, we start with random functions to obtain random fields, not the other way round.

Section 2.2 turns to expansions into infinite series, and then summability of

$$C_W(x,x) = \sum_{k=1}^{\infty} w_k^2(x),\ x \in \Omega$$

must be assumed and allows to define a kernel via (2) for $n = \infty$. It induces a Hilbert space $\mathscr{H}(C_W)$ where the w_k of $W := \{w_k\}_{k\in\mathbb{N}}$ are orthonormal, and all theoretical and computational tools based on kernels are available [see 4, 10, 30, 38, with the references therein].

The case $n = \infty$ requires special care: the pointwise values of such random functions are values of *random series*. They are what Section 2.2 calls *generalized* random variables, ignoring the probability spaces for series of random variables. Then Theorem 2 in Section 2.3 states that the theory of random functions based on series with random coefficients is equivalent to the theory of reproducing kernel Hilbert spaces where orthonormal expansions have uncorrelated random coefficients. The values of random functions at points are *generalized* random variables, defining a *generalized random field*.

Section 3.1 uses this to show in Theorem 3 that pointwise properties of random functions based on orthonormal expansions into $W = \{w_k\}_{k\in\mathbb{N}}$ are equivalent to pointwise properties of deterministic functions in the Hilbert space $\mathscr{H}(C_W)$. The connection is by the admissible functionals in the dual space $\mathscr{H}(C_W)^*$, for instance pointwise derivatives. They define generalized random variables whose variance equals the square of their norm in $\mathscr{H}(C_W)^*$.

Such a close connection between deterministic and stochastic regularity does not hold in the global case, i.e. when norms of random functions are considered.

Since the random expansion coefficients of a random function in terms of an orthonormal system $W = \{w_k\}_{k\in\mathbb{N}}$ are almost surely not square summable, the function does almost surely not lie in the Hilbert space $\mathscr{H}(C_W)$. But it may lie in larger spaces with weaker norms almost surely, and such spaces are characterized by *weighted* expansions in Section 3.3.

Then Section 4 uses the efficiently computable *Newton basis* of any kernel-based Hilbert space $\mathscr{H}(C)$ for actual computation of (pseudo-) random functions. While the standard calculation of values of random fields works on finite sets only, using the Cholesky decomposition of kernel matrices, our use of the Newton basis provides a series of computable random *functions*. The finite partial sums of the series are functions in the Hilbert space $\mathscr{H}(C_W)$, and their values are classical random variables, but the limit in the Hilbert space norm does not exist. However, the variance of the truncation error can be calculated pointwise and it converges to zero as fast as the deterministic kernel expansion converges pointwise, see Theorem 7.

Section 5 is the only one that addresses readers with a full background in Probability and Statistics. It compares the approach of this paper with the bottom-up construction of random functions via their random values, i.e. as *paths* of *random fields*. The main result in Theorem 8 is that all given second-order random fields coincide pointwise, and up to random variables with mean and variance zero, *i.e.* almost surely, with a generalized random field constructed by the approach of this paper. Our top-down approach from expansions to random functions to random fields does not lose generality when compared to the bottom-up strategy starting from random fields to define random functions.

The connection of expansions to kernels and Hilbert spaces is strong enough to allow a generalization to the Hilbert space $L_2(\Omega)$ in Section 6, using the duality between functions and functionals there. Going to weighted expansions, this leads to generalized Sobolev spaces H^β as subspaces of $L_2(\Omega)$, and Theorem 9 proves that orthonormal systems in $L_2(\Omega)$ lead to generalized random functionals that are in $H^{-\beta}$ if $\beta > 1/2$, but almost surely not if $\beta \le 1/2$, in particular not in $H^0 = L_2$. Haar wavelets are an orthonormal basis in L_2 that serves as an example for generalized random functions in Section 6.2.

The transition from a Sobolev space of lower order to a space of higher order is a *smoothing* map, and in the scale of classical global Sobolev spaces on $\mathbb{R}^d$, it can be realized as a convolution with a Matérn kernel [24]. When starting from cheaply computable *white noise*, this yields another bottom-up technique for constructing random functions. This is sketched in Section 6.3, including the case of smoothing a randomized finite element expansion. While the standard literature [15, 16] on random fields implements smoothing maps via stochastic partial differential equations, this approach avoids PDEs and smoothes the basis or uses weighted expansions.

Section 7 concludes the paper, providing numerical examples.

2 Randomized Expansions

The simplest way to construct random functions is to take a basis of a space of real-valued functions on a domain Ω and form linear combinations (1) with random coefficients.

2.1 Finitely many basis functions

For a basis $w_1, \ldots, w_n$ of a space of real-valued functions on a domain Ω, we form linear combinations (1) with random coefficients $v_1, \ldots, v_n$ that are realizations of random variables $V_1, \ldots, V_n$. From here on, all random variables will be defined over $\mathbb{R}$ and always assumed to be second-order, i.e. they have well-defined mean and variance. With no loss of generality, the random variables V_k are assumed to have zero mean and unit variance. Such random functions are easy to compute, if real-valued (pseudo-) random number generators are available.

For fixed $x \in \Omega$, the random values $f_{W,V}(x)$ are events (realizations) from a random variable $f_W(x)$ having zero mean. If the random coefficients V_k are uncorrelated, two such random variables have covariance

$$Cov(f_W(x), f_W(y)) = \sum_{k=1}^{n} w_k(x) w_k(y),$$

and the variance of $f_W(x)$ is

$$Var(f_W(x)) = \sum_{k=1}^{n} w_k(x)^2.$$

The numerical calculation of coefficients is easy if the random variables $V_1, \ldots, V_n$ are independent, but all other cases become computationally cumbersome, especially when n becomes large.

If users want a specified covariance function C, they have to pick a basis $w_1, \ldots, w_n$ with

$$C(x,y) = \sum_{k=1}^{n} w_k(x) w_k(y) \text{ for all } x, y \in \Omega, \tag{3}$$

and all such bases do the job. However, if the user seeks for a specific covariance function in an algebraically closed form, they will often need to go to $n = \infty$, see the next section. For a given basis, the positive semidefinite finite-rank kernel (3) is a consequence. Choosing different uncorrelated distributions for the V_k will change the distributions of the random variables $f_W(x)$ while keeping zero mean and the above variance. If all coefficient distributions are standardized normal (Gaussian) distributions, and if all coefficients are uncorrelated, they are also independent and the distribution of $f_W(x)$ is normal, and fully defined by the above variance.

This is independent of the choice of W and C. Prescribing a specific covariance function will ony rarely work with finitely many functions. Therefore we need to let n go to infinity in the next section.

2.2 Going to Infinity

This section mimicks the previous one, with the additional requirement that the number of elements in the considered linear combination is arbitrarily large ($n \to \infty$). To avoid pathological cases, we need to work under the additional assumption that

$$\sum_{k=1}^{\infty} w_k(x)^2 < \infty, \tag{4}$$

which ensures convergence of the resulting expansion-based covariance, obtained through

$$C(x,y) = \sum_{k=1}^{\infty} w_k(x) w_k(y) \text{ for all } x, y \in \Omega. \tag{5}$$

For a fixed point x in the domain Ω, we have real-valued random events

$$f_{W,V}(x) := \sum_{k=1}^{\infty} w_k(x) v_k,\ x \in \Omega. \tag{6}$$

This is a series of second-order uncorrelated random variables, and the sum of variances is convergent by (4). The mean-square limit $f_W(x)$ exists almost surely, and its variance is given by (4). We ignore the peculiarities of the probability spaces for series of random variables and call sums of such series *generalized random variables*.

More generally , if $\{\alpha_n\}_{n\in\mathbb{N}} \in \ell_2$ is a square summable sequence and if V_n are second-order zero-mean random variables with variance one, the random series

$$V_\alpha := \sum_{n\in\mathbb{N}} \alpha_n V_n$$

is a generalized random variable that has the variance

$$Var(V_\alpha) = \|\alpha\|_{\ell_2}^2 = \sum_{n\in\mathbb{N}} \alpha_n^2.$$

Furthermore, the covariance between two such variables is

$$Cov(V_\alpha, V_\beta) = (\alpha, \beta)_{\ell_2} = \sum_{n\in\mathbb{N}} \alpha\beta.$$

So far, we have

Theorem 1. *Under the assumption (4), the pointwise random values $f_{W,V}(x)$ of (6) define a generalized random variable $f_W(x)$ with variance*

$$Var(f_W(x)) = \sum_{n\in\mathbb{N}} w_n^2(x) < \infty$$

which has finite values almost surely. The covariance between any pair of generalized random variables at points x and y is then attained through

$$Cov(f_W(x), f_W(y)) = \sum_{n\in\mathbb{N}} w_n(x) w_n(y), \tag{7}$$

and the function

$$C(x,y) := \sum_{n\in\mathbb{N}} w_n(x)w_n(y) \tag{8}$$

is a positive semidefinite expansion kernel [8] that is a covariance function via (7). □

Such kernels are abundant in Machine Learning [6, 32, 33].

2.3 Connection to Hilbert Spaces

Defining $(w_n, w_m)_{\mathscr{H}(C)} := \delta_{nm}$ with the Kronecker symbol leads to a Hilbert space $\mathscr{H}(C)$ with reproducing kernel C in which the w_n are orthonormal. The converse is also true: If the w_n are an orthonormal system in a Hilbert space $\mathscr{H}(C)$ where C is reproducing, the above series representation follows. In statistics, it is customary to seek for a process having a specified correlation function (which is by construction positive semidefinite). A simple recipe is to consider the Hilbert space generated by the desired correlation C, and then pick an orthonormal expansion $W = \{w_n\}_{n\in\mathbb{N}}$ that is defined therein.

We summarize:

Theorem 2. *Random functions based on series (6) with (4) and uncorrelated random coefficients lead to the Hilbert space $\mathscr{H}(C)$ for a kernel C defined by (5), and the expansion is orthonormal there. The values of such random functions at points are generalized random variables.* □

3 Properties of Random Functions

Section 3.3 will analyze the classical deterministic regularity of single random functions calculated using expansions. But first we look at regularity statements that hold for all samples. They can be seen as probabilistic properties of our algorithm for computing random functions. We shall connect these properties to standard deterministic properties of the covariance kernel.

3.1 Pointwise Regularity

By Theorem 1, pointwise values of random functions of the form (6) are generalized random variables and have a finite value almost surely. They define a *generalized random field* on Ω. We shall clarify in Section 5 how the standard notion of random fields fits in here, but note that we do not need to go into this at all under a computational viewpoint.

However, we can deal with pointwise derivatives of random functions and show under which conditions they exist and are generalized random variables. They can be defined as limits of divided differences, and the latter are linear combinations

$$\lambda_{X,a}(f) := \sum_{x_j\in X} a_j f(x_j)$$

of point evaluation functionals, for finite sets $X \subset \Omega$ and vectors $a \in \mathbb{R}^{|X|}$. They are in the dual $\mathscr{H}(C)^*$ of the native Hilbert space $\mathscr{H}(C)$ of C and have norm

$$\|\lambda_{X,a}(f)\|^2_{\mathscr{H}(C)^*} := \sum_{x_j,x_k \in X} a_j a_k C(x_j,x_k) = \lambda^u_{X,a}\lambda^v_{X,a}C(u,v).$$

Here, the superscript denotes the variable the functional acts on. We can consider all well-defined limits λ of those in $\mathscr{H}(C)^*$, and they have norms

$$\|\lambda\|^2_{\mathscr{H}(C)^*} = \lambda^u\lambda^v C(u,v).$$

If pointwise functionals like $f \mapsto D^\alpha f(x)$ involving multivariate derivatives of order α have finite norm in $\mathscr{H}(C)^*$, i.e.

$$(D^\alpha)^u(x)(D^\alpha)^v(x)C(u,v) < \infty,$$

they are continuous on $\mathscr{H}(C)$ and characterize the pointwise regularity of these functions.

To move towards randomness, we apply those functionals to random functions based on orthonormal expansions $W = \{w_n\}_{n\in\mathbb{N}}$, i.e.

$$\lambda_{X,a}(f_{W,V}) := \sum_{x_j \in X} a_j f_{W,V}(x_j) = \sum_{n\in\mathbb{N}}\sum_{x_j\in X} a_j w_n(x_j)v_n.$$

This is a generalized random variable if

$$\sum_{n\in\mathbb{N}}\left(\sum_{x_j\in X} a_j w_n(x_j)\right)^2$$

is finite, and this is the squared norm of $\lambda_{X,a}$ in $\mathscr{H}(C)^*$, independent of the chosen expansion. This extends to the limits in $\mathscr{H}(C)^*$, and we get that all continuous functionals on $\mathscr{H}(C)^*$ lead to generalized random variables when applied to the random functions of (6). In a sloppy formulation:

Theorem 3. *Pointwise regularity of functions in a Hilbert space generated by a kernel C is the same as pointwise regularity in the mean-square sense of the generalized random fields based on arbitrary orthonormal expansions of C.* □

3.2 Mean-Square Continuity

If we specialize the previous section to functionals λ_x supported on a single point $x \in \Omega$, like point or derivative evaluations at x, we see that they define a zero-mean generalized random variable $\lambda_{x,W}$ with events

$$\lambda_x(f_{W,V}) = \sum_{n\in\mathbb{N}} \lambda_x(w_n)v_n$$

and variance

$$Var(\lambda_{x,W}) = \|\lambda_x\|^2_{\mathscr{H}(\mathbb{C})^*} = \lambda^u_x\lambda^v_x C(u,v)$$

independent of the expansion W. This defines a generalized random field with covariance

$$Cov(\lambda_{x,W}, \lambda_{y,W}) = \lambda_x^u \lambda_y^v C(u,v) =: C_\lambda(x,y).$$

Theorem 4. *[1]*
Mean-square continuity of this field is equivalent to continuity of its covariance kernel near the diagonal, i.e.

$$\lim_{x\to y} C_\lambda(x,x) = C_\lambda(y,y) \text{ and } \lim_{x\to y} C_\lambda(x,y) = C_\lambda(y,y) \text{ for all } x,y \in \Omega.$$

3.3 Norms of Random Functions

The random functions $f_{W,V}$ are almost surely not in the Hilbert space $\mathscr{H}(C)$, because this would require square summability of the v_n. In which spaces do they lie?

Assume that the pointwise convergence rate of the $w_n(x)$ is good enough to satisfy

$$\sum_{n\in\mathbb{N}} w_n^2(x) n^{2\beta} < \infty \tag{9}$$

for some $\beta > 1/2$. This defines a scale of spaces with a parametric class of kernels

$$C_\gamma(x,y) := \sum_{n\in\mathbb{N}} w_n(x) w_n(y) n^{2\gamma}$$

for $\gamma \in [0,\beta]$, and the elements $w_n n^\gamma$ are orthonormal in the correspondent Hilbert space $\mathscr{H}(C_\gamma)$, leading to

$$(w_n, w_m)_{C_\gamma} = \delta_{nm} n^{-2\gamma},\ n,\ m, \in \mathbb{N}.$$

This construction is well-known from trigonometric or other orthogonal series, and it applies also to Sobolev spaces for certain decay rates and eigenfunction expansions. Improved rates in the sense of (9) means more smoothness of the functions.

Under (9), we have that there exists a class of expansion fields $\{f_{W,V,\gamma},\ \gamma \in [0,\beta]\}$, such that

$$f_{W,V,\gamma}(x) = \sum_n w_n(x) \tilde{v}_n$$

with $\tilde{v}_n := v_n n^\gamma$, having a covariance function C_γ.

The functions of the form (6) now lie in $\mathscr{H}(C_\gamma)$ if and only if

$$\sum_{n\in\mathbb{N}} (w_n, w_n)_{C_\gamma} v_n^2 = \sum_n n^{-2\gamma} v_n^2 < \infty,$$

and this is true for random functions almost surely if $\gamma > 1/2$. Consequently,

Theorem 5. *The above random functions lie almost surely in all spaces $\mathscr{H}(C_\gamma)$ for $\gamma > 1/2$, but almost surely not in $\mathscr{H}(C_{1/2})$ and all smaller spaces like $\mathscr{H}(C) = \mathscr{H}(C_0)$.*

In practice, one has to truncate the series, and then the random functions lie in $\mathscr{H}(C)$. They have an "asymptotic roughness" that forces the limit to a larger space. This calls for an examination of the limit process.

To come closer to Sobolev spaces, one can turn the argument upside-down and work with an orthonormal system $U = \{u_n\}_{n\in\mathbb{N}}$ in $L_2(\Omega)$ on a bounded domain $\Omega \subset \mathbb{R}^d$ for a fixed dimension. Then a scale of spaces $H^m(\Omega)$ norm-equivalent to classical Sobolev spaces $W_2^m(\Omega)$ arises by defining $w_{n,m} := u^n n^{-m/d}$ to be orthonormal in $H^m(\Omega)$. This behaviour mimics what happens for Mercer/Karhunen-Loève expansions of the Matérn kernels $\mathscr{M}_{m-d/2,1}$ in the notation of [24] of standard Sobolev spaces $W_2^m(\mathbb{R}^d)$. By [28], the eigenvalues decay like $n^{-2m/d}$, but we suppress further details. The kernel of $H^m(\Omega)$ then is

$$C_m(x,y) := \sum_{n\in\mathbb{N}} w_{n,m}(x)w_{n,m}(y) = \sum_{n\in\mathbb{N}} n^{-2m/d}u_n(x)u_n(y)$$

and the correspondent random functions are

$$f_{U,m,V}(x) = \sum_{n\in\mathbb{N}} u_n(x)n^{-m/d}v_n.$$

Their norm in $H^\mu(\Omega)$ is given by

$$\begin{aligned}\|f_{U,m,V}\|^2_{H^\mu(\Omega)} &= \sum_{n\in\mathbb{N}} \|u_n\|^2_{H^\mu(\Omega)} n^{-2m/d} v_n^2 \\ &= \sum_{n\in\mathbb{N}} n^{2\mu/d} n^{-2m/d} v_n^2\end{aligned}$$

and this is finite almost surely if $\mu < m - d/2$.

Theorem 6. *Random functions based on orthonormal expansions in $H^m(\Omega)$ lie almost surely in all spaces $H^\mu(\Omega)$ for $\mu < m - d/2$, but almost surely not in $H^{m-d/2}(\Omega)$ and all smaller spaces.* □

However, the above derivation has two shortcomings. First, it uses the asymptotics of the Mercer expansion, which is computationally expensive. Second, the intermediate Sobolev spaces are defined by changes of weights and have nonstandard expansion kernels. They should be norm-equivalent to the classical Sobolev spaces, but this needs proof.

4 Calculation of Random Functions via Newton Bases

It is well-known that kernels have two sorts of orthonormal bases. The Mercer theorem [20] applied above yields eigenfunctions of the integral operator defined by the kernel, and these lead to Karhunen-Loéve expansions under a probabilistic interpretation. The other case uses the Newton basis [21, 23] and is computationally much cheaper. The probabilistic literature uses it for their standard sampling algorithms, ignoring the underlying orthonormal system of functions. Here, we focus on the deterministic Newton

basis expansion and randomize it along the lines of the previous sections. But recall that pointwise regularity of random functions does not depend on the expansion chosen, by Theorem 3.

The Newton basis for a sequence $x_1, x_2, \ldots$, of given points is recursively defined by

$$\begin{aligned} N_j(x_j)^2 &= C(x_j,x_j) - \sum_{m=1}^{j-1} N_m(x_j)^2,\ j \geq 1, \\ N_j(x)N_j(x_j) &= C(x,x_j) - \sum_{m=1}^{j-1} N_m(x)N_m(x_j),\ j \geq 1,\ x \in \Omega. \end{aligned} \tag{10}$$

Then the generated functions have the properties

$$\begin{aligned} N_j(x) &\in \text{span}\{C(x,x_1),\ldots,C(x,x_j)\}, \\ N_j(x_k) &= 0,\ 1 \leq k < j, \end{aligned}$$

where the second property is shared with the Newton basis polynmials

$$N_j(x) = \prod_{k=1}^{j-1}(x - x_k),\ j \geq 1$$

of order j for univariate interpolation. The basis is orthonormal in the Hilbert space $\mathscr{H}(C)$ and (8) takes the form

$$C(x,y) = \sum_{j \in \mathbb{N}} N_j(x)N_j(y) \text{ for all } x,y \in \Omega. \tag{11}$$

The basis has uniformly bounded values due to

$$C(x,x) = \sum_{j \in \mathbb{N}} N_j(x)^2 \text{ for all } x \in \Omega,$$

and it yields Cholesky decompositions of all kernel matrices with entries $C(x_i,x_k),\ 1 \leq i,k \leq n$ via

$$C(x_i,x_k) = \sum_{j=1}^{min(i,k)} N_j(x_i)N_j(x_k),\ 1 \leq i,k \leq n. \tag{12}$$

But here we focus on functions, not on matrices, and we can define generalized random variables $f_N(x)$ with values

$$f_{N,V}(x) := \sum_{j \in \mathbb{N}} N_j(x)v_j \text{ for all } x \in \Omega$$

based on samples v_j from uncorrelated standardized random variables V_j. At points x_k the series is finite, namely

$$f_{N,V}(x_k) := \sum_{j=1}^{k} N_j(x_k)v_j \text{ for all } k \in \mathbb{N},$$

and it defines a classical random variable $f_N(x_k)$ that depends only on the first k samples $v_1,\dots,v_k$. The partial sums

$$f_{N,V,n}(x) := \sum_{j=1}^{n} N_j(x) v_j \text{ for all } x \in \Omega$$

are in the Hilbert space $\mathscr{H}(C)$, defining classical random variables $f_{N,n}(x)$, and their recursion is

$$f_{N,V,n+1}(x) = f_{N,V,n}(x) + N_{n+1}(x) v_{n+1}$$

for a new sample v_{n+1}.

Theorem 7. *The variance of the truncation error of the partial sums is*

$$Var(f_N(x) - f_{N,n}(x)) = C(x,x) - \sum_{j=1}^{n} N_j^2(x) = \sum_{j=n+1}^{\infty} N_j^2(x) \tag{13}$$

converging to zero at the same rate as (11). The decrease of the variance of the truncation at x in step n is $N_n^2(x)$.

The deterministic function (13) is the square of the well-known *Power Function* $P_{x_1,\dots,x_n}(x)$ arising in error bounds for deterministic kernel-based interpolation and as the variance of the *Kriging* error in Spatial Statistics.

It vanishes at $x_1,\dots,x_n$, and a good choice for x_{n+1} is the maximum of it in Ω. This is the well-known *P-greedy* strategy dating back to [9], with asymptotically optimal properties due to [27]. There is a matrix-free implementation [23] of the construction of Newton bases, and it can easily be adapted to generate random functions by picking random coefficients. It will be used for numerical examples in Section 7, and we shall use

$$\|Var(f_N(x) - f_{N,n}(x))\|_\infty \le Var(f_N(x_{n+1}) - f_{N,n}(x_{n+1})) = N_{n+1}(x_{n+1})^2 \tag{14}$$

to control the maximum of all pointwise variances in (13).

Corollary 1. *If the P-greedy method is used for point selection, the variance of the truncation error converges uniformly on the set of eligible points, and the rate is the same as for (11).* □

Much more information on the truncation error can be found in [39] for the Gaussian case, and in [34], both focusing on probability theory.

In Sobolev spaces $W_2^m(\Omega)$ for $\Omega \subset \mathbb{R}^d$ and $m > d/2$, the Matérn kernel $\mathscr{M}_{m-d/2,1}$ in the notation of [24] is reproducing, and one can use the Newton basis there. If the P-greedy point selection is used, [27] shows that

$$Var(f_N(x) - f_{N,n}(x)) \le c n^{1-2m/d} \tag{15}$$

holds uniformly in Ω, but numerical evidence in [27] shows convergence at least like $n^{-2m/d}$. This implies that the partial sums of random expansions converge at a rate the increases with m, but by Theorem 6, the limit is not in $W_2^m(\Omega)$ though the partial sums are. The "limiting" space is $W_2^{m-d/2}(\Omega)$. Figure 2 will show examples.

5 Starting From Random Fields

For readers interested in computations, this section can be skipped. But for readers with a background in Probability and Statistics, we explain the connection to the vast literature on Random Fields here, but leave details to the references [14, 19, 31, 34] close to this paper.

In the nondeterministic literature it is customary [1, 3] to start with a *random field* that has a random variable $R(x)$ at each point x of Ω in such a way that hidden joint distributions behind the scene allow to say that

$$Cov(R(x),R(y)) = C(x,y) \text{ for all } x,y \in \Omega$$

is a positive semidefinite *covariance function*, i.e. a positive semidefinite kernel. To arrive bottom-up at a random function, called a *path* of the random field, is somewhat problematic because it is unclear how to sample infinitely many correlated random variables simultaneously, and how to work with joint distributions of infinitely many random variables. Recall how Section 2 circumvents this problem, allowing to proceed top-down by constructing random functions first, using randomized expansions, and taking their values at points afterwards to define a generalized random field. The literature has randomized expansions as well, but mainly focusing on Karhunen-Loéve expansions that are based on Mercer expansions [20]. Randomized expansions into the Newton basis like in Section 4 are behind the curtains, when sampling is done via Cholesky decompositions of kernel matrices along (12). This is the standard technique for calculating finitely many values of paths of random fields, via a decorrelation.

Since the bottom-up approach is not restricted to expansions, it may be that it is much more general. But we shall show that both approaches are equivalent under mild hypotheses.

We assume a second-order zero-mean random field $x \mapsto R(x)$ on a bounded domain $\Omega \subset \mathbb{R}^d$ that has a well-defined covariance function C_R. Then we define the Hilbert space V_R as the closure of linear combinations of all $R(x)$ under covariance as an inner product, and modulo random variables of mean and variance zero. The map $C_R(x,\cdot) \mapsto R(x)$ then is the Loéve isometry [18] between the reproducing kernel Hilbert space $\mathscr{H}(C_R)$ and V_R. Since we have both C_R and $\mathscr{H}(C_R)$, we can take any orthonormal system $W = \{w_n\}_{n\in\mathbb{N}}$ in $\mathscr{H}(C_R)$ and map it to an uncorrelated system $\{R_{W,n}\}_{n\in\mathbb{N}}$ of generalized random variables in V_R. Then

$$Cov(R(x),R_{W,n}) = (C_R(x,\cdot),w_n)_{\mathscr{H}(C_R)} = w_n(x) \text{ for all } n \in \mathbb{N}, x \in \Omega.$$

The expansion

$$C_R(x,\cdot) = \sum_{n\in\mathbb{N}} w_n(x) w_n(\cdot) \text{ for all } x \in \Omega$$

then maps to an expansion

$$R(x) = \sum_{n\in\mathbb{N}} w_n(x) R_{W,n} \text{ for all } x \in \Omega$$

which is an identity in V_R, i.e. it holds modulo generalized random variables of mean and variance zero.

Theorem 8. *Up to pointwise modifications in the mean-square sense, all given zero-mean second-order random fields with a well-defined covariance function coincide pointwise with an expansion-based generalized random field with uncorrelated coefficients.* □

This holds for all expansions, but each expansion W has its correspondent uncorrelated generalized random variables $R_{W,n}$ for generating the coefficients. Paths of random fields, like functions in $L_2(\Omega)$, admit certain modifications almost everywhere, but the computation of random functions via expansions avoids such ambiguities.

If the Newton basis on a countable point set $X_\infty = \{x_n\}_{n\in\mathbb{N}}$ is used for the orthonormal system, the coefficient distributions can be recursively characterized. The first coefficient is determined by $R(x_1)$, and the $(n+1)$-st follows the joint distribution on $x_1,\ldots,x_{n+1}$ conditioned to the values $y_1,\ldots,y_n$ that $f_{N,V,n}$ has on $x_1,\ldots,x_n$ after determining $v_1,\ldots,v_n$. The variance of that distribution is $P^2_{x_1,\ldots,x_n}(x_{n+1})$ using the Power function, see (13). It is the Kriging variance for prediction at x_{n+1} from values at $x_1,\ldots,x_n$ for given values $y_1,\ldots,y_n$ there. In case of Gaussianity, it is clear how to sample v_{n+1}, but in general the conditional distribution is hard to calculate except for the known variance, and there is no guaranteed independence though there is uncorrelatedness. This is a strong limitation for practical algorithms in the general case.

Conversely, if an expansion-based random function generator is working with independent coefficients for computational simplicity, it will fail to be fully general when there is no Gaussianity. However, the class of generalized random fields calculated via expansions and independent coefficients without Gaussianity is interesting in itself and deserves further study.

6 Expansions Related to L_2

Expansions like (6) can make sense also in $L_2(\Omega)$, provided that convergence takes place in $L_2(\Omega)$.

6.1 The Kernel of L_2

The "kernel" now has no continuous function values, and (8) will not make sense. But a kernel of $L_2(\Omega)$ is well-defined when replacing point evaluation functionals by functionals $\lambda_g \in L_2(\Omega)^*$ for $g \in L_2(\Omega)$ with

$$\lambda_g(f) := (f,g)_{L_2(\Omega)} \text{ for all } f,g \in L_2(\Omega).$$

The kernel then is

$$C(\lambda_f,\lambda_g) := (f,g)_{L_2(\Omega)} = (\lambda_f,\lambda_g)_{L_2(\Omega)^*} \text{ for all } f,g \in L_2(\Omega),$$

generalizing the standard pointwise kernel property

$$C(x,y) = (\delta_x, \delta_y)_{\mathscr{H}(C)^*} \text{ for all } x,y \in \Omega$$

based on point evaluation functionals $\delta_x : f \mapsto f(x)$.

"Pointwise" generalized random variables $f_{W,g}$ are now defined for any "point" $g \in L_2(\Omega)$ or $\lambda_g \in L_2(\Omega)^*$ via sample values

$$V \mapsto f_{W,V,g} := \sum_{k=1}^{\infty} (w_k, g)_{L_2(\Omega)} v_k = \sum_{k=1}^{\infty} \lambda_g(w_k) v_k,\ g \in L_2(\Omega),\ \lambda_g \in L_2(\Omega)^*,$$

and we have

$$Cov(f_{W,g}, f_{W,h}) = (g,h)_{L_2(\Omega)} \text{ for all } g,h \in L_2(\Omega).$$

Theorem 2 holds in that "pointwise" sense.

Global convergence results require a Sobolev-type scale of spaces

$$H^\beta := \left\{ f \ :\ \|f\|_{H^\beta}^2 := \sum_{n\in\mathbb{N}} (f, w_n)_{L_2(\Omega)}^2 n^{2\beta} < \infty \right\} \tag{16}$$

like in (9), with duals

$$H^{-\beta} := \left\{ \lambda \ :\ \|\lambda\|_{H^{-\beta}}^2 := \sum_{n\in\mathbb{N}} \lambda(w_n)^2 n^{-2\beta} < \infty \right\}.$$

We need to identify the above pointwise random values $f_{W,V,g}$ as values of something in a space as functions of g. Using the linearity in g, events

$$\lambda_{W,V} \ :\ g \mapsto \sum_{k=1}^{\infty} (w_k, g)_{L_2(\Omega)} v_k = f_{W,V,g}$$

define a generalized random linear functional λ_W. Its events are in $H^{-\beta}$, if

$$\|\lambda_{W,V}\|_{H^{-\beta}}^2 = \sum_{n\in\mathbb{N}} \lambda_{W,V}(w_n)^2 n^{-2\beta} = \sum_{n\in\mathbb{N}} v_n^2 n^{-2\beta} < \infty,$$

and this holds almost surely if $\beta > 1/2$, but almost surely not if $\beta \leq 1/2$. The $L_2(\Omega)$ version of Theorem 5 then is

Theorem 9. *Orthonormal systems in $L_2(\Omega)$ lead to generalized random functionals that are in $H^{-\beta}$ if $\beta > 1/2$, but almost surely not if $\beta \leq 1/2$.* □

We can define operators D^γ on sequence spaces that map sequences $\{c_n\}_{n\in\mathbb{N}}$ to sequences $\{c_n n^\gamma\}_{n\in\mathbb{N}}$. Applied to expansions in the Sobolev scale (16), this maps H^β isometrically to $H^{\beta-\gamma}$, and therefore the inverse $D^{-\gamma}$ can be called a *smoothing* map that takes H^β isometrically to $H^{\beta+\gamma}$. Applying the smoothing map $D^{-\gamma}$ to any orthonormal expansion in $L_2(\Omega)$ therefore yields

Corollary 2. *Transforming random functions based on an orthonormal expansion in $L_2(\Omega)$ by $D^{-\beta}$ leads to random functions based on an orthonormal expansions in H^β, and these are almost surely in $H^{\beta-\gamma}$ if $\gamma > 1/2$.* □

Simple examples of this are given by expansions into trigonometric series, where D is differentiation and $D(\exp(in\varphi) = in\exp(in\varphi)$, but we aim at a more general case in the next section.

6.2 Random Functions From Haar Wavelets

Consider the well-known Haar mother wavelet

$$\psi(t) = \left\{ \begin{array}{ll} 1 & 0 \le t \le 1/2 \\ -1 & 1/2 \le t \le 1 \\ 0 & \text{otherwise,} \end{array} \right\}$$

and form the univariate L_2 orthonormal system

$$\psi_{n,k}(t) = 2^n \psi(2^n t - k)$$

where n and k vary in $\mathbb{Z}$ for $L_2(\mathbb{R})$, while on $L_2[0,1]$ the indices are in $\mathbb{N}$ with $0 \le k \le 2^n$ and the function 1 must be added. We treat both cases here, ignoring the range of n and k. The transition to arbitrary space dimensions d is easy by using d-variate index vectors.

This aims at L_2 as a space where point evaluations are not continuous, and consequently the formula

$$C(x,y) = \sum_{n,k} \psi_{n,k}(x)\psi_{n,k}(y)$$

yields ∞ for $x = y$. Instead, the formula

$$(f,g)_{L_2} = \sum_{n,k} (f,\psi_{n,k})_{L_2}(\psi_{n,k},g)_{L_2} =: C(f,g) \text{ for all } f,g \in L_2$$

generates the standard kernel in L_2 in function form. Recall the bra-ket notation in Quantum Mechanics when it comes to orthonormal bases.

Using uncorrelated samples $v_{n,k}$ from mean-zero variance-one random variables $V_{n,k}$, we can formally define functions

$$p_{S,\psi} := \sum_{n,k} v_{n,k}\psi_{n,k} \tag{17}$$

whose partial sums are in L_2, but convergence is a problem. The "pointwise" approach now takes arbitrary functionals $\lambda \in L_2^*$ to define events

$$\lambda(p_{S,\psi}) = \sum_{n,k} v_{n,k}\lambda(\psi_{n,k})$$

that are events of generalized random variables. As functions of λ, they are events of a random linear functional λ_ψ. The weights n^β for defining H^β must now be replaced by functions $\beta(n,k)$, and the events are in $H^{-\beta}$ almost surely, if $\sum_{n,k}\beta^{-1}(n,k)<\infty$. Sobolev spaces arise by special choices of weights [35], and smoothing operators like in Section 6 based on coefficients can be applied to reach higher-order Sobolev spaces.

6.3 Random Functions From White Noise

Physics has the idea of *white* noise as a deterministic signal with uniform spectrum. It cannot be realized adequately in L_2. In Spatial Statistics on sets Ω, a standard definition is to sample at each point x independently from a fixed distribution $\mathscr{R}$. This can be called *Gaussian white noise* for $\mathscr{R}$ being $\mathscr{N}(0,1)$, the standardized normal distribution, and it simply is $s(x)$ for the samples at x.

This is a purely bottom-up strategy like in Section 5, and it can easily be handled using independence. After independent sampling, one has a function with well-defined values, but any deterministic conclusion based on the values is unreliable, and the function is not measurable. Furthermore, the spectrum is undefined.

To arrive at random functions starting from white noise, we can use [11] to convolve white noise f_W with sufficiently smooth functions K to get generalized random variables $(f_W * K)(x)$ with covariance

$$(K*K)(x-y) = Cov((f_W*K)(x),(f_W*K)(y)) \text{ for all } x,y\in\mathbb{R}^d.$$

To arrive at a kernel C, one has to take the convolution $C = K*K$. This does not imply that the convolution $f_W * K$ is in $\mathscr{H}(C)$ as a function, but the random field has correct covariance, and this paper applies. Functions in $\mathscr{H}(C)$ have the form $\sqrt[*]{C}*g$ with L_2 functions g and the convolution square root of C, and this matches perfectly, because $f_W \notin L_2$ almost surely. This is a classical way to generate random functions and random fields [1, 13]. It falls into the scope of this paper via Theorem 8.

However, now the calculations are different, and one has to work with convolution kernels. In case of global Sobolev spaces on $\mathbb{R}^d$, a smoothing operator is defined by convolution with Matérn kernels. In particular, in the notation of [24], convolution in $\mathbb{R}^d$ with the Matérn kernel $\mathscr{M}_{\gamma-d/2,1}$ takes $H^\nu(\mathbb{R}^d)$ into $H^{\nu+\gamma}(\mathbb{R}^d)$. We can denote this smoothing operator by $\Delta^{-\gamma}$ by Fourier transform arguments, and its action can be seen as solving an equation $\Delta^\gamma u = f$ for a function $f\in H^\nu$ to get a function in $H^{\nu+\gamma}$. On bounded domains, boundary conditions interfere, but we omit details. This observation is used in various papers in Statistical Sciences starting from [16], solving the equation as a stochastic partial differential equation.

6.4 Random Functions From Finite Elements

For illustration of the above arguments, we consider a convenient computational technique by coarsening the resolution of the noise, and it leads back to an expansion in

$L_2(\Omega)$. Take a random function (1) where each w_k is one on its support W_k, and all supports are disjoint. This is an expansion into piecewise constant finite elements, and it approximates white noise by coarsening the sample. The covariances are

$$Cov(f_{W,V}(x), f_{W,V}(y)) = \sum_k w_k(x) w_k(y)$$

like in Section 2.1. This kernel is zero unless x and y are in a common W_k, and in that case the covariance is one. We get an orthonormal expansion in L_2 if we change to

$$f_{W,V,Vol}(x) := \sum_k v_k (vol(W_k))^{-1/2} w_k(x)$$

because

$$(w_n, w_m)_2 = \delta_{mn} vol(W_n).$$

Let us go on with the original approximation and see what happens after applying a smoothing operator. We could take the operator $D^{-\gamma}$ based on expansion coefficients, but since we already have global functions, we can convolve with a smooth function K. The result is

$$(K * f_{W,V})(x) = \int_{\mathbb{R}} K(x,y) f_{W,V}(y) dy = \sum_k v_k \int_{W_k} K(x,y) dy = \sum_k v_k (K * w_k)(x),$$

which is an expansion into the K-convolved original basis. This approach works similarly when starting from classical finite element bases, but the resulting kernel will not be piecewise diagonal.

Convolutions of kernels with piecewise constant functions w_k generating $L_2(\mathbb{R}^d)$ can be calculated symbolically by MAPLE or other systems, and we shall use this in the next section.

7 Numerical Examples

We start with the P-greedy Newton basis algorithm applied to Matérn kernels with Gaussian coefficients. The domain is the unit circle with the lower left quadrant missing. The point selection for a case with $m = 6$ in 2D is in Figure 1 together with one of the random functions. The P-greedy method selected 266 out of 963 points for a residual variance tolerance of 1.e-12. Figure 2 shows the convergence behaviour for increasing n, for $m = 1.5,\ 2.5,\ 4$, and 6. The observed convergence rates were 0.51, 1.54, 3.19 and 6.85 (green lines fitting the observations), while the red lines are for $m = 1.5,\ 2.5, 4$, and 6 to compare with the rates $2m/d - 1 = m - 1$ mentioned in (15). The agreement is good, but the high-order cases may not be carried out up to their final asymptotics.

Haar wavelets were implemented in $L_2[0,1]$, extended by zero to $\mathbb{R}$ and then convolved with Matérn kernels generating $W_2^m(\mathbb{R})$. These functions can be calculated explicitly as

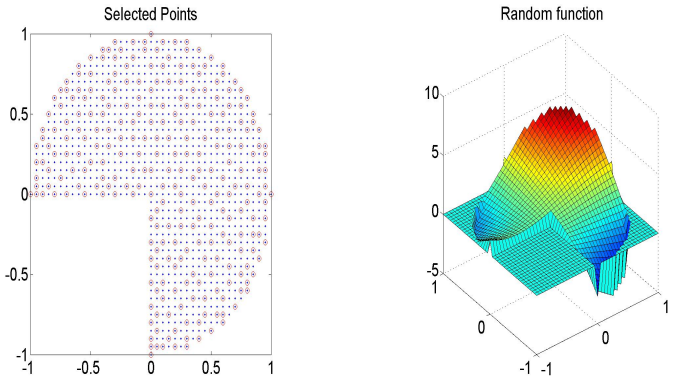

Figure 1: Selected points and a single random function

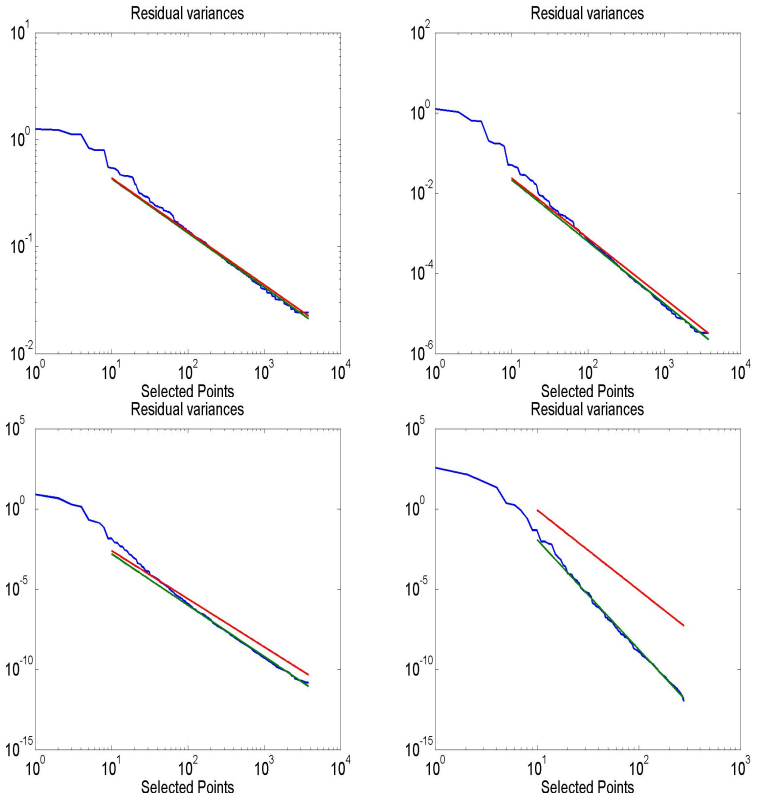

Figure 2: Convergence of residual variances for increasing numbers of selected points, for 2D Matérn kernels with $m = 1.5, 2.5, 3$, and 5.

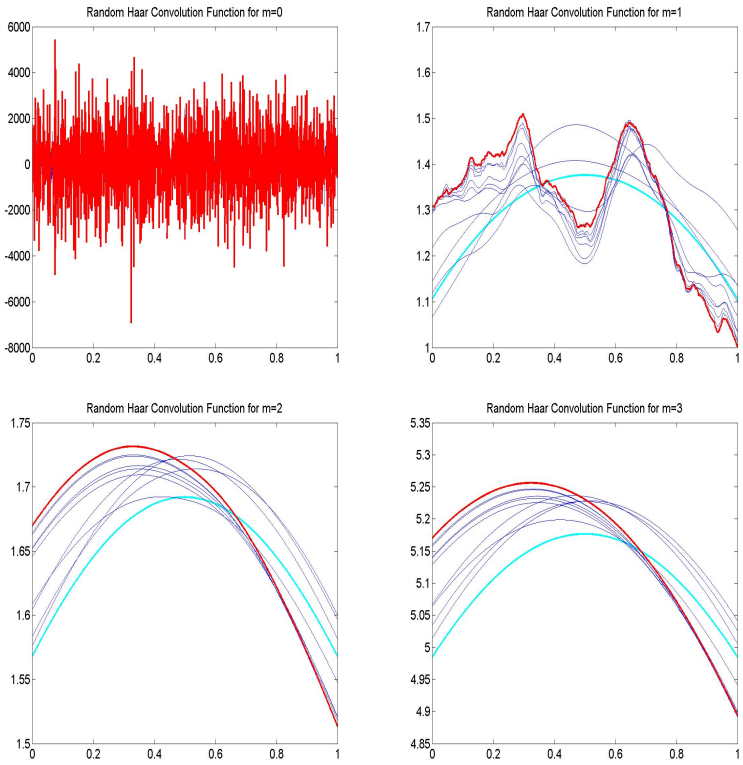

Figure 3: Random Functions based on Haar wavelets on $[0,1]$, smoothed for Sobolev order $m = 0,1,2,3$.

splines consisting of exponentials and polynomials. Starting from "Haar white noise" (17) in the top left of Figure 3, smoothings for $m = 1,2,3$ via

$$\mathscr{M}_{m-1/2,1} * p_{S,\psi} = \sum_{n,k} v_{n,k} (\mathscr{M}_{m-1/2,1} * \psi_{n,k})$$

follow in reading order, using the same random Gaussian coefficients. The final random function is in red, the starting partial sum is in cyan, followed by partial sums for $n = 1,2,\ldots,9$. The $m = 1$ case illustrates the smoothness difference between partial sums and the limit that has "asymptotic roughness".

8 Conclusions and Open Problems

Calculating random functions can be done by randomizing coefficients of basis functions. This can be analyzed without using advanced Probability and Statistics. It

leads to kernels and their native Hilbert spaces, and the chosen bases will be orthonormal there. Paths of random fields require much more probabilistic machinery, but turn out not to be seriously more general. Choosing Newton bases recovers the standard algorithm for sampling random fields on finite subsets via Cholesky decompositions of kernel matrices. Partial sums of these randomized expansions are in the native Hilbert space, but the limits are not. However, the variance of the truncation error can be expressed by the well-known Power Function and proven to be convergent to zero. If points for the Newton basis are selected by the P-greedy method, the convergence rate is close to optimal.

By certain modifications, the suggested randomized orthonormal expansions also work in L_2, and Haar wavelet bases are an example. Smoothing operators take such expansions into Sobolev spaces, allowing scales of spaces defined via weighted expansions. This is well-known for expansions into Mercer-Karhunen-Loéve bases and into Haar wavelets, but it should be proven for Newton expansions based on the P-greedy method.

The connection of smoothing operators to techniques based on Stochastic Partial Differential Equations (SPDEs) should be further investigated. For instance: if an SPDE is numerically solved via finite element bases, is this the same as using convolution, or at least comparable? What is the effect of boundary conditions in both cases? Smoothing operators via convolutions should be calculated explicitly, or good numerical approximations thereof. If these are based on finitely many points, they will possibly be connected to the SPDE methods.

The randomized expansions of this paper provide plenty of algorithms for the implementation of random function generators. But what are their properties? Given a measurable set $M \subset \Omega \times \mathbb{R}$ in graph space, what is the probability that a random function has its graph in M? How does this vary with the expansion and the coefficient distributions? This has a background connection to the statement of Probability Theory that random functions are associated with measures μ on sets A of function spaces via

$$\mu(A) = Prob\{f \in A, \text{ for all events } f\}.$$

These probabilities are a property of the random function generator rather than a property of random functions, and it is interesting to see how random function generators differ.

The standard algorithm using the Newton basis (or Cholesky decompositions, equivalently) is computationally still quite expensive if the smoothness of the kernel C is low and users want an extremely small variance of the truncation error. The probabilistic literature [7, 16, 26, 37] has various algorithms for speed-up, mainly transitions to a Markovian process or to sparse approximate inverses (precision matrices) of kernel matrices. From the viewpoint of Numerical Analysis, this is a form of *localization* that changes the kernel and deserves further investigations that will be useful also for localized deterministic settings [29].

We omit techniques based on Random Fourier Features [17, 25]. These introduce randomness in frequence space and deserve a comparison with the methods proposed here.

References

[1] Adler, R. and Taylor, J. (2007). *Random Fields and Geometry*. Springer.

[2] Beck, J. (2009). *Inevitable Randomness in Discrete Mathematics*, volume 49 of *University Lecture Series*. American Mathematical Society.

[3] Billingsley, P. (2012). *Probability and Measure, Third Edition*. Wiley.

[4] Buhmann, M. (2003). *Radial Basis Functions, Theory and Implementations*. Cambridge University Press, Cambridge,UK.

[5] Cockayne, J., Oates, C. J., Sullivan, T. J., and Girolami, M. (2019). Bayesian probabilistic numerical methods. *SIAM Review*, 61(4):756–789.

[6] Cristianini, N. and Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press, Cambridge.

[7] Datta, A., Banerjee, S., Finley, A., and Gelfand, A. (2016). Hierarchical Nearest-Neighbor Gaussian Process Models for Large Geostatistical Datasets. *Journal of the American Statistical Association*, 111(514):800–812.

[8] De Marchi, S. and Schaback, R. (2009). Nonstandard kernels and their applications. *Dolomites Research Notes on Approximations*, 2:16–43.

[9] De Marchi, S., Schaback, R., and Wendland, H. (2005). Near-optimal data-independent point locations for radial basis function interpolation. *Adv. Comput. Math.*, 23(3):317–330.

[10] Fasshauer, G. and McCourt, M. (2015). *Kernel-based Approximation Methods using MATLAB*, volume 19 of *Interdisciplinary Mathematical Sciences*. World Scientific, Singapore.

[11] Higdon, D. (2002). Space and space-time modeling using process convolutions. In Anderson, C. W., Barnett, V., Chatwin, P. C., and El-Shaarawi, A. H., editors, *Quantitative Methods for Current Environmental Issues*, pages 37–56. Springer London.

[12] Kallenberg, O. (2017). *Random Measures, Theory and Applications*, volume 77 of *Probability Theory and Stochastic Modelling*. Springer.

[13] Kallenberg, O. (2021). *Foundations of Modern Probability*, volume 99 of *Probability Theory and Stochastic Modelling*. Springer, 3rd edition.

[14] Korte-Stapff, M., Karvonen, T., and Moulines, E. (2025). Smoothness estimation for Whittle-MatÃ©rn processes on closed Riemannian manifolds. *Stochastic Processes and their Applications*, 189.

[15] Lindgren, F., Bolin, D., and Rue, H. (2022). The SPDE approach for Gaussian and non-Gaussian fields: 10 years and still running. *Spatial Statistics*, page 100599.

[16] Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B*, 73:423–498.

[17] Liu, F., Huang, X., Chen, Y., and Suykens, J. A. K. (2022). Random features for kernel approximation: A survey on algorithms, theory, and beyond. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):7128–7148.

[18] Loéve, M. (1948). *Fonctions aléatoires du second ordre, Supplement to P. Lévy, Processus Stochas- tiques et Mouvement Brownien*. Gauthier-Villars.

[19] Lukić, M. and Beder, J. (2001). Reproducing Kernel Hilbert Spaces and Paths of Stochastic Processes. *Trans. Am. Math. Soc.*, 353:3945–3969.

[20] Mercer, J. (1909). Functions of positive and negative type and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society A*, 209:441–458.

[21] Müller, S. and Schaback, R. (2009). A Newton basis for kernel spaces. *Journal of Approximation Theory*, 161:645–655.

[22] Oates, C. and Sullivan, T. (2019). A modern retrospective on probabilistic numerics. *Stat. Comput.*, 29:1335–1351.

[23] Pazouki, M. and Schaback, R. (2011). Bases for kernel-based spaces. *Computational and Applied Mathematics*, 236:575–588.

[24] Porcu, E., Bevilacqua, M., Alegria, A., Oates, C., and Schaback, R. (2024). The Matérn model: A journey through statistics, numerical analysis and machine learning. *Statist. Sci.*, 39(3):469–492.

[25] Rahimi, A. and Recht, B. (2007). Random features for large-scale kernel machines. *Advances in Neural Information Processing Systems*, 20:1177–1184.

[26] Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields: Theory and Applications (1st ed.)*. Chapman and Hall.

[27] Santin, G. and Haasdonk, B. (2017). Convergence rate of the data-independent *P*-greedy algorithm in kernel-based approximation. *Dolomites Res. Notes Approx.*, 10(Special Issue):68–78.

[28] Santin, G. and Schaback, R. (2016). Approximation of eigenfunctions in kernel-based spaces. *Adv. Comput. Math.*, 42(4):973–993.

[29] Schaback, R. (2025). Greedy adaptive local recovery of functions in Sobolev spaces. *Numerical Algorithms*.

[30] Schaback, R. and Wendland, H. (2006). Kernel techniques: from machine learning to meshless methods. *Acta Numerica*, 15:543–639.

[31] Scheuerer, M. (2010). Regularity of the sample paths of a general second order random field. *Stoch. Proc. Appl.*, 120:1879–1897.

[32] Schölkopf, B. and Smola, A. (2002). *Learning with Kernels*. MIT Press, Cambridge.

[33] Shawe-Taylor, J. and Cristianini, N. (2004). *Kernel Methods for Pattern Analysis*. Cambridge University Press, Cambridge,UK.

[34] Steinwart, I. (2019). Convergence types and rates in generic Karhunen-Loéve expansions with applications to sample path properties. *Potential Anal.*, 51:361–395.

[35] Triebel, H. (2013). Characterizations of some function spaces in terms of Haar wavelets. *Commentationes Mathematicae*, 53:35–53.

[36] Uspensky, V. and Semenov, A. (1993). *Algorithms: main ideas and applications*. Kluwer Academic Publishers, USA.

[37] Vecchia, A. V. (1988). Estimation and Model Identification for Continuous Spatial Processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 297–312.

[38] Wendland, H. (2005). *Scattered Data Approximation*. Cambridge University Press, Cambridge,UK.

[39] Winkle, D., Steinwart, I., and Haasdonk, B. (2025). Convergence rates for realizations of Gaussian random variables. arxiv 2508.13940.