A class of random functions in non-standard smoothness spaces

S. Dahlke, N. Döhring, S. Kinzel

Preprint 169
DFG-Schwerpunktprogramm 1324
„Extraktion quantifizierbarer Information aus komplexen Systemen”

A class of random functions in non-standard smoothness spaces

S. Dahlke, N. Döhring, S. Kinzel

Preprint 169
The consecutive numbering of the publications is determined by their chronological order.

The aim of this preprint series is to make new research rapidly available for scientific discussion. Therefore, the responsibility for the contents is solely due to the authors. The publications will be distributed by the authors.
A class of random functions in non-standard smoothness spaces

S. Dahlke, N. Döring, S. Kinzel

September 11, 2014

Abstract This paper is concerned with the construction of random functions on bounded domains which possess a well-defined, prescribed smoothness in specific function spaces. In particular, we consider Besov spaces for arbitrary smoothness parameters, anisotropic Besov spaces and tensor spaces with mixed smoothness, respectively. The random functions are designed by means of wavelet expansions with random coefficients. The proofs heavily rely on the equivalence of the different smoothness norms with weighted sequence norms of wavelet expansion coefficients.

Mathematics Subject Classification 2010: 42B35, 42C40, 46A32, 46B28, 46E39, 60H25, 60G60, 65T60.

Key words: Random fields, random wavelet expansions, random tensor expansions, Besov spaces, anisotropic Besov spaces, mixed smoothness spaces, wavelets.

1 Introduction

In recent years, the relations of stochastic analysis and the theory of function spaces has become a field of increasing interest. As an example, the regularity of the solutions to SDEs and SPDEs in certain function spaces has been intensively studied, see, e.g., [9, 34, 35, 37, 38, 51, 52]. Although these studies are of interest on their own, very often there is a concrete practical motivation in the background. Indeed, when it comes to numerical approximations of the objects of interest, e.g., the solutions of SDEs/SPDEs, then the approximation order that can be achieved usually depends on the membership of the objects under consideration in specific scales of smoothness spaces. As an example, let us mention approximation schemes based on wavelets. Wavelets are Riesz bases for $L_2$ that can be derived by scaling, dilating and translating a finite set of functions, the so-called mother wavelets. We refer to [10, 21, 40, 41] for more detailed information. It is well-known that the approximation order of linear (uniform) wavelet algorithms depends on the Sobolev smoothness of the underlying object, whereas the approximation order of nonlinear algorithms such as best $n$-term approximation in $L_2$ depends on the regularity in the specific scale

$$B^s_r(L_r), \quad \text{where} \quad \frac{1}{r} = \frac{s}{d} + \frac{1}{2}, \quad s > 0,$$

(1)

*This work has been supported by Deutsche Forschungsgemeinschaft (DFG, grants DA 360/13-2, RI 599/4-2)
of Besov spaces. We refer to [16, 23, 24] and the references therein for further information. These relationships are a consequence of the fact that wavelets are able to characterize smoothness spaces such as Besov and Sobolev spaces, respectively, in the sense that the corresponding smoothness norms are equivalent to weighted sequence norms of wavelet expansion coefficients, see, e.g., [24, 27, 41, 45, 49].

Motivated by these observations, in the last years some effort has been spent to create stochastic processes whose realizations possess, almost everywhere, a prescribed regularity in Besov or Sobolev spaces, see, e.g., [1–3, 8, 11, 13, 15, 36]. Very often, one major tool has been exactly the wavelet characterization of smoothness spaces. Based on a fixed wavelet basis, random coefficients have been designed which, by means of the norm equivalences, guarantee the desired regularity. In, for instance, the paper [2], the random wavelet coefficients $w_{j,k}$ have been modelled as an independent mixture of Bernoulli distributions $Y_j$ and standard normal distributions $Z_j$:

$$w_{j,k} \sim (1 - \pi_j)Y_j + \pi_j \tau_j Z_j,$$

where $\pi_j \in [0, 1]$ and $\tau_j > 0$. In particular, $\tau_j^2 = 2^{-\alpha j}C_1$ and $\pi_j = \min(1, 2^{-\beta j}C_2)$ have been studied, and it has been investigated how the parameters $\alpha$ and $\beta$ have to be tuned to yield a certain prescribed Besov smoothness. However, in [2] only the one-dimensional setting, smoothness parameters $s > 0$, and integrability parameters $p, q > 1$ have been considered. Quite recently, this analysis has been widely generalized in [8] and [3]. In [8], the multivariate setting and regularity estimates in Besov spaces related to arbitrary parameters $p, q > 0$ have been studied. That is, in that manuscript also the case of quasi-Banach spaces corresponding to $p < 1$ has been included. Moreover, the power of approximation schemes based on these random wavelet expansions has been studied. The very recent paper [3] considers more general parametrizations and focuses on Bayesian nonparametric wavelet regression.

This paper is concerned with further generalizations of the analysis presented in [2, 8] in the following directions. First of all, we discuss stochastic fields with realizations in Besov spaces with negative smoothness. These spaces naturally arise, e.g., in the context of elliptic operator equations with random right-hand sides such as the stochastic Poisson equation

$$-\Delta U = X \quad \text{in} \quad D,$$

$$U = 0 \quad \text{on} \quad \partial D. \quad (2)$$

In this case, the Laplacian is a bounded operator from $H^1_0$ onto $H^{-1}$, the normed dual of $H_0^1$. Therefore it is natural to consider random right-hand sides with negative smoothness. Numerical studies which substantiate this observation have already been carried out in [8]. Moreover, stochastic elliptic equations of the form (2) have necessarily to be treated whenever a stochastic evolution equation is discretized by means of a Rothe scheme, see, e.g., [4, 7, 14, 22, 30, 33, 43]. Therefore, in Section 3 of this paper, we state the conditions under which negative smoothness can be achieved. Moreover, the approximation results of [8] are also generalized to this case as far as possible.

The main results of this paper are contained in Sections 4 and 5. In Section 4, we discuss additional important classes of smoothness spaces, namely the anisotropic Sobolev and Besov spaces. Once again, to design random functions in these spaces is of independent interest, but nevertheless, we are convinced that there are a lot of
possible applications. As an example, let us mention certain elliptic equations with random coefficients as they occur, e.g., in the modelling of groundwater flow problems [25,26,50]. Usually, the random coefficients are modelled by highly isotropic lognormal distributions. However, due to certain anisotropic features that might show up in the physical environment, it could be more appropriate to use stochastic models that reflect these kinds of anisotropies. Therefore, we derive stochastic fields with prescribed smoothness in anisotropic smoothness spaces. The major tool is again the wavelet characterization of these spaces as derived, e.g., in [28,29].

Finally, in Section 5, we construct new bases of stochastic tensor wavelets. To our best knowledge, these kinds of random fields have not been considered before. Tensor wavelets are in a certain sense the wavelet version of the sparse grid approach, see, e.g., [5] for a detailed discussion on sparse grids. They are very important for the following reason: Similar to sparse grids, (adaptive) approximation schemes based on these wavelets can give rise to dimension-independent convergence rates, see [46]. In this sense, tensor wavelets provide a way to break the famous curse of dimensionality. The spaces that can be characterized by tensor wavelets are generalized dominated mixed smoothness spaces, see Section 2 for details. Therefore, in Section 5, we derive stochastic fields with prescribed regularity in these spaces.

We denote by \( f(\alpha) \preceq g(\alpha) \) that there exists some constant \( c_1 > 0 \), which is independent of \( \alpha \), such that \( f(\alpha) \leq c_1 g(\alpha) \). Analogously, we write \( f(\alpha) \succeq g(\alpha) \) if there is a constant \( c_2 > 0 \) such that \( f(\alpha) \geq c_2 g(\alpha) \). Clearly, \( f \asymp g \) means that \( f \preceq g \) and \( g \preceq f \).

## 2 Function spaces

In this section, we briefly introduce the function spaces in which we consider the random functions. To this end, let \( D \subseteq \mathbb{R}^d \) be a Lipschitz domain and let \( L_p(D) \) with \( 0 < p \leq \infty \) be the standard real-valued Lebesgue space, which is quasi-normed by

\[
\| f \|_{L_p(D)} := \left( \int_D |f(x)|^p \, dx \right)^{1/p}
\]

and with the usual modifications if \( p = \infty \).

### 2.1 The Besov spaces \( B^s_q(L_p(D)) \)

For a function \( f : D \to \mathbb{R} \) let

\[
\Delta^k_h f(x) := \prod_{i=0}^k \mathbb{1}_{D}(x + ih) \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} f(x + jh), \quad x \in \mathbb{R}^d,
\]

be the \( k \)-th difference, \( k \in \mathbb{N} \), of \( f \) with step size \( h \in \mathbb{R}^d \). For \( p \in (0, \infty] \) the modulus of smoothness is defined by

\[
\omega^k(t, f)_p := \sup_{|h| < t} \| \Delta^k_h f \|_{L_p(D)}, \quad t > 0.
\]

Now, for \( 0 < p, q \leq \infty \) and \( d(1/p - 1)^+ < s < \infty \), as well as \( k > s \), the Besov space \( B^s_q(L_p(D)) \) consists of all \( f \in L_p(D) \), such that

\[
\| f \|_{B^s_q(L_p(D))} := \| f \|_{L_p(D)} + \left( \int_0^\infty \left( t^{-s} \omega^k(t, f)_p \right)^q \frac{dt}{t} \right)^{1/q} < \infty
\]
(with the usual modifications if \( q = \infty \)). Additionally, with \( 1 \leq p, p', q, q' < \infty \), \( 1/p + 1/p' = 1/q + 1/q' = 1 \), we set

\[
B^s_q(L_p(D)) := \left( B^{-s}_{q'}(L_{p'}(D)) \right)' \quad \text{for} \quad s < 0.
\]

In this way, we have defined the Besov spaces for parameters located within the hatched area shown in Figure 1, where each point represents a Besov space defined by the parameters in the diagram (at which the fine tuning parameter \( q \) is omitted). We could have defined Besov spaces also for \( 0 < s < d(1/p - 1)_+ \) and \( 0 < p < 1 \), but they would not allow a characterization by a wavelet basis satisfying Assumption 3.1 in Section 3 below.

![Figure 1: Defined Besov spaces in a DeVore-Triebel diagram](image)

It is well–known that for \( p = q, p > 1 \) and \( s > 0, s \notin \mathbb{N} \) Sobolev and Besov spaces coincide, i.e., \( B^s_p(L_p(D)) = W^s(L_p(D)) \). The space \( B^s_2(L_2(D)) \), coincides for all \( s > 0 \) with the standard Sobolev space \( H^s(D) \), i.e., the space of all \( f \in L_2(D) \) for which

\[
\| f \|_{H^s(D)} := \inf \left\{ \| g \|_{H^s(\mathbb{R}^d)} : g \in H^s(\mathbb{R}^d), g|_D = f \right\} < \infty,
\]

where \( H^s(\mathbb{R}^d) := \{ f \in L_2(\mathbb{R}^d) : \mathcal{F}^{-1}(1+|\xi|^2)^{s/2}\mathcal{F}f \in L_2(\mathbb{R}^d) \} \). For more details about Besov spaces we refer to, e.g., [10, Ch. 3], [45,49], and the references therein.

### 2.2 The anisotropic Besov spaces \( B^s_{q,a}(L_p(\mathbb{R}^d)) \)

Let us now consider the anisotropic setting. First, we fix an anisotropy

\[
a = (a_1, \ldots, a_d) \in \mathbb{R}^d_+, \quad \text{with} \quad \sum_{i=1}^d \frac{1}{a_i} = d. \tag{3}
\]

Let \( \{e_1, \ldots, e_d\} \) denote the canonical basis of \( \mathbb{R}^d \). For a function \( f : \mathbb{R}^d \to \mathbb{R} \) let

\[
\Delta^k_h f(x) := (\Delta_{h_1 e_1} \circ \ldots \circ \Delta_{h_d e_d}) f(x), \quad x \in \mathbb{R}^d,
\]

be the mixed difference of order \( k = (k_1, \ldots, k_d) \in \mathbb{N}^d \) and step \( h = (h_1, \ldots, h_d) \in \mathbb{R}^d \). For \( p \in (0, \infty) \) the mixed modulus of smoothness with respect to \( a \) is defined by

\[
\omega^k_a(t, f)_p := \sup_{|k|_a < t} \| \Delta^k_h f \|_{L_p(D)}, \quad t > 0,
\]
where
\[ |h|_a := \sum_{j=1}^{d} |h_j|^a, \quad h \in \mathbb{R}^d, \]
is the \textit{anisotropic pseudo-distance} of the step \( h \) related to the anisotropy \( a \).

Now, let \( 0 < p, q \leq \infty \) and \( d(1/p - 1)_+ < s < \infty \). Furthermore, let \( s_i := sa_i \), \( i = 1, ..., d \), and \( \mathbb{N} \ni K > \max\{s_1, ..., s_d\} \). The \textit{anisotropic Besov space} \( B_{q,a}^s(L_p(\mathbb{R}^d)) \) consists of all functions \( f \in L_p(\mathbb{R}^d) \) for which
\[
\|f\|_{B_{q,a}^s(L_p(\mathbb{R}^d))} := \|f\|_{L_p(\mathbb{R}^d)} + \sum_{|\alpha|=K} \left( \int_0^\infty \left(t^{-s} \omega^k_{a}(t,f)\right)^q \frac{dt}{t} \right)^{1/q} < \infty \tag{4}
\]
(with the usual modifications if \( q = \infty \)). An \textit{anisotropic Besov space on a domain} \( B_{q,a}^s(L_p(D)) \) is defined by restriction, i.e.,
\[
B_{q,a}^s(L_p(D)) := \left\{ f \in L_p(D) : \exists g \in B_{q,a}^s(L_p(\mathbb{R}^d)), g|_D = f \right\}
\]
with norm
\[
\|f\|_{B_{q,a}^s(L_p(D))} := \inf \left\{ \|g\|_{B_{q,a}^s(L_p(\mathbb{R}^d))} : g \in B_{q,a}^s(L_p(\mathbb{R}^d)), g|_D = f \right\}.
\]

\textbf{Remark 2.1.} The definition of anisotropic Besov spaces given by (4) is equivalent to the definition used in [29, Sec. 2], see [28, Prop. 2.2] and the references therein.

The space \( B_{2,a}^{s}(L_2(\mathbb{R}^d)) \) coincides with the standard fractional anisotropic Sobolev space
\[
H^{s,a}(\mathbb{R}^d) := \left\{ f : \mathcal{F}^{-1}(1 + |\xi|)^{sa_i/2} \mathcal{F} f \in L_2(\mathbb{R}^d), i = 1, \ldots, d \right\}.
\]
In the case \((sa_1, \ldots, sa_d) \in \mathbb{N}^d\), it coincides with the classical anisotropic Sobolev space, i.e.,
\[
H^{s,a}(\mathbb{R}^d) = \left\{ f \in L_2(\mathbb{R}^d) : \sum_{i=1}^{d} \left\| \frac{\partial^{sa_i}}{\partial x_i^{sa_i}} f \right\|_{L_2(\mathbb{R}^d)} < \infty \right\}.
\]
Also note, whenever \( a = 1 \) we are in the isotropic case. More details about anisotropic Besov spaces can be found in, e.g., [49, Ch. 5] and the references therein.

\subsection{The tensor space \( H^{t,\ell}(D) \)}

Let the domain \( D \) be an \( n \)-fold product of component domains \( D_m \subset \mathbb{R}^{d_m} \), \( m = 1, \ldots, n \), \( n \geq 2 \), with \( \sum_{m=1}^{n} d_m = d \). The tensor space \( H^{t,\ell}(D) \subset L_2(D) \) is defined as follows. Let \( H^{s}(D_m) \), \( s \geq 0 \), be the standard Sobolev space, or a closed subspace of it, in which boundary conditions are incorporated if required. Let \( t \in [0, \infty)^n \), \( \ell \in [0, \infty) \), and \( \delta_{m,i} \) be the Kronecker delta. Then \( H^{t,\ell}(D) \) consist of all functions \( f \in L_2(D) \) for which
\[
\|f\|_{H^{t,\ell}(D)} := \sum_{i=1}^{n} \prod_{m=1}^{n} \|f_i\|_{H^{t_{m} + s_{m},\ell_{m}}(D_m)}, \quad f = f_1 \otimes \cdots \otimes f_n,
\]
is finite, that is,
\[
H^{t,\ell}(D) := \bigcap_{i=1}^{n} \bigotimes_{m=1}^{n} H^{t_{m} + \delta_{m,i},\ell_{m}}(D_m).
\]
These spaces are generalizations of spaces with dominating mixed derivatives as introduced in [39], see also [31,46]. In particular, the space $H^{t,0}(D)$ is known as the Sobolev space with dominating mixed derivatives, while $H^{0,\ell}(D)$ is isomorphic to the standard Sobolev space $H^\ell(D)$.

3 A class of random functions in Besov spaces

In this section, we study a new class of random functions based on wavelet expansions and derive conditions under which such a random function (almost surely) has a certain smoothness in a given Besov space $B^s_{p,q}(L_p(D))$, where $s \in \mathbb{R}$ and $0 < p,q \leq \infty$ as defined in Subsection 2.1. After recalling the wavelet characterization of these Besov spaces, we state the stochastic model upon which we construct the random functions. The main result is then given in Theorem 3.9 which generalizes the findings in [8] to the case of negative smoothness. Additional approximation results for these random functions are also presented.

3.1 Wavelet characterization

In general a wavelet basis $\Psi := \{\psi_\lambda : \lambda \in \nabla\}$ on a domain $D$ is a basis for $L^2(D)$, where the indices $\lambda \in \nabla$ encode several types of information; most importantly the scale, denoted by $|\lambda|$, the spatial location, and the type of the wavelet. On the real line, e.g., the scale $|\lambda| = j \in \mathbb{Z}$ denotes the dyadic refinement level and $2^{-j}k$ with $k \in \mathbb{Z}$ denotes the spatial location of the wavelet. In this section, we will disregard any explicit dependence on the type of the wavelet since this only produces additional constants. Thus, we will frequently use $\lambda = (j,k)$ and $\nabla = \{(j,k) : j \geq 0, k \in \nabla_j\}$, where $|(j,k)| = j$ and $\nabla_j$ is some countable index set understood to encode the spatial location and type of the wavelets.

Assumption 3.1. We assume that the domain $D$ under consideration enables us to construct a wavelet basis $\Psi = \{\psi_{j,k}\}_{j \geq 0, k \in \nabla_j}$ with the following properties:

(i) $\Psi$ forms a Riesz basis for $L^2(D)$, i.e., there exist positive constants $c_R, C_R$, such that

$$c_R \left( \sum_{j,k} |a_{j,k}|^2 \right) \leq \left\| \sum_{j,k} a_{j,k} \psi_{j,k} \right\|_{L^2(D)}^2 \leq C_R \left( \sum_{j,k} |a_{j,k}|^2 \right)$$

holds for all $(a_{j,k})_{j,k} \in \ell_2$ and $\text{clos}(\text{span } \psi_{j,k}) = L^2(D)$.

(ii) The wavelets are compactly supported and satisfy the locality assumption

$$\text{diam}(\text{supp } \psi_{j,k}) \asymp 2^{-j}, \quad k \in \nabla_j.$$

(iii) The wavelets satisfy the cancellation property

$$|\langle v, \psi_{j,k} \rangle_{L^2(D)}| \leq 2^{-j(d/2 + \tilde{m})} \left| v \right|_{W^{\tilde{m}}(L_\infty(\text{supp } \psi_{j,k}))}$$

for $j > 0$ and some parameter $\tilde{m} \in \mathbb{N}_0$.

(iv) The cardinalities of the index sets $\nabla_j$ satisfy $\# \nabla_j \asymp 2^{jd}$. 
(v) There exist a second (dual) wavelet basis $\tilde{\Psi} = \{\tilde{\psi}_{j,k}\}_{j \geq 0, k \in \nabla_j}$ also fulfilling the assumptions (i) – (iv) that is biorthogonal to $\Psi$, i.e.,
$$
\langle \psi_{j,k}, \tilde{\psi}_{j',k'} \rangle = \delta_{j,j'} \delta_{k,k'}.
$$

(vi) The biorthogonal wavelet bases $\Psi$ and $\tilde{\Psi}$ induce a characterization of the Besov space $B^{s}_{q}(L^p(D))$ — within $-s_2 < s < s_1$, where $s_1, s_2 > 0$ are bounds which are determined by the smoothness and the approximation properties of $\Psi$ and $\tilde{\Psi}$ — of the form
$$
\| \cdot \|_{B^{s}_{q}(L^p(D))} \asymp \left( \sum_{j=0}^{\infty} 2^{jq(s+d(\frac{1}{2}-\frac{1}{p}))} \left( \sum_{k \in \nabla_j} |\langle \cdot, \tilde{\psi}_{j,k} \rangle_{L^2(D)}|^{p} \right)^{q/p} \right)^{1/q}
$$
for $d(1/p - 1)_+ < s < s_1$. For $-s_2 < s < 0$ and $1 \leq p, q < \infty$ it is of the form
$$
\| \cdot \|_{B^{s}_{q}(L^p(D))} \asymp \left( \sum_{j=0}^{\infty} 2^{jq(s+d(\frac{1}{2}-\frac{1}{p}))} \left( \sum_{k \in \nabla_j} |\langle \cdot, \psi_{j,k} \rangle_{L^2(D)}|^{p} \right)^{q/p} \right)^{1/q}.
$$

Remark 3.2. (i) Suitable constructions of wavelets on domains satisfying Assumption 3.1 can be found in, e.g., [6,12,18–20,42]. We refer to [10] for a detailed discussion.

(ii) In practical applications, e.g., in the context of elliptic boundary value problems, Sobolev and Besov spaces involving boundary conditions come into play. The most prominent example is the Sobolev space $W^1_0(L^2(D))$ which is used to describe Dirichlet boundary conditions for second order elliptic differential operators. In this case, the dual space looks slightly different since it is defined as $W^{-1}(L^2(D)) = (W^1_0(L^2(D)))'$. In many cases, it is possible to find a boundary adapted wavelet basis that characterizes $W^1_0(L^2(D))$ in the sense of Assumption 3.1 (vi), while the dual basis gives rise to similar norm equivalences for the dual spaces, see, e.g., [18, Thm. 3.4.3]. Moreover, these biorthogonal wavelet bases very often also exist for much more general boundary conditions. Once more, we refer to [18] for further information. Once such a wavelet basis is available, the analysis presented in this paper can also be generalized to this case.

3.2 The stochastic model

Let
$$
\alpha, \gamma \in \mathbb{R}, \quad \beta \in [0,1], \quad \sigma_j^2 := \begin{cases} \frac{\gamma d 2^{-\alpha jd}}{2^{-\beta jd}} : j > 0, \\ 1 : j = 0, \end{cases} \quad \text{and} \quad \rho_j := 2^{-\beta jd},
$$
and consider an independent family of random variables $(Z_{j,k}, Y_{j,k})_{j \in \mathbb{N}_0, k \in \nabla_j}$ on a probability space $(\Omega, \mathcal{F}, P)$, where
$$
Z_{j,k} \sim N(0,1), \quad \text{and} \quad P(Y_{j,k} = 1) = 1 - P(Y_{j,k} = 0) = \rho_j.
$$

Furthermore, given biorthogonal wavelet bases $\Psi$, $\tilde{\Psi}$ which satisfy Assumption 3.1, we define the random functions
$$
X := \sum_{j=0}^{\infty} \sum_{k \in \nabla_j} \sigma_j Y_{j,k} Z_{j,k} \psi_{j,k}^*,
$$
Then \( \sum \) Let \( \nu \)

Remark 3.3. Suppose \( X \in H^s(D) \) and let \( \xi, \zeta \in H^s(D) \). Furthermore, since \( H^s(D) \)

is a Hilbert space and \( \{2^{-js} \psi_{j,k}, j \geq 0, k \in \mathbb{N}_j \} \) is a Riesz basis for \( H^s(D) \), we know

that there exists an orthonormal basis \( \{e_{j,k}\}_{j,k} \) of \( H^s(D) \) with \( 2^{-js} \psi_{j,k} = \Phi e_{j,k} \) for

a bounded linear bijection \( \Phi : H^s(D) \to H^s(D) \), using Gram-Schmidt. We have

\[
E(\langle \xi, \Phi^{-1} X \rangle_{H^s(D)}) = 0, \text{ and since}
\]

we obtain for the covariance operator \( \hat{Q} \) associated with \( \Phi^{-1} X \)

\[
\hat{Q} \xi = \sum_{j=0}^{\infty} 2^{2js} \sigma_j^2 \sum_{k \in \mathbb{N}_j} \langle \xi, e_{j,k} \rangle_{H^s(D)} e_{j,k}.
\]

Before we present the main result of this section, we continue by stating some technical Lemmas which will be used in the sequel. Lemmas 3.5 – 3.7 are taken from [8]. The proofs are given in the appendix. We set

\[
S_{j,p} := \sum_{k \in \mathbb{N}_j} Y_{j,k} |Z_{j,k}|^p
\]

for an independent family of random variables \( \{Y_{j,k}, Z_{j,k}\}_{j \in \mathbb{N}_k, k \in \mathbb{N}_j} \) as defined by (6). Note, \( (S_{j,p})_{j=0}^{\infty} \) forms an independent sequence for every fixed \( 0 < p < \infty \). Also, with \( \nu_p \) denoting the \( p \)-th absolute moment of the standard normal distribution, we have

\[
E(S_{j,p}) = \# \mathbb{N}_j \rho_j \nu_p.
\]

Lemma 3.4. Let \( \{X_i\}_{i \in \mathbb{N}} \) be a family of independent, non-negative random variables. Then \( \sum_{i=1}^{\infty} X_i < \infty \), P-a.s., if and only if \( \sum_{i=1}^{\infty} E \left( \frac{X_i}{1 + X_i} \right) < \infty \).

Lemma 3.5. Let \( n \in \mathbb{N}, p \in [0,1] \) and \( X_{n,p} \sim \text{Bin}(n, p) \). For all \( n \) there exists a constant \( c = c(n) > 0 \) such that for all \( r > 0 \) and \( p \)

\[
E(X_{n,p}^r) \leq c (1 + (np)^r).
\]

Lemma 3.6. Let \( \beta \in [0,1) \). Then

\[
\lim_{j \to \infty} \frac{S_{j,p}}{\# \mathbb{N}_j \rho_j} = \nu_p
\]

holds with probability one, where \( \nu_p \) denotes the \( p \)-th absolute moment of the standard normal distribution. Further, for every \( r > 0 \)

\[
\sup_{j \geq 0} \frac{E(S_{j,p}^r)}{(\# \mathbb{N}_j \rho_j)^r} < \infty.
\]
Lemma 3.7. Let $\beta = 1$ and
\[
\lim_{j \to \infty} \#\nabla_j 2^{-jd} = C_0 \quad \text{for some} \quad C_0 > 0.
\] (11)

Let $\mu_p$ denote the distribution of $|Z_{j,k}|^p$, and let $S_p$ be a compound Poisson distributed random variable with intensity measure $C_0 \mu_p$. Then $(S_{j,p})_j$ converges in distribution to $S_p$, and for every $r > 0$
\[
\sup_{j \geq 0} \mathbb{E} (S_{j,p}^r) < \infty.
\]

Remark 3.8. In general, we only have Assumption 3.1 (iv) instead of (11). For $\beta = 1$ the upper bound (10) remains valid in the general case, too. In all known constructions of wavelet bases on bounded domains, see, e.g., [6, 12, 18–20, 42], the number $\#\nabla_j$ of wavelets per level $j > 0$ is a constant multiple of $2^{jd}$. For those kinds of bases, (11) trivially holds.

Now, the following theorem states the conditions on the parameters $\alpha, \beta, \gamma$ in (5) of the stochastic model which guarantee that a random function, defined by (7), almost surely is contained in a given Besov space $B^s_q(L_p(D))$. The case $s > d(1/p - 1)_+$ has been studied in [8, Theorem 6]. Here, we generalize this result to negative values of $s$.

Theorem 3.9. Let Assumption 3.1 hold, and let $X$ be a random function as defined in (7) with respect to the dual basis $\tilde{\Psi}$. Then $X$ is $\mathbb{P}$-almost surely contained in $B^s_q(L_p(D))$ with $s < 0$, $1 \leq p, q < \infty$, if and only if
\[
s < d \left( \frac{\alpha - 1}{2} + \frac{\beta}{p} \right)
\] (12)
or
\[
s \leq d \left( \frac{\alpha - 1}{2} + \frac{\beta}{p} \right) \quad \text{and} \quad q \gamma d < -2.
\] (13)

In both cases
\[
\mathbb{E} \|X\|_{B^s_q(L_p(D))}^q < \infty.
\] (14)

Proof. Using Assumption 3.1 (vi) we have $X \in B^s_q(L_p(D))$ with $s < 0$ $\mathbb{P}$-almost surely if and only if
\[
\|X\|_{B^s_q(L_p(D))}^q = \sum_{j=0}^{\infty} 2^{j(s+d(1/2-1/p))q} \sigma_j^q S_{j,p}^{q/p} < \infty, \quad \text{P-a.s.}
\]

Thus, using the abbreviation $a_j := 2^{j(s+d(1/2-1/p))q} \sigma_j^q$, we have to show when
\[
\sum_{j=0}^{\infty} a_j S_{j,p}^{q/p} < \infty, \quad \text{P-a.s.} \tag{15}
\]

It is enough to show that (15) is equivalent to
\[
\sum_{j=0}^{\infty} a_j (\#\nabla_j \rho_j)^{q/p} < \infty, \tag{16}
\]
because inserting Assumption 3.1 (iv) and (5) into (16) yields
\[ \sum_{j=0}^{\infty} a_j (\#\nabla_j \rho_j)^{q/p} \sum_{j=0}^{\infty} j^{q\gamma d/2} 2^{j\gamma d(s/d-(\alpha-1)/2-\beta/p)}, \]
and we see that (16) holds if and only if the conditions (12) or (13) are satisfied.

We continue to show the equivalence of (15) and (16). In the case \( 0 \leq \beta < 1 \) it follows from Lemma 3.6. In the case \( \beta = 1 \) observe that (16) with Assumption 3.1 (iv) reduces to
\[ \sum_{j=0}^{\infty} a_j < \infty, \] (17)
while (15) is, due to Lemma 3.4, equivalent to
\[ \sum_{j=0}^{\infty} E \left( \frac{a_j S_{j,p}^{q/p}}{1 + a_j S_{j,p}^{q/p}} \right) < \infty. \] (18)

The equivalence of (17) and (18) is shown in two parts. To show that (17) implies (18), we use Lemma 3.7 to conclude
\[ \sum_{j=0}^{\infty} E \left( \frac{a_j S_{j,p}^{q/p}}{1 + a_j S_{j,p}^{q/p}} \right) \leq \sum_{j=0}^{\infty} a_j E(S_{j,p}^{q/p}) < \infty, \] if \( \sum_{j=0}^{\infty} a_j < \infty. \)

The second part, i.e., (18) implies (17), is shown by contradiction. We assume (18) and \( \sum_{j=0}^{\infty} a_j = \infty \) to hold. Now, by (11) we obtain that \( c_p := \inf_{j \geq 0} P(S_{j,p} \geq 1) > 0, \) and, using the specific form of \( a_j, \) we can conclude
\[ \sum_{j=0}^{\infty} E \left( \frac{a_j S_{j,p}^{q/p}}{1 + a_j S_{j,p}^{q/p}} \right) \geq c_p \sum_{j=0}^{\infty} \frac{a_j}{1 + a_j} = \infty, \]
which contradicts the assumption (18). All together, the equivalence of (15) and (16) is proven.

It remains to show (14). We use the norm equivalence of Assumption 3.1 (vi), Lemma 3.6, Lemma 3.7, and (16) to derive
\[ E\|X\|_{B^s_q(L_p(D))}^q \leq \sum_{j=0}^{\infty} a_j E(S_{j,p}^{q/p}) \leq \sum_{j=0}^{\infty} a_j (\#\nabla_j \rho_j)^{q/p} < \infty. \]

\[ \square \]

**Remark 3.10.** (i) Combining Theorem 3.9 with [8, Theorem 6], we have that \( X \) which is defined by (7) is \( P \)-almost surely contained in \( B^s_q(L_p(D)) \) with \( s \in \mathbb{R} \setminus 0 \) and \( 0 < p, q < \infty, \) where \( p, q \geq 1 \) if \( s < 0, \) if and only if
\[ s < d \left( \frac{\alpha - 1}{2} + \frac{\beta}{p} \right) \] (19)
or
\[ s \leq d \left( \frac{\alpha - 1}{2} + \frac{\beta}{p} \right) \quad \text{and} \quad q \gamma d < -2. \] (20)
Furthermore, in both cases we have

$$E \|X\|^q_{B^s_q(L_p(D))} < \infty.$$  \hfill (21)

(ii) In the case $X \in H^s(D)$ one can compute the moment (21) directly, i.e.,

$$E \|X\|^2_{H^s(D)} = \sum_{j=0}^{\infty} 2^{j\nu} \sigma_j^2 \mathbb{E} \left( S_j, 2^{-j\beta} \right) \\ = \sum_{j=0}^{\infty} 2^{j\nu} \sigma_j^2 \# \nabla_j 2^{-j\beta} \asymp \sum_{j=0}^{\infty} 2^{-j(\alpha + \beta - 1 + 2s/d)} j^\gamma d.$$

As a special case of Theorem 3.9 we emphasize the regularity of $X$ in $B^s_p(L_p(D))$, where

$$s^* := \frac{d}{1 - \beta} \left( \frac{\alpha - 1}{2} + \frac{\beta}{p} \right) - \frac{\beta \nu}{1 - \beta}.$$  \hfill (22)

The reason is, that this scale (22) generalizes the $L_2$-adaptivity scale (1). The regularity of a function in the scale (1) determines the approximation order that can be achieved by nonlinear wavelet approximation in $L_2$, whereas the regularity in the scale (22) is related with nonlinear wavelet approximation in $B^s_p(L_p(D))$. Results of this type also hold for negative $\nu$, for details see, e.g., [17]. This context is visualized in Figure 2.

**Corollary 3.11.** Let $\beta \in [0, 1)$, $p \in (1, \infty)$, and $-d/p \leq \nu < d((\alpha - 1)/2 + \beta/p)$. Then

$$X \in B^s_p(L_p(D)) \text{ for all } s < s^*$$

in the scale (22), where

$$s^* := \frac{d}{1 - \beta} \left( \frac{\alpha - 1}{2} + \frac{\beta}{p} \right) - \frac{\beta \nu}{1 - \beta}.$$  \hfill (23)

**Proof.** Using Remark 3.10 (i) we have to show that

$$s^* = d \left( \frac{\alpha - 1}{2} + \frac{\beta}{\tau^*} \right).$$  \hfill (24)

Inserting (22) with $s = s^*$ and $\tau = \tau^*$ into (23), which is $s^* = d((\alpha - 1)/2 + \beta/p) + \beta(s^* - \nu)$, yields the claim. \hfill \qed

![DeVore-Triebel diagram](image-url)
Remark 3.12. (i) Corollary 3.11 implies that by choosing $\beta$ closer to one, an arbitrary high regularity in the adaptivity scale (22) can be achieved, provided that the underlying wavelet basis is sufficiently smooth.

(ii) The lower bound on $\nu$ in Corollary 3.11 is caused by the fact that we must ensure the fundamental condition $d(1/\tau - 1)_+ < s$, which guarantees the existence of a wavelet characterization, see again Figure 2 for an illustration.

3.3 Linear and nonlinear approximation results

In this subsection we state error bounds for linear and nonlinear approximation schemes for random functions $X : \Omega \to B_p^{\nu+m}(L_p(D))$ with respect to weaker Besov norms $B_p^{\nu+m}(L_p(D))$, i.e., $m > 0$.

We define the linear approximation error of $X$ with respect to $B_p^{\nu+m}(L_p(D))$ by

$$ e_{N,p,\nu}(X) := \inf \left( E \| X - \hat{X} \|_{B_p^{\nu+m}(L_p(D))}^p \right)^{1/p} $$

with the infimum taken over all measurable mappings $\hat{X}$ such that

$$ \dim(\text{span}(\hat{X}(D))) \leq N, $$

where $N \in \mathbb{N}_0$ can be understood as the number of degrees of freedom in this context.

Theorem 3.13. Let $\beta \in [0, 1)$, $p \in (1, \infty)$, and $m > 0$. For a fixed approximation space $B_p^{\nu+m}(L_p(D))$, let $X$ be given by (7) with

$$ \nu + m < d((\alpha - 1)/2 + \beta/p) =: \nu + m^*, \quad (25) $$

i.e., $X \in B_p^{\nu+m}(L_p(D))$ for all $m < m^*$. The linear approximation error with respect to $B_p^{\nu+m}(L_p(D))$ satisfies

$$ e_{N,p,\nu}(X) \lesssim (\log_2 N)^{\frac{2d}{\nu}} N^{-(\frac{n-1+\frac{2}{p}-\frac{\nu}{d}}{2})}. \quad (26) $$

Proof. As a specific linear approximation, let us consider a uniform approximation of the form

$$ X_{j_1} := \sum_{j=0}^{j_1} \sum_{k \in \nabla_j} \sigma_j Y_{j,k} Z_{j,k} \psi_{j,k} $$

for some $j_1 > 0$, where in particular $N \asymp 2^{j_1 d}$. With $S_{j,p}$ as defined in (8) and with (9), we obtain

$$ E \| X - X_{j_1} \|_{B_p^{\nu+m}(L_p(D))}^p \asymp E \left( \sum_{j=j_1+1}^{\infty} 2^{jp((\nu+m^*)+d(\frac{1}{2}-\frac{1}{\rho}))} \sigma_j^p S_{j,p} \right) $$

$$ = E \left( \sum_{j=j_1+1}^{\infty} 2^{jp((\nu+m^*)+d(\frac{1}{2}-\frac{1}{\rho}))} 2^{-jpm^*} \sigma_j^p S_{j,p} \right) $$

$$ = \sum_{j=j_1+1}^{\infty} 2^{jp((\nu+m^*)+d(\frac{1}{2}-\frac{1}{\rho}))} 2^{-jpm^*} \sigma_j^p \| \nabla_j \rho_j \|. $
Theorem 3.15. Let a lower bound for the linear approximation error can be derived of smoothness spaces. From (27), we observe that, similar to the well-known deterministic setting, see, e.g., \([16, 23]\), the approximation order which can be achieved by (uniform) linear schemes depends on the regularity of the object under consideration in the same scale.

Proof. Given \((e_{j,k})_{j,k}\), we get

\[
\begin{align*}
\mathbb{E}\|X - X_{j_1}\|_{B^p_p(L_p(D))}^p &= \sum_{j=j_1+1}^{\infty} 2^{jp(\frac{\alpha - 1}{2} + \frac{\beta}{d} - \frac{\beta}{2} - \frac{1}{2})} 2^{-jpm} \beta^{\frac{j}{2}} 2^{-\frac{jdp}{p} - \frac{2jd}{d} - \frac{2jd}{d}} \\
&= \sum_{j=j_1+1}^{\infty} j^{\frac{dp}{p}} 2^{-jpm} \times \left(j^1 \times 2^{-jpm} \times (\log_2 N)^{\frac{dp}{p}} N^{-p(\frac{\alpha - 1}{2} + \frac{\beta}{d} - \frac{\beta}{2})}\right),
\end{align*}
\]

which yields (26).

\(\square\)

**Remark 3.14.** In the setting of Theorem 3.13, with a slightly coarser error estimation, for all \(m < m^*\) we get

\[
\begin{align*}
\mathbb{E}\|X - X_{j_1}\|_{B^p_p(L_p(D))}^p &\leq 2^{-j_1pm} \mathbb{E}\left(\sum_{j=0}^{\infty} 2^{jp(\nu + m + d(\frac{\beta}{d} - \frac{1}{2}) + \frac{1}{2})} \sigma_j H_{j,m}\right) \\
&\leq N^{-pm/d} \mathbb{E}\|X\|_{B^p_p+m(L_p(D))}^p.
\end{align*}
\]

Since we have \(\mathbb{E}\|X\|_{B^p_p+m(L_p(D))} < \infty\), \(m < m^*\), by (21), we derive that the linear approximation error satisfies

\[
e_{N,p,\nu}^{\text{lin}}(X) \leq N^{-m/d} \left(\mathbb{E}\|X\|_{B^p_p+m(L_p(D))}^p\right)^{1/p}.
\]

From (27), we observe that, similar to the well-known deterministic setting, see, e.g., \([16, 23]\), the approximation order which can be achieved by (uniform) linear schemes depends on the regularity of the object under consideration in the same scale of smoothness spaces.

For the case \(p = 2\), i.e., for linear wavelet approximation with respect to \(H^\nu\), also a lower bound for the linear approximation error can be derived.

**Theorem 3.15.** Let \(\beta \in [0, 1]\) and \(m > 0\). For a fixed approximation space \(H^\nu(D)\), let \(X\) be given by (7) with \(\nu + m < d(\alpha - 1)/2 + \beta/2) =: \nu + m^*\), i.e., \(X \in H^{\nu+m}(D)\) for all \(m < m^*\). The linear approximation error with respect to \(H^\nu(D)\) satisfies

\[
e_{N,2,\nu}^{\text{lin}}(X) \leq (\log_2 N)^{\frac{dp}{d}} N^{-\left(\frac{\alpha - 1 + \beta}{2} - \frac{\beta}{2}\right)}.
\]

**Proof.** Given \((e_{j,k})_{j,k}\) and \(\Phi\) as in Remark 3.3, we know that

\[
e_{N,2,\nu}^{\text{lin}}(X) \leq e_{N,2,\nu}^{\text{lin}}(\Phi^{-1}X).
\]

Furthermore, we also know from Remark 3.3 that the covariance operator \(\tilde{Q}\) of \(\Phi^{-1}X\) is given by

\[
\tilde{Q}\xi = \sum_{j=0}^{\infty} 2^{2j\nu} \sigma_j^2 \rho_j \sum_{k \in \nabla_j} \langle \xi, e_{j,k} H^\nu(D) e_{j,k} \rangle,
\]

which means, that the functions \(e_{j,k}\) form an orthonormal basis of eigenfunctions of \(\tilde{Q}\) with associated eigenvalues \(2^{2j\nu} \sigma_j^2 \rho_j\). Using methods, e.g. shown in \([44, \text{Ch. III}]\), we get

\[
e_{N,2,\nu}^{\text{lin}}(\Phi^{-1}X) = \left(\sum_{j=j_1+1}^{\infty} \#\nabla_j 2^{2j\nu} \sigma_j^2 \rho_j\right)^{1/2}
\]

if \(N = \sum_{j=0}^{j_1} \#\nabla_j\). Inserting Assumption 3.1 (iv) and (5) into (28) yields the claim.

\(\square\)
Now we define the average nonlinear approximation error of $X : \Omega \to B^s_p(L_\tau(D))$ with respect to $B^s_p(L_\tau(D))$, where

$$\frac{1}{\tau} = \frac{s - \nu}{d} + \frac{1}{p}, \quad p \in (1, \infty), \quad \text{and} \quad \nu < s,$$

(29)
cf. Corollary 3.11, by

$$e_{N,p,\nu,\tau}^{\text{avg}}(X) := \inf \left( \mathbb{E} \| X - \hat{X} \|_{B^s_p(L_\tau(D))}^p \right)^{1/p}$$

with the infimum taken over all measurable mappings $\hat{X}$ such that $\mathbb{E}(\eta(\hat{X})) \leq N$. Here,

$$\eta(g) := \# \left\{ \lambda \in \nabla : g = \sum_{\lambda \in \nabla} c_\lambda \psi_\lambda, \ c_\lambda \neq 0 \right\}$$
denotes the number of nonzero wavelet coefficients of $g$.

**Theorem 3.16.** Let $\beta \in [0, 1)$ and $p \in (1, \infty)$. For a fixed approximation space $B^s_p(L_\tau(D))$, let $X$ be given by (7) with $-d/p \leq \nu < d((\alpha - 1)/2 + \beta/p)$, that is, $X \in B^s_p(L_\tau(D))$ in the scale (29) for all $s < s^*$, where $s^*$ is given by (23). Then the average nonlinear approximation error with respect to $B^s_p(L_\tau(D))$ satisfies

$$e_{N,p,\nu}^{\text{avg}}(X) \preceq \left( \log_2 N \right)^{\frac{d}{2} s^*} N^{-\frac{1}{p} \left( \frac{s^* - 1}{2} + \frac{1}{d} \right)}.$$

(30)

**Proof.** As a specific nonlinear approximation of $X$, let us consider

$$\hat{X}_{j_1} := \sum_{j=0}^{j_1} \sum_{k \in \nabla_j} \sigma_j Y_{j,k} Z_{j,k} \psi_{j,k}$$

for some $j_1 > 0$, where we only retain the non-zero coefficients $N := \mathbb{E}(\eta(\hat{X}_{j_1}))$. It holds that

$$N = \sum_{j=0}^{j_1} \# \nabla_j \rho_j \asymp 2^{(1-\beta)j_1 d}.$$

With $S_{j,p}$ being defined in (8) and with (9), we use (29), where $s = s^*$ and $\tau = \tau^*$, to obtain

$$\mathbb{E} \| X - \hat{X}_{j_1} \|_{B^s_p(L_\tau(D))}^p \asymp \mathbb{E} \left( \sum_{j=j_1+1}^{\infty} 2^{jp} \nu + d \left( \frac{1}{2} - \frac{1}{p} \right) \right) \sigma_j^p S_{j,p}$$

$$= \mathbb{E} \left( \sum_{j=j_1+1}^{\infty} 2^{jp} \sigma_j^p S_{j,p} \right)$$

$$= \sum_{j=j_1+1}^{\infty} 2^{jp} \sigma_j^p \# \nabla_j \rho_j.$$

Inserting Assumption 3.1 (iv), (5), (24), and (29), where $s = s^*$ and $\tau = \tau^*$, we get

$$\mathbb{E} \| X - \hat{X}_{j_1} \|_{B^s_p(L_\tau(D))}^p \asymp \sum_{j=j_1+1}^{\infty} 2^{jp} \left( \frac{\alpha - 1}{2} + \frac{1}{d} - \frac{1}{p} \right) \sigma_j^p \# \nabla_j \rho_j 2^{-\beta j d} 2^{j d/2 - \beta j d}$$

14
\[
\begin{align*}
&= \sum_{j=j_1+1}^{\infty} j^{-\frac{dp}{2}} 2^{-j\beta(1-\beta)(s^*-\nu)} \\
&\leq j^{-\frac{2dp}{2}} 2^{-j\beta(1-\beta)(s^*-\nu)} \\
&\leq (\log_2 N)^{\frac{2dp}{2}} N^{-\frac{1}{r}(\frac{\alpha-1}{2}+\frac{\beta}{p}-\frac{\nu}{2})},
\end{align*}
\]

which yields (30). \qed

An analogous statement to Remark 3.14 also holds for the average nonlinear approximation error.

**Remark 3.17.** Let \( \varepsilon > 0 \) and \( s := s^*-\varepsilon \) with \( s^* \) being defined in (23). In the setting of Theorem 3.16, with a slightly coarser error estimation, i.e., using (5) we get

\[
\begin{align*}
E \| X - \hat{X}_{j_1} \|_{B^p_{r}(L_p(D))}^p &\leq E \left( \sum_{j=j_1+1}^{\infty} 2^{j(p(\tau-\nu)(s^*-\nu)+\delta s)} j^p \| \frac{\partial}{\partial} \E \| X - \hat{X}_{j_1} \|_{B^p_{r}(L_p(D))}^p \leq E \left( \sum_{j=j_1+1}^{\infty} 2^{j(p(\tau-\nu)(s^*-\nu)+\delta s)} j^p \| \frac{\partial}{\partial} \right) \right),
\end{align*}
\]

for any \( \delta > 0 \). Inserting \( s = s^* - \varepsilon \) and \( \delta := \frac{p(s - \nu)(1 - \beta)(\varepsilon \tau)}{d} \), as well as using (24), (29), and also (29) with \( s = s^* \) and \( \tau = \tau^* \), which yields \( 1/\tau^* = 1/\tau + \varepsilon/d \), we get

\[
\begin{align*}
E \| X - \hat{X}_{j_1} \|_{B^p_{r}(L_p(D))}^p &\leq E \left( \sum_{j=j_1+1}^{\infty} 2^{j(p-\frac{1}{\tau}-\nu)(s^*-\nu)+\delta s} j^p \| \frac{\partial}{\partial} \right) \right),
\end{align*}
\]

with \( E(S_{j,p}) = \# \nabla_j \rho_j \mu_p \# \nabla_j \rho_j \nu_{s\mu_p} = E(S_{j,p}) \frac{\mu_p}{\nu_{s\mu_p}} \). Since we have \( E \| X \|_{B^p_{r}(L_r(D))}^r < \infty \) for \( s < s^* \), by (21), we see that the average nonlinear approximation error satisfies

\[
\varepsilon_{N,p,\nu}^{\text{avg}}(X) \leq N^{-\frac{p-1}{2}} \left( E \| X \|_{B^p_{r}(L_r(D))}^r \right)^{1/p}.
\]
From (31) we observe that, similar to the deterministic setting, the approximation order which can be achieved by nonlinear approximation does not depend on the regularity in the same scale of smoothness spaces of the object under consideration, but on the regularity in the corresponding adaptivity scale (29) of Besov spaces.

For the case \( p = 2 \), i.e., for nonlinear wavelet approximation with respect to \( H^\nu \), also a lower bound for the average nonlinear approximation error can be derived.

**Theorem 3.18.** Let \( \beta \in [0,1) \). For a fixed approximation space \( H^\nu(D) \), let \( X \) be given by (7) with \(-d/2 \leq \nu < d((\alpha - 1)/2 + \beta/2)\), i.e., \( X \in B^s_\nu(L_2(D)) \) in the scale (29) for all \( s < s^* \), where \( s^* \) is given by (23) with \( p = 2 \). Then the average nonlinear approximation error in \( H^\nu(D) \) satisfies

\[
e_{N,2,\nu}(X) \geq (\log_2 N)^{\frac{d}{\nu}} N^{-\frac{1}{\nu} \left( \frac{\alpha - 1 + \beta}{2} - \frac{d}{2} \right)}
\]

(32)

**Proof.** Let \( X \) be defined by (7). For every level \( j \), we define the number of scaled coefficients of \( X \) larger than \( \delta_j > 0 \) as

\[
M(j, \delta_j) := \# \{ 2^{\nu j} \sigma_j Y_{j,k} | Z_{j,k} | > \delta_j : k \in \mathcal{N}_j \}.
\]

(33)

We set \( Y_{j,\beta} := \sum_{k \in \mathcal{N}_j} Y_{j,k} \) and obtain \( Y_{j,\beta} \sim \text{Bin}(2^{jd}, 2^{-\beta jd}) \). Since the \( (Y_{j,k})_{j,k} \) are discrete and \( (Z_{j,l})_{j,l} \) are identically distributed we can compute

\[
E(M(j, \delta_j)) = \sum_{l=0}^{2^{jd}} E(M(j, \delta_j) \mid \sum_{k \in \mathcal{N}_j} Y_{j,k} = l) P\left( \sum_{k \in \mathcal{N}_j} Y_{j,k} = l \right)
= \sum_{l=0}^{2^{jd}} l P\left( 2^{\nu j} \sigma_j | Z_{j,l} | > \delta_j \right) P(Y_{j,\beta} = l)
= E(Y_{j,\beta}) P\left( 2^{\nu j} \sigma_j | Z_{j,k} | > \delta_j \right)
= 2^{jd(1-\beta)} 2^{jd} \left( 1 - \Phi\left( \frac{\delta_j}{2^{\nu j} \sigma_j} \right) \right),
\]

where \( \Phi(x) \) denotes the cumulative distribution function of the standard normal distribution. Now, we choose

\[
\delta_j := 2^{\nu j} \sigma_j
\]

(34)

and we obtain \( E(M(j, 2^{\nu j} \sigma_j)) = c_1 2^{jd(1-\beta)} \) with \( c_1 := 2(1 - \Phi(1)) \). For a given \( N \in \mathbb{N} \) we set \( j_0 := \min\{ j : N \leq 2^{jd} \} \) and determine a level \( j_1 \), such that

\[
E(M(j_1, 2^{\nu j_1} \sigma_{j_1})) \geq c_1 N^2.
\]

(35)

This holds for

\[
j_1 = \left\lfloor \frac{j_0}{1 - \beta} \right\rfloor.
\]

(36)

Up to this point we have shown that, for \( X \) and any given \( N \in \mathbb{N} \), we can find a level \( j_1 \), which contains on average at least \( c_1 N^2 \) coefficients, that are larger than \( \delta_{j_1} \).

Let \( \hat{X}_N := \sum_{j=0}^{\infty} \sum_{k \in \mathcal{N}_j} c_{j,k} \psi_{j,k} := \sum_{\lambda \in \mathcal{V}} c_\lambda \psi_\lambda \) with \( E(\#\hat{X}) = E(\eta(\hat{X}_N)) \leq N \) be any approximation of \( X = \sum_{j=0}^{\infty} \sum_{k \in \mathcal{N}_j} \sigma_j Y_{j,k} Z_{j,k} \psi_{j,k} := \sum_{\lambda \in \mathcal{V}} d_\lambda \psi_\lambda \). We set \( |\lambda| := j \). Then, by using the norm equivalence from Assumption 3.1 (vi), we obtain

\[
E \| X - \hat{X}_N \|_{H^\nu(D)}^2 = E \left\| \sum_{\lambda \in \mathcal{V}} d_\lambda \psi_\lambda - \sum_{\lambda \in \mathcal{V}} c_\lambda \psi_\lambda \right\|_{H^\nu(D)}^2
\]
\[
\begin{align*}
&= E \left\| \sum_{\lambda \in \nabla} d_{\lambda} \psi_{\lambda} + \sum_{\lambda \in \nabla} (d_{\lambda} - c_{\lambda}) \psi_{\lambda} \right\|_{H^\nu(D)}^2 \\
&\approx E \left( \sum_{\lambda \in \nabla} 2^{2|\lambda|\nu} |d_{\lambda}|^2 + \sum_{\lambda \in \nabla} 2^{2|\lambda|\nu} |d_{\lambda} - c_{\lambda}|^2 \right).
\end{align*}
\]

If we omit the second sum and by (33) and (34), we get
\[
E \|X - \hat{X}_N\|_{H^\nu(D)}^2 \geq E \left( \sum_{\lambda \in \nabla} 2^{2|\lambda|\nu} |d_{\lambda}|^2 \right)
\geq E \left( \sum_{\lambda \in \nabla} 2^{2|\lambda|\nu} |d_{\lambda}|^2 \right)
\geq E \left( \sum_{k \in \nabla_{j_1}} 2^{2|\sigma_{j_1}|} |d_{j_1,k}|^2 \right)
\geq E \left( \# \{ k \in \nabla_{j_1} : 2^{2|\sigma_{j_1}|} |d_{j_1,k}|^2 > \delta_{j_1}^2 \} \right) \cdot \delta_{j_1}^2
\geq E \left( M(j_1,2^{|\sigma_{j_1}|}) - \# \nabla \right) \cdot 2^{2j_1\nu}\sigma_{j_1}^2
= \left( E(M(j_1,2^{j_1\nu}\sigma_{j_1})) - E(\# \nabla) \right) \cdot 2^{2j_1\nu}\sigma_{j_1}^2,
\]
so that, by inserting (5), (35), (36), \(E(\# \nabla) \leq N \leq 2^{\frac{j_0d}{2}}\), we can conclude
\[
E \|X - \hat{X}_N\|_{H^\nu(D)}^2 \geq \left( 2^{j_0d} - 2^{\frac{j_0d}{2}} \right) 2^{2j_1\nu} j_1^{d} 2^{-\alpha j_1^{d}}
\geq j_0^{-d} 2^{j_0d} + 2^{\frac{j_0d}{2}} \frac{\alpha j_0d}{11\nu}
\geq \left( \log_2 N \right)^{\gamma d} N^{-\frac{1}{11\nu} (\alpha - 1 + \beta - \frac{2}{\nu})},
\]
which yields (32).

\[\square\]

**Remark 3.19.** (i) For the proof of Theorem 3.18 it is essential to be able to compute the expected value of \(M(j,\delta_j)\), i.e., the average number of coefficients on level \(j\) which are larger than a threshold. This random variable can be derived solely due to the structure of \(X\). Note that, since the threshold \(\delta_j = 2^{\nu j^{-d/2} - \alpha j^{d/2}}\) decays with increasing level \(j\), the growth of \(E(M(j,\delta_j))\) does not contradict Theorem 3.9.

(ii) Observe that the upper bound in Theorem 3.16 for \(p = 2\) coincides with the lower bound in Theorem 3.18.

### 4 A class of random functions in anisotropic Besov spaces

In this section, we study a new class of random functions in anisotropic Besov spaces. Based on a similar stochastic model as considered in the previous section, we derive conditions, under which such a random function (almost surely) has a certain smoothness in a given anisotropic Besov space \(B_{q,a}^s(L_p(D))\) for a wide range of parameters, cf. Theorem 4.9. The anisotropic Besov spaces are defined in Subsection 2.2.
Analogously to the previous section, we employ a wavelet characterization of the space under consideration, cf. Theorem 4.5. In contrast to the basic Assumption 3.1, which was concerned with the isotropic case, in the anisotropic case the dilatation of the wavelets depend on the anisotropy, as we will now explain.

Throughout this section we fix an anisotropy \( \mathbf{a} = (a_1, \ldots, a_d) \in \mathbb{R}_+^d \), cf. (3). We employ suitable \( M \)-scaling functions \( \varphi : \mathbb{R}^d \to \mathbb{R} \), which satisfy

\[
\varphi(\cdot) = |\det(M)|^{1/2} \sum_{k \in \mathbb{Z}^d} h_k \varphi(M \cdot - k)
\]

with a finite number of non-zero coefficients \( h_k \in \mathbb{R} \). Here, \( M \) is an anisotropic integer scaling matrix of the form

\[
M := \text{diag}(\lambda^{1/a_1}, \ldots, \lambda^{1/a_d})
\]

for some \( \lambda > 1 \), (38)

Note that with \( \sum_{i=1}^d \frac{1}{a_i} = d \), we get

\[
\lambda^d = |\det(M)| =: m.
\]

**Remark 4.1.** Matrices of the form (38) are the only ones compatible with the anisotropy for the type of multiscale decomposition we wish to apply, see [29, Sec. 3.3].

The following assumptions are needed, to construct a suitable wavelet basis that characterizes the given anisotropic Besov space \( B_{s,a}^{0,q}(L_p(D)) \).

**Assumption 4.2.** We assume to have an \( M \)-scaling function \( \varphi \) at hand, which satisfies the following properties:

1. \( \varphi \in H^s(\mathbb{R}^d) \) for some \( s > d/2 \),
2. \( \varphi \) is compactly supported and \( \int_{\mathbb{R}^d} \varphi(x) \, dx = 1 \),
3. \( \varphi \) is a refinable function in the sense of (37),
4. \( \{\varphi(\cdot - k)\}_{k \in \mathbb{Z}^d} \) is a Riesz basis of the space it spans, i.e.,

\[
\sum_{k \in \mathbb{Z}^d} |a_k|^2 \leq \| \sum_{k \in \mathbb{Z}^d} a_k \varphi(\cdot - k) \|^2_{L_2(\mathbb{R}^d)};
\]
5. \( \varphi \in B_{s_0,a}^{0,q}(L_p(\mathbb{R})) \cap H^{L,a}(\mathbb{R}^d) \) for some \( s_0 > 0 \) and \( \mathbb{N} \ni L > d/2 \).
6. Furthermore, there exists an \( M \)-scaling function \( \tilde{\varphi} \), which also satisfies (i)-(v) with potentially different constants, that is biorthogonal to \( \varphi \), i.e., \( \{\varphi(\cdot - k)\}_{k \in \mathbb{Z}^d} \) and \( \{\tilde{\varphi}(\cdot - k)\}_{k \in \mathbb{Z}^d} \) satisfy

\[
\int_{\mathbb{R}^d} \varphi(x) \tilde{\varphi}(x - k) \, dx = \delta_{0,k} \quad \text{for all } k \in \mathbb{Z}^d.
\]

**Remark 4.3.** (i) The existence of nontrivial scaling functions satisfying Assumption 4.2 is of course not obvious, nevertheless a lot of examples exist. We refer to [29, Sec. 3.1] for a detailed discussion.

(ii) By the Sobolev embedding theorem Assumptions 4.2 (i) and (ii) imply \( \varphi \) and \( \tilde{\varphi} \) to be continuous and that they are contained in \( L_p(\mathbb{R}^d) \) for all \( 1 \leq p \leq \infty \).

(iii) The anisotropic smoothness of \( \varphi \) and \( \tilde{\varphi} \) in Assumption 4.2 (v) directly affects the range of smoothness that can be characterized later, cf. Theorem 4.5 below.
From now on, let \( \varphi \) be an \( \mathbf{M} \)-scaling function satisfying Assumption 4.2. Furthermore, let \( 1 \leq p < \infty \) and \( 1/p + 1/p' = 1 \), where \( p' = \infty \), if \( p = 1 \), and for any real-valued functions \( f_1 \in L_p(\mathbb{R}^d) \) and \( f_2 \in L_{p'}(\mathbb{R}^d) \) we set
\[
\langle f_1, f_2 \rangle := \int_{\mathbb{R}^d} f_1(x)f_2(x) \, dx.
\]

Since, in particular, \( \varphi, \tilde{\varphi} \in L_\infty(\mathbb{R}^d) \) they generate MRAs \( (V_j^{(p)})_{j \in \mathbb{Z}} \) and \( (\tilde{V}_j^{(p')})_{j \in \mathbb{Z}} \) by
\[
V_j^{(p)} := \text{clos span}_{L_p(\mathbb{R}^d)} \left\{ \varphi_{j,k}^{(p)} := |\det(M)|^{j/p} \varphi(M^j \cdot - k) : k \in \mathbb{Z}^d \right\}, \quad j \in \mathbb{Z},
\]
and
\[
\tilde{V}_j^{(p')} := \text{clos span}_{L_{p'}(\mathbb{R}^d)} \left\{ \tilde{\varphi}_{j,k}^{(p')} := |\det(M)|^{j/p'} \tilde{\varphi}(M^j \cdot - k) : k \in \mathbb{Z}^d \right\}, \quad j \in \mathbb{Z},
\]
see [29, Sec. 3.1]. Moreover, \( (V_j^{(p)})_{j \in \mathbb{Z}} \) and \( (\tilde{V}_j^{(p')})_{j \in \mathbb{Z}} \) are biorthogonal, that is
\[
\langle \varphi_{j,k}, \tilde{\varphi}_{j',k'} \rangle = \delta_{k,k'}, \quad j \in \mathbb{Z}.
\]

**Remark 4.4.** An MRA \( (V_j^{(q)})_{j \in \mathbb{Z}}, 1 \leq q \leq \infty \), is a sequence of closed linear subspaces of \( L_q(\mathbb{R}^d) \) with the following properties. Note, in the case \( q = \infty \) the space \( L_q(\mathbb{R}^d) \) is replaced by \( C_0(\mathbb{R}^d) \), the space of continuous functions with compact support, and \( \ell_q(\mathbb{Z}^d) \) is replaced by \( c_0(\mathbb{Z}^d) \), the space of sequences converging to zero.

- \( V_j^{(q)} \subset V_{j+1}^{(q)} \) for all \( j \in \mathbb{Z} \),
- \( \bigcup_{j \in \mathbb{Z}} V_j^{(q)} \) is dense in \( L_q(\mathbb{R}^d) \) and \( \bigcap_{j \in \mathbb{Z}} V_j^{(q)} = \{0\} \),
- there exists a function \( \varphi \in V_0^{(q)} \) such that \( \{\varphi(\cdot - k)\}_{k \in \mathbb{Z}^d} \) is a \( q \)-stable basis of \( V_0^{(q)} \), i.e., \( \text{span}\{\varphi(\cdot - k) : k \in \mathbb{Z}^d\} = V_0^{(q)} \) and
\[
\left\| \sum_{k \in \mathbb{Z}^d} a_k \varphi(\cdot - k) \right\|_{L_q(\mathbb{R}^d)} \asymp \left( \sum_{k \in \mathbb{Z}^d} |a_k|^q \right)^{1/q},
\]
- \( f(\cdot) \in V_j^{(q)} \) if and only if \( f(M \cdot) \in V_{j+1}^{(q)} \) for all \( j \in \mathbb{Z} \).

Wavelets come into play as basis functions for the detail spaces \( W_j^{(p)} \) and \( \tilde{W}_j^{(p')} \) to the MRAs, which are defined by
\[
W_j^{(p)} := \left\{ f \in V_{j+1}^{(p)} : \langle f, \tilde{f} \rangle = 0 \text{ for all } \tilde{f} \in \tilde{V}_j^{(p')} \right\}, \quad j \in \mathbb{Z},
\]
and \( \tilde{W}_j^{(p')} \) is defined analogously. In this way, one obtains a decomposition of the form
\[
L_p(\mathbb{R}^d) = V_0^{(p)} \oplus \left( \bigoplus_{j=0}^{\infty} W_j^{(p)} \right), \quad 1 \leq p < \infty.
\]

19
Moreover, the following norm equivalence holds:

\[
\{ \psi_{e,j,k}^{(p)} \}_{e,j,k} := \{ \psi_e(M^j - k) : e = 1, \ldots, m - 1, j \in \mathbb{Z}, k \in \mathbb{Z}^d \}, \tag{39}
\]

and analogously for \( \{ \tilde{\psi}_{e,j,k}^{(p')} \}_{e,j,k} \), are uniformly \( p, p' \)-stable bases of \( W_j^{(p)} \) and \( \tilde{W}_j^{(p')} \), respectively, see [29, Sec. 3.2] for details. Accordingly, two wavelet bases \( \Psi, \tilde{\Psi} \) for \( L_p(\mathbb{R}^d), 1 \leq p < \infty \), are given by

\[
\Psi := \{ \varphi_{0,k}^{(p)} : k \in \mathbb{Z}^d \} \cup \{ \psi_{e,j,k}^{(p)} : e = 1, \ldots, m - 1, j \in \mathbb{N}_0, k \in \mathbb{Z}^d \}
\]

and

\[
\tilde{\Psi} := \{ \tilde{\varphi}_{0,k}^{(p')} : k \in \mathbb{Z}^d \} \cup \{ \tilde{\psi}_{e,j,k}^{(p')} : e = 1, \ldots, m - 1, j \in \mathbb{N}_0, k \in \mathbb{Z}^d \}.
\]

Moreover, \( \Psi \) and \( \tilde{\Psi} \) are biorthogonal, that is \( \langle \psi_{e,j,k}, \tilde{\psi}_{e',j',k} \rangle = \delta_{k,k'} \delta_{j,j'} \). In particular, for every \( f \in L_p(\mathbb{R}^d) \) we have the multiscale decomposition

\[
f = \sum_{k \in \mathbb{Z}^d} \langle f, \tilde{\psi}_{0,k}^{(p')} \rangle \varphi_{0,k}^{(p')} + \sum_{j=0}^{\infty} \sum_{e=1}^{m-1} \sum_{k \in \mathbb{Z}^d} \langle f, \tilde{\psi}_{e,j,k}^{(p')} \rangle \psi_{e,j,k}^{(p')}.
\]

The following theorem states the wavelet characterization of the anisotropic Besov space \( B_q^{s,a}(L_p(\mathbb{R}^d)) \). Part 1 is cited from [29, Thm. 5.3] and considers the case \( 1 \leq p, q < \infty \), while Part 2 considers the quasi-Banach case and is taken from [28, Thm. 6.2]. Since the biorthogonal setting described above is restricted to \( L_p \)-spaces with \( p \geq 1 \), only anisotropic quasi-Banach spaces which are embedded in some \( L_p \)-space with \( p \geq 1 \) can be characterized in this way. In [28], it has been shown that \( B_q^{s,a}(L_\tau(\mathbb{R}^d)) \hookrightarrow L_p(\mathbb{R}^d) \) for every finite \( p \) such that \( \tau \leq p \leq p(s, \tau) \), where

\[
p(s, \tau) := \begin{cases} \left(1/\tau - s/d\right)^{-1} & : s < d/\tau, \\ \infty & : \text{otherwise}. \end{cases} \tag{40}
\]

This explains the restrictions in Part 2.

**Theorem 4.5. Part 1** [29, Thm. 5.3]. Suppose Assumption 4.2 is satisfied and let \( 1 \leq p, q < \infty \). If \( 0 < s < \min\{s_0, L/a_1, \ldots, L/a_d\} \), then

\[
B_q^{s,a}(L_p(\mathbb{R}^d)) = \left\{ f \in L_p(\mathbb{R}^d) : \sum_{k \in \mathbb{Z}^d} |\langle f, \varphi_{0,k}^{(p)} \rangle|_p^p + \sum_{j=0}^{\infty} \sum_{e=1}^{m-1} \sum_{k \in \mathbb{Z}^d} |\langle f, \psi_{e,j,k}^{(p')} \rangle|_p^p / |\langle \varphi_{0,k}^{(p')} \rangle|_p^{q/p} < \infty \right\}.
\]

Moreover, the following norm equivalence holds:

\[
\|f\|_{B_q^{s,a}(L_p(\mathbb{R}^d))} \asymp \left( \sum_{k \in \mathbb{Z}^d} |\langle \cdot, \tilde{\varphi}_{0,k}^{(p')} \rangle|_p^p \right)^{1/p} + \left( \sum_{j=0}^{\infty} \sum_{e=1}^{m-1} \sum_{k \in \mathbb{Z}^d} |\langle \cdot, \tilde{\psi}_{e,j,k}^{(p')} \rangle|_p^p \right)^{q/p} \right)^{1/q}. \tag{41}
\]

**Part 2** [28]. Suppose Assumption 4.2 is satisfied. Let \( 0 < \tau < 1, d(1/\tau - 1) < s < \min\{s_0, L/a_1, \ldots, L/a_d\} \), and \( p(s, \tau) \) be defined by (40). Then, for any \( p \in (1, \infty) \) with \( p \leq p(s, \tau) \) we have

\[
B_q^{s,a}(L_\tau(\mathbb{R}^d))
\]
\[ f \in L^p(\mathbb{R}^d) : \sum_{k \in \mathbb{Z}^d} |\langle f, \tilde{\varphi}_{0,k}^{(p')} \rangle|^\tau + \sum_{j=0}^{\infty} m^{\tau j} \sum_{e=1}^{m-1} \sum_{k \in \mathbb{Z}^d} |\langle f, \tilde{\psi}_{e,j,k}^{(p')} \rangle|^\tau < \infty \].

Moreover, the following quasi-norm equivalence holds:

\[ \| \cdot \|_{B^{s,a}_{p,q}(L^\tau(\mathbb{R}^d))} \asymp \left( \sum_{k \in \mathbb{Z}^d} |\langle \cdot, \tilde{\varphi}_{0,k}^{(p')} \rangle|^p \right)^{1/\tau} + \left( \sum_{j=0}^{\infty} m^{\tau j} \sum_{e=1}^{m-1} \sum_{k \in \mathbb{Z}^d} |\langle \cdot, \tilde{\psi}_{e,j,k}^{(p')} \rangle|^\tau \right)^{1/\tau}. \] (42)

**Remark 4.6.** In [48] a wavelet characterization for the anisotropic Besov spaces in the general case \( 0 \leq p,q < \infty \), \( s \in \mathbb{R} \) is discussed, where orthogonal wavelet bases, e.g., Daubechies type wavelets, are employed. The (more general) biorthogonal case is not considered.

It is our goal to construct random functions on bounded (Lipschitz) domains \( D \subset \mathbb{R}^d \) taking values in anisotropic Besov spaces with the aid of the wavelet characterization of Besov spaces as outlined in Theorem 4.5. However, here we are facing a nontrivial problem. The analysis in Section 3 was particularly designed for bounded domains, and since the cardinality of \( \nabla_j \), cf. Assumption 3.1 (iv), plays a central role, it is more or less restricted to this case. Unfortunately, Theorem 4.5 is concerned with function spaces on the whole Euclidean \( d \)-plane, and a generalization to bounded domains is at least not obvious, since this would require the construction of specific boundary-adapted anisotropic wavelet bases on domains. To our best knowledge no result in this direction has been reported so far. It seems reasonably, that, e.g., the construction outlined in [20] can be generalized to the anisotropic case, but that would be beyond the scope of this paper. As a possible way out, we proceed in the following way.

Let \( D \subset \mathbb{R}^d \) be a bounded Lipschitz domain, then there exists a cube \( \Box \supset D \) such that

\[ \varphi_{0,k}^{(p)} \cap D \neq \emptyset \quad \text{and} \quad \psi_{e,j,k}^{(p)} \cap D \neq \emptyset \]

implies \( \text{supp} \varphi_{0,k}^{(p)} \subset \subset \Box \) and \( \text{supp} \psi_{e,j,k}^{(p)} \subset \subset \Box \). Based on this cube \( \Box \) we define the sets

\[ \nabla_0 := \{ k \in \mathbb{Z}^d : \text{supp} \varphi_{0,k}^{(p)} \subset \subset \Box \}, \]

\[ \nabla_{e,j} := \{ k \in \mathbb{M}^{-j}\mathbb{Z}^d : \text{supp} \psi_{e,j,k}^{(p)} \subset \subset \Box \} \] (43)

for \( e = 1, \ldots, m-1 \) and \( j \in \mathbb{N}_0 \).

**Assumption 4.7.** We assume to have compactly-supported wavelets at hand such that \( \nabla_0 \) is a finite set and the sets \( \nabla_{e,j} \) in (43) fulfill

\[ \# \nabla_{e,j} \asymp m^j = |\text{det}(\mathbb{M})|^j. \]

**Remark 4.8.** Assumption 4.7 on the cardinality of \( \nabla_{e,j} \) slightly varies from Assumption 3.1 (iv) for the isotropic setting. It is motivated by the following observation. If we start with a tensor wavelet construction on the whole Euclidean plane, then, since \( \mathbb{M} \) is a diagonal matrix, the relevant grid points can be computed coordinate-wise and are of the order \( \lambda_i^{j/a_i} \), \( i = 1, \ldots, d \). By observing \( \prod_{i=1}^d \lambda_i^{j/a_i} = |\text{det}(\mathbb{M})|^j \) Assumption 4.7 is satisfied in this case.
Now, we construct a class of random functions analogously to Subsection 3.2. Here, we define
\[ X := \sum_{k \in \nabla_0} \sigma_0 Y_{0,k} Z_{0,k} \varphi_{0,k}^{(p)} + \sum_{j=0}^{\infty} \sum_{e=1}^{m-1} \sum_{k \in \nabla_{e,j}} \sigma_j Y_{e,j,k} Z_{e,j,k} \psi_{e,j,k}^{(p)}, \] (44)
where \((Y_{e,j,k}, Z_{e,j,k})_{e,j,k}, e = 1, \ldots, m - 1, j \in \mathbb{N}_0, k \in \nabla_{e,j},\) is an independent family of random variables on a probability space \((\Omega, \mathfrak{A}, P).\) As in Subsection 3.2, this stochastic model also depends on three parameters \(\alpha, \gamma \in \mathbb{R}, \beta \in [0, 1].\) (45)

The variables \(Y_{e,j,k}\) are Bernoulli distributed with parameter \(\rho_j := 2^{-\beta j d},\) where \(P(Y_{e,j,k} = 1) = \rho_j\) and \(P(Y_{e,j,k} = 0) = 1 - \rho_j.\)

The variables \(Z_{e,j,k}\) are \(N(0, 1)\)-distributed, and we set \(\sigma_j^2 := j^{-\gamma d} 2^{-a j d}, j \in \mathbb{N},\) and \(\sigma_0 := 1.\)

Now, we can state conditions on the parameters (45) under which a random function \(X\) according to (44) (almost surely) belongs to an anisotropic Besov space \(B_{s,a}^q(L_p(D)).\)

The following theorem is the second main result of this paper.

**Theorem 4.9.** Let Assumptions 4.2 and 4.7 be satisfied, that is (41) and (42) hold. Let \(X\) be a random function as defined in (44). Then \(X|_D\) is \(P\)-almost surely contained in \(B_{s,a}^q(L_p(D))\), for \(s > d(1/p - 1)\) and either \(1 \leq p, q < \infty\) or \(0 < p = q < 1,\) if
\[ s < \frac{d^2}{\log_2 m} \left( \frac{\alpha}{2} + \frac{\beta}{p} \right) - \frac{d}{p}, \] (46)
or
\[ s \leq \frac{d^2}{\log_2 m} \left( \frac{\alpha}{2} + \frac{\beta}{p} \right) - \frac{d}{p} \text{ and } \ q \gamma d < -2. \] (47)

In both cases
\[ \mathbb{E}\|X|_D\|^q_{B_{s,a}^q(L_p(D))} < \infty. \] (48)

**Proof.** Since
\[ \|f\|_{B_{s,a}^q(L_p(D))} := \inf \{ \|g\|_{B_{s,a}^q(L_p(\mathbb{R}^d))} : g \in B_{s,a}^q(L_p(\mathbb{R}^d)), g|_D = f \}, \]
it is sufficient to show that \(X\) as defined by (44) is \(P\)-a.s. contained in \(B_{s,a}^q(L_p(\mathbb{R}^d)).\)

Using Theorem 4.5 and since, by the definition (43), the support of \(X\) is contained in the cube \(\square,\) we have \(X \in B_{s,a}^q(L_p(\mathbb{R}^d))\) \(P\)-a.s. if and only if \(P\)-a.s.
\[ \|X\|_{B_{s,a}^q(L_p(\mathbb{R}^d))} \leq \left( \sum_{k \in \nabla_0} |\langle X, \varphi_{0,k}^{(p)} \rangle|^{p/q} \right)^{1/p} + \sum_{j=0}^{\infty} m^{q \gamma d} \left( \sum_{e=1}^{m-1} \sum_{k \in \nabla_{e,j}} |\langle X, \psi_{e,j,k}^{(p)} \rangle|^{p/q} \right)^{1/p} < \infty. \]
Now, observe that the first sum in the above formula is finite, since by Assumption 4.7 the set $\nabla_0$ is finite. Therefore we are left to show that the second sum is also finite. Inserting (44) and using the abbreviation

$$S_{e,j,p} := \sum_{k \in \nabla_{e,j}} Y_{e,j,k} |Z_{e,j,k}|^p,$$

yields

$$\sum_{j=0}^{\infty} m^{\frac{2q}{q-p}} \left( \sum_{e=1}^{m-1} \sum_{k \in \mathbb{Z}^d} |\langle X, \tilde{\psi}_{e,j,k} \rangle|^p \right)^{q/p} \asymp \sum_{j=0}^{\infty} m^{\frac{2q}{q-p}} \sigma_j^q \left( \sum_{e=1}^{m-1} S_{e,j,p} \right)^{q/p}. $$

So, with $a_j := m^{\frac{2q}{q-p}} \sigma_j^q$, we have to show that

$$\sum_{j=0}^{\infty} a_j \left( \sum_{e=1}^{m-1} S_{e,j,p} \right)^{q/p} < \infty, \quad \text{P-a.s.} \quad (49)$$

Analogously to the corresponding steps in the proof of Theorem 3.9 it can be shown that (49) is equivalent to

$$\sum_{j=0}^{\infty} a_j \left( \sum_{e=1}^{m-1} \# \nabla_{e,j} \rho_j \right)^{q/p} < \infty. \quad (50)$$

Inserting Assumption 4.7 and the stochastic model into (50) yields

$$\sum_{j=0}^{\infty} a_j \left( \sum_{e=1}^{m-1} \# \nabla_{e,j} \rho_j \right)^{q/p} \asymp \sum_{j=0}^{\infty} a_j \left( (m-1)m^d \rho_j \right)^{q/p} \asymp \sum_{j=0}^{\infty} (m-1)^{q/p} j^{\frac{2m^d}{q} - 2qR},$$

with $R := (\frac{q}{2} \log_2 m - \frac{od}{2} + \frac{1}{p} \log_2 m - \frac{3d}{p})$. Therefore, (50) holds if and only if the conditions (46) or (47) are satisfied.

It remains to prove (48). Note, since (46) or (47) are satisfied, we have that (50) holds. Using the norm equivalences (41) and (42), as well as Lemma 3.6 and Lemma 3.7, with $\nabla_j := \bigcup_{e=1}^{m-1} \nabla_{e,j}$ and $S_{j,p} := \sum_{e=1}^{m-1} S_{e,j,p}$ instead of (8), we can conclude

$$E \|X\|_{B^a_{q,p}(L_p(\mathbb{R}^d))}^q \leq \sum_{j=0}^{\infty} a_j E \left[ \left( \sum_{e=1}^{m-1} S_{e,j,p} \right)^{q/p} \right] \leq \sum_{j=0}^{\infty} a_j \left( \sum_{e=1}^{m-1} \# \nabla_{e,j} \rho_j \right)^{q/p} < \infty. \quad \square$$

**Remark 4.10.** (i) Note, for $p = 2$ and in the isotropic case, where $a = 1$, $\lambda = 2$, and $\log_2 m = d$, the statements (46) and (47) of Theorem 4.9 coincide with (19) and (20). (ii) For $p \neq 2$ and in the isotropic case, note that the wavelets in (39) are normalized in $L_p(\mathbb{R}^d)$. A renormalization of these wavelets to $L_2(\mathbb{R}^d)$ would produce the additional factor $m^d(1/2-1/p)$ and yield that the statements (46) and (47) coincide with (19) and (20) also in this case.
5 Random tensor wavelet expansions

In this section, we study a new class of random tensor wavelet expansions, based on a similar stochastic model as given in the previous sections. We derive conditions, cf. Theorem 5.5, under which such a random tensor expansion (almost surely) has a certain smoothness in the tensor space $H^{t,\ell}(D)$, $t \in [0,\infty)^n$, $\ell \in [0,\infty)$, as defined in Subsection 2.3.

In the spirit of the previous sections, we employ a wavelet characterization of the space under consideration. Here, note that, the domain $D \subset \mathbb{R}^d$, $d > 1$ is assumed to be an $n$-fold product of component domains $D_m \subset \mathbb{R}^{d_m}$, $m = 1, \ldots, n$, $n \geq 2$, with $\sum_{m=1}^n d_m = d$.

**Assumption 5.1.** We assume that all domains $D_m$, $m = 1, \ldots, n$, allow the construction of a wavelet basis $(\psi_{\lambda_m}^{(m)})_{\lambda_m \in \Lambda_m}$, which is sufficiently smooth and has sufficiently many vanishing moments, such that for all $\ell' \in [t_m, t_m + \ell]$ the scaled wavelets $\{2^{-|\lambda_m|\ell'} \psi_{\lambda_m}^{(m)} : \lambda_m \in \Lambda_m\}$ are Riesz bases for $H^{\ell'}(D_m)$.

Similar to the previous sections, the wavelet indices $\lambda_m \in \Lambda_m$, $m = 1, \ldots, n$, are of the form $\lambda_m = (j_m, k_m)$, where $|\lambda_m| = j_m \in \mathbb{N}_0$ is the scale of the wavelet and $k_m \in \nabla_{j_m}$ encodes the shift and type of the wavelet. For the remaining part, we impose the following assumption which is the natural generalization of Assumption 3.1 (iv).

**Assumption 5.2.** We assume to have suitable compactly supported wavelets at hand for which

$$
\#\nabla_{j_m} \asymp 2^{j_m d_m}.
$$

Now, a tensor wavelet basis on the domain $D$ is defined as the collection of all functions $(\psi_\lambda)_{\lambda \in \Lambda}$ of the form

$$
\psi_\lambda := \bigotimes_{m=1}^n \psi_{\lambda_m}^{(m)}
$$

with $\lambda := (\lambda_1, \ldots, \lambda_n)$, $|\lambda| := (|\lambda_1|, \ldots, |\lambda_n|)$, and $\Lambda := \prod_{m=1}^n \Lambda_m$.

**Remark 5.3.** In comparison to standard isotropic wavelets, tensor wavelets are in a certain sense the wavelet version of the sparse grid approach, see, e.g., [5] for a detailed discussion on sparse grids. The application of (adaptive) tensor product approximation schemes gives rise to convergence rates that only depend on the component domains. We refer to, e.g., [46] and the references therein for further information. In particular, if all the component domains are one-dimensional one obtains dimensional independent convergence rates.

The wavelet characterization for $H^{t,\ell}(D)$ is now given by the following theorem.

**Theorem 5.4** [31,32]. Suppose Assumption 5.1 is satisfied. Then

$$
\{2^{-|\lambda|\ell} - \ell^{\|\lambda\|\infty} \psi_\lambda : \lambda \in \Lambda\}
$$

is a Riesz basis for $H^{t,\ell}(D)$. In particular, for every $f = \sum_{\lambda \in \Lambda} a_\lambda(f) \psi_\lambda$ the following two statements are equivalent:

i) $f \in H^{t,\ell}(D)$,

ii) $\sum_{\lambda \in \Lambda} |a_\lambda(f)|^2 2^{2t|\lambda| + 2\ell\|\lambda\|\infty} < \infty$.
The aim of this section is to derive random tensor expansions of the form

\[ X = \sum_{\lambda \in \Lambda} a_\lambda(X) \otimes_{m=1}^{n} \psi_{\lambda_m}^{(m)}, \quad \text{where} \quad a_\lambda(X) = \prod_{\lambda_m \in \Lambda_m} a_{\lambda_m}(X), \quad (51) \]

which have a prescribed smoothness in \( H^{t,\ell}(D) \). Here, the sequence \( (a_{\lambda_m}(X))_{\lambda_m \in \Lambda_m} \) of random variables \( a_{\lambda_m} : \Omega \to \mathbb{R} \) is based on a stochastic model, which is similar to the stochastic models of the previous sections. Given \( \alpha_m, \gamma_m \in \mathbb{R} \), and \( \beta_m \in [0, 1], m = 1, \ldots, n \), and a probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \), we set

\[ a_{\lambda_m} := a_{j_m, k_m} := \sigma_{j_m} Y_{j_m, k_m}^{(m)} Z_{j_m, k_m}^{(m)}, \quad j_m \in \mathbb{N}_0, \; k_m \in \nabla_{j_m}, \; m = 1, \ldots, n, \quad (52) \]

where \( Z_{j_m, k_m}^{(m)} \sim N(0, 1) \) are standard-normally distributed,

\[ \sigma_{j_m}^2 := j_m^{\gamma_m d_m} 2^{-\alpha_m j_m d_m}, \quad \sigma_{j_m} := 1 \text{ for } j_m = 0, \quad (53) \]

and \( Y_{j_m, k_m}^{(m)} \) are Bernoulli distributed random variables with parameter

\[ \rho_{j_m} := 2^{-\beta_m j_m d_m} \quad \text{and} \quad P(Y_{j_m, k_m}^{(m)} = 1) = 1 - P(Y_{j_m, k_m}^{(m)} = 0) = \rho_{j_m}. \quad (54) \]

Also, we assume the family of random variables \( \{Y_{j_m, k_m}^{(m)}, Z_{j_m, k_m}^{(m)}\}_{j_m, k_m} \) to be independent.

Now, we state the conditions on the parameters \( \alpha_m, \beta_m, \gamma_m, m = 1, \ldots, n \), such that the expansion \( (51), \) with \( (52) \), is (almost surely) contained in a given tensor space \( H^{t,\ell}(D) \). The following theorem is the third main result of this paper.

**Theorem 5.5.** Let Assumptions 5.1 and 5.2 be satisfied. Let \( X \) be a random tensor expansion of the form \( (51), \) \( (52) \). Then

i) \( X \) is contained in \( H^{t,\ell}(D) \), \( t = (t_1, \ldots, t_n) \), \( \mathbb{P} \)-almost surely if and only if

\[ t_m + \ell < d_m \left( \frac{\alpha_m + \beta_m - 1}{2} \right), \quad \gamma_m = 0, \quad m = 1, \ldots, n. \quad (55) \]

ii) \( X \) is contained in \( H^{t,\ell}(D) \), \( t = (t_1, \ldots, t_n) \), \( \mathbb{P} \)-almost surely if

\[ t_m + \ell \leq d_m \left( \frac{\alpha_m + \beta_m - 1}{2} \right), \quad \gamma_m d_m < -1, \quad m = 1, \ldots, n. \quad (56) \]

In both cases

\[ E\|X\|_{H^{t,\ell}(D)}^2 < \infty. \quad (57) \]

**Proof.** In order to prove \( (55) \) and \( (56) \), according to Theorem 5.4, we have to check under which conditions on \( \alpha_m, \beta_m, \) and \( \gamma_m, m = 1, \ldots, n \), we have

\[ \sum_{\lambda \in \Lambda} |a_\lambda(X)|^2 2^{|2\lambda|} \|X\|_{\infty} < \infty, \quad \mathbb{P} \text{-a.s.} \quad (58) \]

Step 1. Inserting \( (52) \) into \( (58) \) and using the abbreviations

\[ S_{j_m, 2} := \sum_{k_m \in \nabla_{j_m}} \left| Y_{j_m, k_m}^{(m)} \right|^2 \left| Z_{j_m, k_m}^{(m)} \right|^2, \quad m = 1, \ldots, n. \quad (59) \]
as well as $\bar{a}_j := 2^{2t_j + 2\ell ||j||_\infty} \sigma^2_{j_1} \cdots \sigma^2_{j_n}$, where $j = (j_1, \ldots, j_n)$, yields that we have to check when

$$\sum_{j \in \mathbb{N}_0^n} \bar{a}_j \prod_{m=1}^n S_{j_m, 2} < \infty, \quad \text{P-a.s.} \quad (60)$$

Applying Lemma 3.6 and Lemma 3.7, (60) is equivalent to

$$\sum_{j \in \mathbb{N}_0^n} \bar{a}_j \prod_{m=1}^n \#\nabla_{j_m} \rho_{j_m} < \infty, \quad (61)$$

which is shown analogously to the corresponding steps in the proof of Theorem 3.9.

Step 2. We first prove the claim for the case $n = 2$ to better illustrate the key ideas. Moreover, we will restrict this case to $\gamma_1 = \gamma_2 = 0$ in (53) of the stochastic model. The general case will be proven in Step 4.

With $n = 2$ we have $t = (t_1, t_2)$, $\Lambda = \Lambda_1 \times \Lambda_2$, and $|\lambda| = (|\lambda_1|, |\lambda_2|) = (j_1, j_2)$ for $\lambda \in \Lambda$, as well as $a_\lambda = a_{\lambda_1}a_{\lambda_2}$ in (51). Furthermore, we have $\lambda_m = (j_m, k_m)$ with $j_m \in \mathbb{N}_0$ and $k_m \in \nabla_{j_m}$, where $\nabla_{j_m}$ are finite sets with $\#\nabla_{j_m} \asymp 2^{j_m \alpha_m}$, $m = 1, 2$, cf. Assumption 5.2. Hence, we have to check under which conditions

$$\sum_{\lambda \in \Lambda} |a_\lambda|^2 2^{2t||\lambda||_\infty} = \sum_{\lambda_1 \in \Lambda_1} \sum_{\lambda_2 \in \Lambda_2} |a_{\lambda_1}|^2 |a_{\lambda_2}|^2 2^{2t_1 |\lambda_1|_1 + t_2 |\lambda_2|_2 + 2\ell ||(\lambda_1, \lambda_2)||_\infty}$$

$$= \sum_{j_1 \in \mathbb{N}_0} \sum_{k_1 \in \nabla_{j_1}} \sum_{j_2 \in \mathbb{N}_0} \sum_{k_2 \in \nabla_{j_2}} |a_{j_1, k_1}|^2 |a_{j_2, k_2}|^2 2^{2t_1 j_1 + t_2 j_2 + 2\ell ||(j_1, j_2)||_\infty}$$

$$< \infty, \quad \text{P-a.s.}$$

That is, inserting (52) and the abbreviation (59) we have to show under which conditions on $\alpha_m, \beta_m, \gamma_m, m = 1, 2$, we get

$$\sum_{j_1 \in \mathbb{N}_0} 2^{2t_1 j_1} \sigma^2_{j_1} S_{j_1, 2} \sum_{j_2 \in \mathbb{N}_0} 2^{2t_2 j_2 + 2\ell ||(j_1, j_2)||_\infty} \sigma^2_{j_2} S_{j_2, 2} < \infty, \quad \text{P-a.s.} \quad (62)$$

Applying Lemma 3.6 and Lemma 3.7, (62) is equivalent to

$$\sum_{j_1 \in \mathbb{N}_0} 2^{2t_1 j_1} \sigma^2_{j_1} \#\nabla_{j_1} \rho_{j_1} \sum_{j_2 \in \mathbb{N}_0} 2^{2t_2 j_2 + 2\ell ||(j_1, j_2)||_\infty} \sigma^2_{j_2} \#\nabla_{j_2} \rho_{j_2} < \infty.$$

Using Assumption 5.2 and (53) with $\gamma_1 = \gamma_2 = 0$, as well as (54) we continue with the calculation

$$\sum_{j_1 \in \mathbb{N}_0} 2^{2t_1 j_1} \sigma^2_{j_1} \#\nabla_{j_1} \rho_{j_1} \sum_{j_2 \in \mathbb{N}_0} 2^{2t_2 j_2 + 2\ell ||(j_1, j_2)||_\infty} \sigma^2_{j_2} \#\nabla_{j_2} \rho_{j_2}$$

$$\asymp \sum_{j_1 \in \mathbb{N}_0} 2^{j_1 (2t_1 - d_1 (\alpha_1 + \beta_1 - 1))} \sum_{j_2 \in \mathbb{N}_0} 2^{j_2 (2t_2 - d_2 (\alpha_2 + \beta_2 - 1)) + 2\ell ||(j_1, j_2)||_\infty}$$

$$= \sum_{j_1 \in \mathbb{N}_0} 2^{j_1 (2t_1 - d_1 (\alpha_1 + \beta_1 - 1))} \left( \sum_{j_2 < j_1} 2^{j_2 (2t_2 - d_2 (\alpha_2 + \beta_2 - 1)) + 2\ell j_1} + \sum_{j_2 \geq j_1} 2^{j_2 (2t_2 - d_2 (\alpha_2 + \beta_2 - 1) + 2\ell) j_2} \right)$$

26
\[=: \sum_{j_1 \in \mathbb{N}_0} A_{j_1} \left( \sum_{j_2 < j_1} B_{j_1, j_2} + \sum_{j_2 \geq j_1} C_{j_1, j_2} \right) =: \sum_{j_1 \in \mathbb{N}_0} A_{j_1} (B_{j_1} + C_{j_1}). \tag{63}\]

Now, observe that \(\sum_{j_2 \geq j_1} C_{j_1, j_2} < \infty\) if and only if
\[t_2 + \ell < d_2 \left( \frac{\alpha_2 + \beta_2 - 1}{2} \right). \tag{64}\]
Applying the geometric series formula we obtain
\[\sum_{j_2 \geq j_1} C_{j_1, j_2} \asymp 2^{j_1(2t_2 - d_2(\alpha_2 + \beta_2 - 1) + 2\ell)}\]
and hence, we have
\[\sum_{j_1 \in \mathbb{N}_0} A_{j_1} C_{j_1} \asymp \sum_{j_1 \in \mathbb{N}_0} 2^{j_1(2t_1 - d_1(\alpha_1 + \beta_1 - 1))} 2^{j_1(2t_2 - d_2(\alpha_2 + \beta_2 - 1) + 2\ell)},\]
which is finite if and only if
\[t_1 + t_2 + \ell < d_1 \left( \frac{\alpha_1 + \beta_1 - 1}{2} \right) + d_2 \left( \frac{\alpha_2 + \beta_2 - 1}{2} \right). \tag{65}\]
To show \(\sum_{j_1 \in \mathbb{N}_0} A_{j_1} B_{j_1} < \infty\) in (63), we again apply the geometric series formula on \(\sum_{j_2 < j_1} B_{j_1, j_2}\), and we are left to determine when
\[\sum_{j_1 \in \mathbb{N}_0} 2^{j_1(2t_1 - d_1(\alpha_1 + \beta_1 - 1) + 2\ell)} - 2^{j_1(2t_1 - d_1(\alpha_1 + \beta_1 - 1) + 2t_2 - d_2(\alpha_2 + \beta_2 - 1) + 2\ell)} < \infty. \tag{66}\]
Summing the difference in formula (66) separately, we see that (66) is finite if and only if (65) and
\[t_1 + \ell < d_1 \left( \frac{\alpha_1 + \beta_1 - 1}{2} \right) \tag{67}\]
hold, or the exponents are both zero or equal. The latter cases yield the condition \(t_2 = d_2(\alpha_2 + \beta_2 - 1)/2\) which contradicts (64). Observe that (64) and (67) imply (65).

Step 3. In order to generalize Step 2 to the case \(n > 2\) and to show (55), observe that the conditions on \(\alpha_m, \beta_m, m = 1, \ldots, n\), such that (61) holds are determined by just iterating the ideas of Step 2. In particular, splitting the appearing sums iteratively as in (63) and proceeding with analogous arguments leads to the conditions (55).

Step 4. Now, we show (56). Using Assumption 5.2 and inserting (53) and (54) into (61) yields that we have to determine when
\[\sum_{j_1 \in \mathbb{N}_0} j_1^{\gamma_1 d_1} 2^{j_1(2t_1 - d_1(\alpha_1 + \beta_1 - 1))} \ldots \sum_{j_n \in \mathbb{N}_0} j_n^{\gamma_n d_n} 2^{j_n(2t_n - d_n(\alpha_n + \beta_n - 1)) + 2\ell} \| (j_1, \ldots, j_n) \|_\infty < \infty.\]
We proceed with the same idea as in Step 3 by iteratively splitting these sums. Then, using that \(\gamma_n d_n < -1\) the sum over \(j_n\), where \(j_n = \| (j_1, \ldots, j_n) \|_\infty\), converges also if we choose the parameters \(\alpha_n, \beta_n\) such that \(t_n + \ell = d_n(\alpha_n + \beta_n - 1)/2\) holds. Using that \(j_n^{\gamma_n d_n} \leq 1\) in all other places it appears, we see that the sum where \(j_{n-1} = \| (j_1, \ldots, j_n) \|_\infty\). Analogously, using \(\gamma_{n-1} d_{n-1} < -1\) this sum converges if we have \(t_{n-1} + \ell = d_{n-1}(\alpha_{n-1} + \beta_{n-1} - 1)/2\). Repeating these arguments yields (56).
Finally we prove (57) provided that (61) is satisfied. Using the Riesz basis property
in Theorem 5.4, as well as Lemma 3.6 and Lemma 3.7 together with the independence
of the $S_{j_m,2}$, $m = 1, \ldots, n$, we have

$$E\left\|X\right\|_{H^{t,\ell}(D)}^2 \lesssim \sum_{j \in \mathbb{N}_0} a_j E\left(\prod_{m=1}^n S_{j_m,2}\right) \leq \sum_{j \in \mathbb{N}_0} a_j \prod_{m=1}^n \# \nabla_{j_m,\rho_{j_m}} < \infty. \quad \Box$$

Remark 5.6. Observe that, in the case $t = 0$, where $H^{0,\ell}(D)$ is isomorphic to the
standard Sobolev space, the conditions (55) and (56) coincide with (19) and (20) for
$p = q = 2$.

Appendix

Proof of Lemma 3.4

On the one hand, let $\sum_{i=1}^\infty X_i < \infty$, P-a.s. With $1_A$ being the index set of $A$ we have

$$\sum_{i=1}^\infty E\left(\frac{X_i}{1+X_i}\right) \leq \sum_{i=1}^\infty E(X_i 1_{\{X_i \leq 1\}} + 1_{\{X_i > 1\}}) = \sum_{i=1}^\infty E(X_i 1_{\{X_i \leq 1\}}) + \sum_{i=1}^\infty P(X_i > 1).$$

Both sums on the right-hand side are finite, due to Kolmogorov’s three-series theorem,
see, e.g., [47]. On the other hand, let $\sum_{i=1}^\infty E\left(\frac{X_i}{1+X_i}\right) < \infty$. Then we get

$$\sum_{i=1}^\infty P(X_i > 1) = 2 \sum_{i=1}^\infty E\left(\frac{1}{2} 1_{\{X_i > 1\}}\right) \leq 2 \sum_{i=1}^\infty E\left(\frac{X_i}{1+X_i} 1_{\{X_i > 1\}}\right) \leq 2 \sum_{i=1}^\infty E\left(\frac{X_i}{1+X_i}\right) < \infty$$

and

$$\sum_{i=1}^\infty E\left(X_i 1_{\{X_i \leq 1\}}\right) = 2 \sum_{i=1}^\infty E\left(\frac{1}{2} X_i 1_{\{X_i \leq 1\}}\right) \leq 2 \sum_{i=1}^\infty E\left(\frac{X_i}{1+X_i} 1_{\{X_i \leq 1\}}\right) \leq 2 \sum_{i=1}^\infty E\left(\frac{X_i}{1+X_i}\right) < \infty.$$

This yields

$$\sum_{i=1}^\infty \text{var}(X_i 1_{\{X_i \leq 1\}}) = \sum_{i=1}^\infty E(X_i^2 1_{\{X_i \leq 1\}}) - \sum_{i=1}^\infty E(X_i 1_{\{X_i \leq 1\}})^2, \quad \leq 2 \sum_{i=1}^\infty E\left(\frac{X_i}{1+X_i}\right) < \infty$$

which is equivalent to $\sum_{i=1}^\infty X_i < \infty$, P-a.s., again due to Kolmogorov’s three-series
theorem.
Proof of Lemma 3.5

For \( r = \frac{s}{t} \) with \( s, t \in \mathbb{N} \) we have

\[
E(X_{n,p}^r) = E(X_{n,p}^{s/t}) \leq (E(X_{n,p}^s))^{1/t}
\]

using Jensen’s inequality. We have \( E(X_{1,p}^s) = p \leq 1 \), and for \( n \geq 2 \)

\[
E(X_{n,p}^s) = \sum_{k=1}^{n} k^s \binom{n}{k} p^k (1 - p)^{n-k}
= \sum_{k=0}^{n-1} (k+1)^{s-1} n \binom{n-1}{k} p^{k+1} (1 - p)^{n-1-k}
= np \ E(1 + X_{n-1,s})^{s-1}.
\]

Inductively over \( s \), we obtain

\[
E(X_{n,p}^s) \leq np \ (E(1) + E(X_{n-1,s})^{s-1}) \leq \tilde{c} (np + (np)^s) \leq c (1 + (np)^s),
\]

and for \( r \in \mathbb{Q} \) this leads to

\[
E(X_{n,p}^r) \leq np \ (E(1) + E(X_{n-1,s})^{s-1})^{1/t} \leq \tilde{c} (np + (np)^s)^{1/t} \leq c (1 + (np)^r).
\]

Using the density of \( \mathbb{Q} \) in \( \mathbb{R} \) we get the result for all \( r \in \mathbb{R} \). \( \square \)

Proof of Lemma 3.6

It is \( E(S_{j,p}) = \#\nabla_j \rho \nu_p \) and \( \text{Var}(S_{j,p}) = \#\nabla_j \nu_p (\nu_{2p} - \rho_j \nu_p^2) \). Using Chebyshev’s inequality we get

\[
P(|S_{j,p}/(\#\nabla_j \rho_j) - \nu_p| \geq \varepsilon) \leq \varepsilon^{-2}(\#\nabla_j \rho_j)^{-1}(\nu_{2p} - \rho_j \nu_p^2) \leq c (\varepsilon^{-2(1-\beta)j_d}).
\]

Since \( \beta < 1 \), applying the Borel-Cantelli Lemma yields almost sure convergence. Let \( r > 0 \) and \( y_{j,k} \in \{0, 1\} \). By the equivalence of moments of Gaussian measures there exists a constant \( c_1 > 0 \) such that

\[
E\left(\sum_{k \in \nabla_j} |y_{j,k} Z_{j,k}|^r\right)^\rho_p \leq c_1 \left(\sum_{k \in \nabla_j} |y_{j,k} Z_{j,k}|^p\right)^\rho_p = \sum_{k \in \nabla_j} c_1 \nu_p\left(\sum_{k \in \nabla_j} y_{j,k}\right)^r.
\]

The sequences \( (Y_{j,k})_k \) and \( (Z_{j,k})_k \) are independent. This yields \( E(S_{j,p}^r) \leq c_1 (\nu_p)^r E(S_{j,0}^r) \). Using Lemma 3.5 there exists a constant \( c_2 > 0 \) such that \( E(S_{j,0}^r) \leq c_2 (\#\nabla_j \rho_j)^r \). \( \square \)

Proof of Lemma 3.7

Let \( Z \) be \( N(0,1) \)-distributed. The characteristic function \( \varphi_{S_p} \) of \( S_p \) is given by

\[
\varphi_{S_p}(t) = E(\exp(it S_p)) = \exp(C_0(\varphi_{|Z|^p}(t) - 1)).
\]

Furthermore, for the characteristic function \( \varphi_{S_{j,p}} \) of \( S_{j,p} \),

\[
\varphi_{S_{j,p}}(t) = (\rho_j \varphi_{|Z|^p}(t) + 1 - \rho_j)^{\#\nabla_j}
\]

29
\[
(1 + \frac{1}{\#\nabla_j} \cdot \rho_j \#\nabla_j (\varphi_j|Z|^p(t) - 1))^{\#\nabla_j}.
\]

We use (11) to conclude that

\[
\lim_{j \to \infty} \varphi_{S_{j,p}}(t) = \varphi_{S_p}(t),
\]

which yields the convergence in distribution as claimed. Suppose that \( p \geq 1 \) and \( r > 0 \). Then we take \( c_1 > 0 \) such that \( z^r \leq c_1 \exp(z) \) for every \( z \geq 0 \) and we put \( c_2 = \mathbb{E}(\exp(|Z|)) \) to obtain

\[
\mathbb{E}(S^*_j,p) \leq \mathbb{E}(S^*_{j,1}) \leq c_1 (1 + \rho_j(c_2 - 1))^{\#\nabla_j}.
\]

Note that the upper bound converges to \( c_1 \exp(C_0(c_2 - 1)) \). In the case \( 0 < p < 1 \) we have \( S_{j,p} \leq S_{j,0} + S_{j,1} \). Hence it remains to observe that \( \sup_{j \geq j_0} \mathbb{E}(S^*_{j,0}) < \infty \), which follows from Lemma 3.5. \( \square \)

References


Stephan Dahlke, Stefan Kinzel
Philipps-Universität Marburg
FB Mathematik und Informatik, AG Numerik/Optimierung
Hans-Meerwein-Strasse
35032 Marburg, Germany
{dahlke, kinzel}@mathematik.uni-marburg.de

Nicolas Döhring
TU Kaiserslautern
Fachbereich Mathematik, AG Computational Stochastics
Postfach 3049
67653 Kaiserslautern, Germany
doehring@mathematik.uni-kl.de
Reports


