STUDIA MATHEMATICA

Online First version

Besov regularity of random wavelet series

by

Andreas Horst, Thomas Jahn and Felix Voigtlaender

Abstract. We study the Besov regularity of wavelet series on $\mathbb{R}^d$ with randomly chosen coefficients. More precisely, each coefficient is a product of a random factor and a parameterized deterministic factor that decays (or grows) with the scale j and the norm of the shift m. Compared to the literature, we impose relatively mild conditions on the moments of the random variables in order to characterize the almost sure convergence of the wavelet series in Besov spaces $B^s_{p,q}(\mathbb{R}^d)$ and the finiteness of the moments as well as of the moment generating function of the Besov norm. In most cases, we achieve a complete characterization, i.e., the conditions derived are both necessary and sufficient.

1. Introduction. We study the Besov regularity of functions $f : \mathbb{R}^d \to \mathbb{R}$ (or tempered distributions $f \in \mathcal{S}'(\mathbb{R}^d)$) given by Karhunen–Loève-type expansions of the form

$$f = \sum_{j,t,m} 2^{j\alpha}(j+1)^{\theta}(1+\|m\|_\infty/2^j)^{\beta}\xi_{j,t,m}\Psi_{j,t,m}, \tag{1.1}$$

$$f = \sum_{j,t,m} 2^{j\alpha}(1+\|m\|_\infty/2^j)^{\beta}\lambda_{j,t,m}\xi_{j,t,m}\Psi_{j,t,m} \tag{1.2}$$

with respect to a sufficiently "nice" compactly supported orthonormal wavelet basis $(\Psi_{j,t,m})_{j,t,m}$ of $L_2(\mathbb{R}^d)$. The functions given by (1.1) and (1.2) are deemed *random* as the factors $\xi_{j,t,m}$ are chosen to be independent and identically distributed random variables, and the factors $\lambda_{j,t,m}$ are Bernoulli-distributed random variables with success probabilities depending on j and m. The functions in (1.2) can be seen as a model for random *sparse* wavelet series. (A similar approach for periodic wavelet series has been investigated in [1].) In addition to purely mathematical interest, studying such random wavelet series is motivated by their use to define prior distributions in Bayesian inversion [19, 8]; see also Section 1.1.

2020 *Mathematics Subject Classification*: Primary 42B35; Secondary 42C40, 60G17, 60G60, 60H50.

Key words and phrases: Besov space, Karhunen–Loève-type expansion, prior, wavelet.

Received 5 December 2024; revised 8 October 2025.

Published online 29 October 2026.

 © Instytut Matematyczny PAN, 2026

Our first main result is to show that under mild conditions on the moments of the random variables $\xi_{j,t,m}$ and on the regularity of the wavelet orthonormal basis Ψ, the series (1.1) almost surely belongs to the Besov space $B^s_{p,q}(\mathbb{R}^d)$ *if and only if* the parameters s, p, q of the Besov space and the parameters α, β, and θ of the prior (1.1) satisfy

$$\beta < -d/p, \quad s + d/2 + \alpha < 0 \tag{1.3}$$

or

$$\beta < -d/p, \quad s + d/2 + \alpha = 0, \quad \theta < -1/q; \tag{1.4}$$

see Theorems 3.3, 4.1 and 4.3. (When $p = \infty$, the condition $\beta < -d/p$ has to be replaced by $\beta \leq 0$. When $q = \infty$, the condition $\theta < -1/q$ has to be replaced by $\theta \leq 0$.) A similar analysis for the series in (1.2) is given in Theorems 3.8, 4.4 and 4.5. Furthermore, given $r > 0$ and assuming that (1.3) or (1.4) holds, we show that under (different) mild conditions on the common distribution of the random variables $\xi_{j,t,m}$, the moments $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}]$ and the moment generating function $\mathbb{E}[\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)})]$ are finite (for some $c > 0$); see Theorems 3.4 and 3.5. Thus, somewhat surprisingly, the condition that guarantees that $\|f\|_{B^s_{p,q}(\mathbb{R}^d)}$ is finite almost surely already implies the finiteness of $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}]$ or of $\mathbb{E}[\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)})]$, if only the common distribution of the random variables $\xi_{j,t,m}$ is sufficiently "nice". A similar analysis for the moments $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}]$ of the series in (1.2) is given in Theorem 3.8.

As $c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)} \leq \exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)})$, the condition $\mathbb{E}[\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)})] < \infty$ implies $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}] < \infty$, which in turn implies $\|f\|_{B^s_{p,q}(\mathbb{R}^d)} \overset{\text{a.s.}}{<} \infty$. Therefore, necessary conditions for the latter are also necessary for the former. In other words, if f is of the form (1.1) and the distribution of the random variables $\xi_{j,t,m}$ is sufficiently well concentrated, we achieve a *complete characterization* in terms of the parameters d, s, p, q, and α, β, θ for each of the following conditions:

- $f \in B^s_{p,q}(\mathbb{R}^d)$ almost surely,
- $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}] < \infty$,
- $\mathbb{E}[\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)})] < \infty$ for some $c > 0$.

1.1. Motivation and related work. While our investigations in the present paper are of theoretical nature, they are motivated by the Bayesian approach to inverse problems, where random Fourier series and random wavelet series are used as prior distributions for a stochastic interpretation of inverse problems. There, the finiteness of the exponential moments plays a role in the study of discretization invariance and well-posedness of the

Bayesian inverse problems; see [19, 8] and in particular [19, Theorem 1]. Lassas, Saksman, and Siltanen [19, Lemma 2] prove that for $s \in \mathbb{R}$ and $p \in [1, \infty)$, the series $f(x) = \sum_{\ell=1}^{\infty} \ell^{-(s/d+1/2-1/p)} \xi_\ell \Psi_\ell(x)$ almost surely belongs to $B^{\tilde{s}}_{p,p}(\mathbb{T}^d)$ iff $\mathbb{E}\big[\exp\big(\frac{1}{2}\|f\|^p_{B^{\tilde{s}}_{p,p}(\mathbb{T}^d)}\big)\big] < \infty$ iff $\tilde{s} < s - d/p$. Here $(\Psi_j)_{j\in\mathbb{N}}$ denotes an enumeration (by increasing scale) of a compactly supported orthonormal wavelet basis of $L_2(\mathbb{T}^d)$ on the torus $\mathbb{T}^d$, and $(\xi_j)_{j\in\mathbb{N}}$ denotes a family of i.i.d. real-valued random variables with probability density function proportional to $\exp(-|x|^p)$. The series $f = \sum_{\ell=1}^{\infty} \ell^{-(s/d+1/2-1/p)} \xi_\ell \psi_\ell$ induces a probability measure on $B^{\tilde{s}}_{p,p}(\mathbb{T}^d)$ and is referred to as a *Besov prior* in the inverse problem community.

The Besov prior has been extended in several works by introducing an additional random coefficient into the wavelet expansion of f. For example, the authors of [16] choose the indices of the nonzero coefficients in the wavelet expansion according to a proper random tree generated by a Galton–Watson process, and in [18], time-dependent random coefficients from a so called q-exponential process are investigated. In [16, Theorem 3.3] and [18, Theorem 4.2], these authors prove that their priors almost surely belong to some Besov space depending on newly introduced parameters.

Apart from the applications in Bayesian inversion, wavelet series with randomly chosen coefficients have been studied in the literature for more than two decades now. Aubry and Jaffard [3] study the *local* Hölder regularity of random wavelet series. Because of the local nature of their analysis, the authors focus on the one-dimensional 1-periodic case, i.e., functions $f : \mathbb{T} \to \mathbb{R}$ defined on the torus.

Besov regularity enters the discussion in [3] through the *scaling function* $\eta_f(p) := \sup\{s : f \in B^{s/p}_{p,\infty}(\mathbb{T})\}$, a function which is shown in [3] to be closely related to histograms of wavelet coefficients at all scales. Note that in contrast to the setting of [19], the parameter p in [3] is an element of $(0, \infty)$.

Abramovich, Sapatinas, and Silverman [1] investigate sparse random wavelet expansions on the interval $[0, 1]$. Specifically, the wavelet coefficients $d_{j,k}$ are chosen at random according to $d_{j,k} \sim (1 - \pi_j)\delta_0 + \pi_j \mathcal{N}(0, \tau_j^2)$; see [1, Eq. (6)]. Here, $\pi_j \in [0, 1]$ and $\tau_j > 0$ are numbers, $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution with mean μ and variance σ^2, and δ_0 is the Dirac measure at 0. Special attention is paid to the case where $\pi_j = \min\{1, 2^{-j\beta} C_2\}$ and $\tau_j^2 = 2^{-j\alpha} C_1$ for some absolute constants C_1 and C_2, and parameters α and β, which are then subject to the characterization of the almost sure membership of the corresponding wavelet series in the Besov space $B^s_{p,q}([0, 1])$; see [1, Theorem 1 and Equation (9)]. Bochkina [6] extends this setting to more general distributions of the random variables.

For Besov spaces $B^s_{p,q}(Q)$ with parameters $s \in \mathbb{R}$ and $p, q \in (0, \infty)$ on bounded Lipschitz domains $Q \subset \mathbb{R}^d$ which admit an orthonormal basis

of compactly supported wavelets, the authors of [7], similar to [1], choose products of Bernoulli variables (with success probabilities decaying exponentially in j) and Gaussians for the coefficients of a wavelet series. In [7, Theorem 5.7], a characterization of the almost sure membership of such a wavelet series g in $B^s_{p,q}(Q)$ is given in terms of inequalities involving d, s, p, q, and the parameters governing the random variables. Furthermore, the same inequalities are shown to be sufficient for $\mathbb{E}[\|g\|^q_{B^s_{p,q}(Q)}] < \infty$.

Another account on random wavelet series on $\mathbb{R}^d$ is given by Aziznejad and Fageot [4, Proposition 3.4]. There it is shown that if the wavelet coefficients of f are i.i.d. standard Gaussian random variables (which corresponds to $\alpha = \beta = \theta = 0$ and $\xi_{j,t,m} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ in (1.1)) and $p \in (0,\infty]$, then f almost surely belongs to a weighted Besov space $B^s_{p,p}(\mathbb{R}^d, w_\sigma)$ if $s < -d/2$ and $\sigma < -d/p$, and almost surely does not belong to $B^s_{p,p}(\mathbb{R}^d, w_\sigma)$ if $s \geq -d/2$ or $\sigma \geq -d/p$; see Remark 2.7 for a definition of w_σ.

Esser, Jaffard, and Vedel [10] study regularity properties of wavelet series through randomization. More precisely, for wavelet bases on the d-dimensional torus $\mathbb{T}^d$ and on $\mathbb{R}^d$, they compare the Hölder continuity and boundedness properties of a wavelet series before and after multiplying every coefficient with a member of a family of i.i.d. random variables. Their analysis shows that regarding the conditions on the random variables, there is an essential distinction between bounded and unbounded random variables.

In the paper at hand, we also provide necessary and sufficient conditions for random wavelet series to almost surely belong to a given Besov space (or for its Besov quasi-norm to have finite moments and finite moment-generating function.) But, different from the existing literature, we simultaneously allow the following conditions: (i) we allow much more general distributions of the random coefficients, only imposing mild conditions such as finiteness of certain moments or of the moment generating function of certain powers of $|\xi_{j,t,m}|$, (ii) we allow Besov spaces $B^s_{p,q}$ with $p \neq q$, (iii) we work on $\mathbb{R}^d$ instead of $\mathbb{T}^d$, and (iv) we consider the full range of exponents $p, q \in (0,\infty]$.

1.2. Structure of the paper. Section 2 is a review of the necessary preliminaries on Besov spaces and their characterization by wavelets. Moreover, it introduces the Besov priors in (1.1) and (1.2) more formally and establishes a convenient terminology regarding conditions on the various parameters.

In Section 3, we show that under suitable conditions on the random variables $\xi_{j,t,m}$, the conditions (1.3) and (1.4) are *sufficient* for having $f \in B^s_{p,q}(\mathbb{R}^d)$ almost surely and $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}] < \infty$ when f is of the form (1.1). Additionally, we show the sufficiency of an analog of (1.3) and (1.4) for

having $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}] < \infty$ for the series (1.2). Section 3 is concluded with the proof of the sufficiency of (1.3) and (1.4) for having $\mathbb{E}[\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)})] < \infty$ for $c, r > 0$ when f is of the form (1.1).

In Section 4, we show that (1.3) and (1.4) are *necessary* for the series in (1.1) to almost surely belong to the Besov space $B^s_{p,q}(\mathbb{R}^d)$, given a condition on the moments of the random variables $\xi_{j,t,m}$ in (1.1). Additionally, we show the necessity of an analog of (1.3) and (1.4) for the series in (1.2) to almost surely belong to the Besov space $B^s_{p,q}(\mathbb{R}^d)$.

In Section 5, we discuss the necessity of the conditions on the random variables in the theorems about the series (1.1).

Finally, Appendix A contains auxiliary statements and technical proofs.

2. Preliminaries and notation

2.1. General notation. For two sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ of non-negative real numbers, we write $a_n \lesssim b_n$ if there exists a constant $C > 0$ such that $a_n \le Cb_n$ for all $n \in \mathbb{N}$. We write $a_n \gtrsim b_n$ if $b_n \lesssim a_n$, and $a_n \asymp b_n$ if both $a_n \lesssim b_n$ and $b_n \lesssim a_n$. Unless mentioned explicitly, the symbol log refers to the natural logarithm. We write $\log_2^+(x) := \max\{0, \log_2(x)\}$ for $x > 0$.

For an at most countable set $I \neq \emptyset$, $x \in \mathbb{R}^I$, and $p \in (0, \infty]$, we write

$$\|x\|_{\ell_p(I)} := \begin{cases} (\sum_{i\in I} |x_i|^p)^{1/p} & \text{if } 0 < p < \infty, \\ \sup_{i\in I} |x_i| & \text{if } p = \infty. \end{cases}$$

We shall usually omit I from the notation, as it will be clear from the context. When $I = \{1, \dots, d\}$, we write $\|x\|_p := \|x\|_{\ell_p(I)}$.

For a (measurable) function $f : \mathbb{R}^d \to \mathbb{R}$, we write

$$\|f\|_{L_p(\mathbb{R}^d)} := \begin{cases} (\int_{\mathbb{R}^d} |f(x)|^p \,\mathrm{d}x)^{1/p} & \text{if } 0 < p < \infty, \\ \operatorname{ess\,sup}_{x\in\mathbb{R}^d} |f(x)| & \text{if } p = \infty. \end{cases}$$

For $\gamma = (\gamma_1, \dots, \gamma_d) \in \mathbb{N}_0^d$ and sufficiently smooth $f : \mathbb{R}^d \to \mathbb{R}$, we denote the partial derivatives of f by $D^\gamma f = \frac{\partial^{\gamma_1+\dots+\gamma_d}}{\partial x_1^{\gamma_1} \dots \partial x_d^{\gamma_d}} f$.

2.2. Wavelets and Besov spaces. Our discussion of wavelet series and their Besov regularity relies on the wavelet characterization of Besov spaces. We present a brief outline of this characterization here; we follow [22, Chapter 1] for the definitions of the Fourier transform and (weighted) Besov (sequence) spaces. We will not go into the minute details in what sense the occurring integrals are well-defined, how the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ and its dual space $\mathcal{S}'(\mathbb{R}^d)$ are defined, and in which way the Fourier transform is extended from $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}'(\mathbb{R}^d)$; see [20, Section 2.3] for these details.

Instead, our aim is to set the notation and to avoid ambiguities regarding different conventions in the literature.

The *Fourier transform* $\hat{f}$ of $f \in \mathcal{S}(\mathbb{R}^d)$ is defined as

$$\hat{f}(x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}x^\top y} f(y)\,\mathrm{d}y$$

for $x \in \mathbb{R}^d$. Here $x^\top y$ denotes the standard inner product of $x, y \in \mathbb{R}^d$. The inverse of the Fourier transform shall be denoted by $\check{f}$ or $f^\vee$; it is given by $\check{f}(x) = \hat{f}(-x)$.

Definition 2.1. Let $d \in \mathbb{N}$, $p, q \in (0, \infty]$, $s \in \mathbb{R}$. An *admissible weight* [22, Definition 1.22] is a function $w \in C^\infty(\mathbb{R}^d, \mathbb{R})$ such that for every multi-index $\gamma \in \mathbb{N}_0^d$, there exists a constant $c_\gamma > 0$ such that the partial derivative $D^\gamma w$ of w satisfies

$$|D^\gamma w(x)| \leq c_\gamma w(x) \quad \text{for all } x \in \mathbb{R}^d,$$

and there exist $c > 0$ and $\alpha \geq 0$ such that

$$0 < w(x) \leq c w(y)(1 + \|x - y\|_2^2)^{\alpha/2} \quad \text{for all } x, y \in \mathbb{R}^d.$$

Let w be an admissible weight as above. If $(\varphi_j)_{j \in \mathbb{N}_0}$ is a suitable resolution of unity, cf. [22, (1.5), (1.6), (1.7)], then the *(weighted) Besov space* $B^s_{p,q}(\mathbb{R}^d, w)$ is defined as the collection of all $f \in \mathcal{S}'(\mathbb{R}^d)$ such that

$$\|f\|_{B^s_{p,q}(\mathbb{R}^d, w)} := \Big(\sum_{j \in \mathbb{N}_0} 2^{jsq} \|w \cdot (\varphi_j \hat{f})^\vee\|^q_{L_p(\mathbb{R}^d)} \Big)^{1/q} < \infty,$$

with the usual modifications for $q = \infty$; see [22, Definition 1.22]. (This definition makes sense since $\varphi_j \hat{f} \in \mathcal{S}'(\mathbb{R}^d)$ is a compactly supported distribution so that $(\varphi_j \hat{f})^\vee \in C^\infty(\mathbb{R}^d)$ is a (smooth) function by the Paley–Wiener theorem.)

Define $T_0 := \{F\}^d$, $T_j := \{F, M\}^d \setminus T_0$ for $j \in \mathbb{N}$,

$$J := \{(j, t, m) : j \in \mathbb{N}_0,\, t \in T_j,\, m \in \mathbb{Z}^d\},$$

and for a sequence $\mathbf{a} = (a_{j,t,m})_{(j,t,m) \in J} \in \mathbb{R}^J$, following [22, Definition 1.24] with a slight modification, we let

$$\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d, w)} := \Big\| \big(2^{j(s + d/2 - d/p)} \big\| \big(w(2^{-(j - \mathbb{1}_{j \neq 0})} m) |a_{j,t,m}| \big)_{m \in \mathbb{Z}^d} \big\|_{\ell_p} \big)_{j \in \mathbb{N}_0, t \in T_j} \Big\|_{\ell_q},$$

where $\mathbb{1}_{j \neq 0} = 1$ if $j \neq 0$ and $\mathbb{1}_{j \neq 0} = 0$ if $j = 0$.

The (*weighted*) *Besov sequence space* $b^s_{p,q}(\mathbb{R}^d, w)$ is defined as the collection of all sequences $(a_{j,t,m})_{(j,t,m) \in J} \in \mathbb{R}^J$ such that $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d, w)} < \infty$. We write $B^s_{p,q}(\mathbb{R}^d) := B^s_{p,q}(\mathbb{R}^d, w)$ and $b^s_{p,q}(\mathbb{R}^d) := b^s_{p,q}(\mathbb{R}^d, w)$ when $w(x) = 1$ for all $x \in \mathbb{R}^d$.

It turns out that $(B^s_{p,q}(\mathbb{R}^d, w), \|\cdot\|_{B^s_{p,q}(\mathbb{R}^d,w)})$ and $(b^s_{p,q}(\mathbb{R}^d, w), \|\cdot\|_{b^s_{p,q}(\mathbb{R}^d,w)})$ are quasi-Banach spaces, and different choices of the resolution of unity $(\varphi_j)_{j\in\mathbb{N}_0}$ lead to equivalent quasi-norms [14, Section 2.2].

The characterization of Besov spaces in terms of wavelet series is usually done using compactly supported wavelets. We cite the existence result for such "nice" compactly supported wavelets (which originally is due to Daubechies [9]) from [13, Theorem 10].

THEOREM 2.2. *For each $k \in \mathbb{N}$, there exist functions $\psi_F, \psi_M \in L_2(\mathbb{R})$, called* scaling function *and* wavelet function*, respectively, such that*

(i) *ψ_F and ψ_M are real-valued and compactly supported,*
(ii) *ψ_F, $\psi_M \in C^k(\mathbb{R})$,*
(iii) *$\widehat{\psi_F}(0) = (2\pi)^{-1/2}$,*
(iv) *$\int_{\mathbb{R}} x^\ell \psi_M(x)\,\mathrm{d}x = 0$ for all $\ell \in \{0, \dots, k\}$,*
(v) *The one-dimensional inhomogeneous wavelet system $(\psi_{j,m})_{j\in\mathbb{N}_0,\, m\in\mathbb{Z}}$ defined by*

$$\psi_{0,m} = \psi_F(\cdot - m), \qquad \psi_{j,m}(x) = 2^{(j-1)/2}\psi_M(2^{j-1}x - m) \quad \text{for } j \in \mathbb{N}$$

is an orthonormal basis of $L_2(\mathbb{R})$.

The (one-dimensional) scaling function ψ_F and wavelet function ψ_M from Theorem 2.2 are then used to define an orthonormal basis $(\Psi_{j,t,m})_{(j,t,m)\in J}$ of $L_2(\mathbb{R}^d)$ via

(2.1)
$$\Psi_{j,t,m} : \mathbb{R}^d \to \mathbb{R}, \quad x \mapsto \begin{cases} \prod_{k=1}^d \psi_F(x_k - m_k) & \text{if } j = 0, \\ 2^{(j-1)d/2} \prod_{k=1}^d \psi_{t_k}(2^{j-1}x_k - m_k) & \text{if } j \in \mathbb{N}, \end{cases}$$

for $(j, t, m) \in J$ [21, Proposition 1.53].

It turns out that mapping the coefficient sequence to the corresponding wavelet expansion constitutes an isomorphism between the Besov sequence space and the corresponding Besov function space. This is the above-mentioned wavelet characterization of Besov spaces. The following theorem makes this precise; it is taken from [22, Theorem 1.26]. We note that unlike Triebel [22], we use the L_2-normalized wavelet system $(\Psi_{j,t,m})_{(j,t,m)\in J}$ instead of the L_∞-bounded wavelet system $(2^{-jd/2}\Psi_{j,t,m})_{(j,t,m)\in J}$ used in [22]. This is compensated by the factor $2^{jd/2}$ in the definition of $b^s_{p,q}(\mathbb{R}^d)$ which does not occur in [22]; see also Remark 2.4.

THEOREM 2.3. *Let $d \in \mathbb{N}$, $p, q \in (0, \infty]$, $s \in \mathbb{R}$. Denote by $(\Psi_{j,t,m})_{(j,t,m)\in J}$ the orthonormal basis of $L_2(\mathbb{R}^d)$ given in* (2.1).

Assume that $k > \max\{s, d(1/p-1)-s\}$ *if* $p \in (0,1)$ *and* $k > |s|$ *else. Let* $f \in \mathcal{S}'(\mathbb{R}^d)$*. Then* $f \in B^s_{p,q}(\mathbb{R}^d, w)$ *if and only if it can be represented as*

$$f = \sum_{(j,t,m)\in J} a_{j,t,m}\Psi_{j,t,m} \tag{2.2}$$

with $\mathbf{a} = (a_{j,t,m})_{(j,t,m)\in J} \in b^s_{p,q}(\mathbb{R}^d, w)$*. The convergence is unconditional in* $\mathcal{S}'(\mathbb{R}^d)$*. Furthermore, the representation is unique, i.e., for given* $f \in B^s_{p,q}(\mathbb{R}^d, w)$*, the unique element* $(a_{j,t,m})_{(j,t,m)\in J} \in b^s_{p,q}(\mathbb{R}^d, w)$ *such that* (2.2) *holds is given by* $a_{j,t,m} = \langle f, \Psi_{j,t,m}\rangle$*, where the duality pairing is suitably defined. The map*

$$I : f \mapsto (\langle f, \Psi_{j,t,m}\rangle)_{(j,t,m)\in J}$$

defines an isomorphism of quasi-Banach spaces $B^s_{p,q}(\mathbb{R}^d, w) \to b^s_{p,q}(\mathbb{R}^d, w)$*, and the map*

$$S : (a_{j,t,m})_{(j,t,m)\in J} \mapsto \sum_{(j,t,m)\in J} a_{j,t,m}\Psi_{j,t,m}$$

defines an isomorphism of quasi-Banach spaces $b^s_{p,q}(\mathbb{R}^d, w) \to B^s_{p,q}(\mathbb{R}^d, w)$*.*

Remark 2.4. Let us comment on the notational differences to [22]. Triebel [22, (1.89) and (1.90)] uses different index sets for the wavelet system, namely $G_0 := \{F, M\}^d$, $G_j := \{F, M\}^d \setminus \{F\}^d$. The wavelet system [22, (1.91)] then takes the form

$$\widetilde{\Psi}_{j,t,m}(x) = 2^{jd/2} \prod_{k=1}^{d} \psi_{t_k}(2^j x_k - m_k).$$

If we let $\widetilde{J} := \{(j,t,m) : j \in \mathbb{N}_0,\, t \in G_j,\, m \in \mathbb{Z}^d\}$, we note that Triebel's definition

$$\|\mathbf{a}\|_{\widetilde{b^s_{p,q}}(\mathbb{R}^d,w)} = \Big(\sum_{j\in\mathbb{N}_0} 2^{jq(s-d/p)} \sum_{t\in G_j} \Big(\sum_{m\in\mathbb{Z}^d} w(2^{-j}m)^p |a_{j,t,m}|^p\Big)^{q/p}\Big)^{1/q}$$

of the Besov sequence quasi-norm of $\mathbf{a} = (a_{j,t,m})_{(j,t,m)\in\widetilde{J}}$ does not involve a factor of $2^{jd/2}$ (see [22, (1.119)]), but the wavelet characterization in [22, Theorem 1.26] does:

$$\|f\|_{B^s_{p,q}(\mathbb{R}^d,w)} \asymp \|(2^{jd/2}\langle f, \widetilde{\Psi}_{j,t,m}\rangle)_{(j,t,m)\in\widetilde{J}}\|_{\widetilde{b^s_{p,q}}(\mathbb{R}^d,w)}.$$

But note that we also have

$$\|(2^{jd/2}\langle f, \widetilde{\Psi}_{j,t,m}\rangle)_{(j,t,m)\in\widetilde{J}}\|_{\widetilde{b^s_{p,q}}(\mathbb{R}^d,w)} \asymp \|(\langle f, \Psi_{j,t,m}\rangle)_{(j,t,m)\in J}\|_{b^s_{p,q}(\mathbb{R}^d,w)}.$$

2.3. Besov priors and Besov sequence priors. Once we fix the orthonormal basis $(\Psi_{j,t,m})_{(j,t,m)\in J}$ of $L_2(\mathbb{R}^d)$, we may consider random wavelet series f as in (1.1). In view of the application in Bayesian inverse problems as a prior distribution, we then say that f follows a Besov prior. Note that

the following definition is slightly more general than (1.1), since we allow the additional parameter γ.

DEFINITION 2.5. Let $d, k \in \mathbb{N}$, $\alpha, \beta, \gamma, \theta \in \mathbb{R}$, $\mu, \nu \in (-\infty, 0]$, and let X be a (real-valued) random variable.

(a) We say that f *follows a Besov prior with parameters* $(\alpha, \beta, \gamma, \theta, k)$ *and template random variable* X if the orthonormal basis $(\Psi_{j,t,m})_{(j,t,m)\in J}$ of $L_2(\mathbb{R}^d)$ is the one from (2.1) (with $\psi_F, \psi_M \in C^k(\mathbb{R})$ and ψ_M having k vanishing moments as in Theorem 2.2(iv)) and there are i.i.d. random variables $\xi_{j,t,m} \sim X$ for $(j,t,m) \in \mathbb{N}_0 \times \{F,M\}^d \times \mathbb{Z}^d$ such that

$$f = \sum_{(j,t,m)\in J} 2^{j\alpha}(j+1)^{\theta}(1+\|m\|_\infty/2^j)^{\beta+(\gamma-\beta)\mathbb{1}_{j=0}}\xi_{j,t,m}\Psi_{j,t,m}, \tag{2.3}$$

with almost sure unconditional convergence of the series in the sense of tempered distributions.

(b) We shall say that f *follows a Besov–Bernoulli prior with parameters* $(\alpha, \beta, \gamma, \mu, \nu, k)$ *and template random variable* X if the orthonormal basis $(\Psi_{j,t,m})_{(j,t,m)\in J}$ of $L_2(\mathbb{R}^d)$ is the one from (2.1) (with $\psi_F, \psi_M \in C^k(\mathbb{R})$ and ψ_M having k vanishing moments), there are i.i.d. random variables $\xi_{j,t,m} \sim X$ for $(j,t,m) \in \mathbb{N}_0 \times \{F,M\}^d \times \mathbb{Z}^d$, and Bernoulli$(\varrho_{j,m})$-distributed random variables $\lambda_{j,t,m}$ for $(j,t,m) \in \mathbb{N}_0 \times \{F,M\}^d \times \mathbb{Z}^d$, independent among each other and from all the $\xi_{j,t,m}$, with

$$\varrho_{j,m} = 2^{j\mu}(1+\|m\|_\infty/2^j)^{\nu}$$

such that

$$f = \sum_{(j,t,m)\in J} 2^{j\alpha}(1+\|m\|_\infty/2^j)^{\beta+(\gamma-\beta)\mathbb{1}_{j=0}}\lambda_{j,t,m}\xi_{j,t,m}\Psi_{j,t,m}, \tag{2.4}$$

with almost sure unconditional convergence of the series (2.4) in the sense of tempered distributions.

Here $\mathbb{1}_{j=0} := 1$ if $j = 0$, and $\mathbb{1}_{j=0} := 0$ else. Furthermore, if $j = 0$ for a summand in the expansion (2.3), then also $t = (F, \dots, F)$ is fixed, and in that case we shorten notation by omitting both indices j, t and only keeping the relevant index $m \in \mathbb{Z}^d$, i.e., we write $\xi_m := \xi_{0,(F,\dots,F),m}$ and $\lambda_m := \lambda_{0,(F,\dots,F),m}$.

The analysis for the functions/distributions and for their coefficient sequences is not entirely the same; see the discussion in Remark 4.2. Therefore it is appropriate to introduce analogous notions for the coefficient sequences. The notational conventions from Definition 2.5 apply.

DEFINITION 2.6. Let $d \in \mathbb{N}$, $\alpha, \beta, \gamma, \theta \in \mathbb{R}$, $\mu, \nu \in (-\infty, 0]$, and let X be a (real-valued) random variable.

(a) We say that $\mathbf{a} = (a_{j,t,m})_{(j,t,m)\in J} \in \mathbb{R}^J$ *follows a Besov sequence prior with parameters* $(\alpha, \beta, \gamma, \theta)$ *and template random variable* X if there are i.i.d. random variables $\xi_{j,t,m} \sim X$ for $(j,t,m) \in \mathbb{N}_0 \times \{F,M\}^d \times \mathbb{Z}^d$ such that

$$a_{j,t,m} = 2^{j\alpha}(j+1)^{\theta}(1 + \|m\|_\infty/2^j)^{\beta+(\gamma-\beta)\mathbb{1}_{j=0}}\xi_{j,t,m}. \tag{2.5}$$

(b) We say that $\mathbf{a} = (a_{j,t,m})_{(j,t,m)\in J} \in \mathbb{R}^J$ *follows a Besov–Bernoulli sequence prior with parameters* $(\alpha, \beta, \gamma, \mu, \nu)$ *and template random variable* X if there are i.i.d. random variables $\xi_{j,t,m} \sim X$ for $(j,t,m) \in \mathbb{N}_0 \times \{F,M\}^d \times \mathbb{Z}^d$ and Bernoulli$(\varrho_{j,m})$-distributed random variables $\lambda_{j,t,m}$ for $(j,t,m) \in \mathbb{N}_0 \times \{F,M\}^d \times \mathbb{Z}^d$, independent of each other and of all the $\xi_{j,t,m}$, with

$$\varrho_{j,m} = 2^{j\mu}(1 + \|m\|_\infty/2^j)^{\nu}$$

such that

$$a_{j,t,m} = 2^{j\alpha}(1 + \|m\|_\infty/2^j)^{\beta+(\gamma-\beta)\mathbb{1}_{j=0}}\lambda_{j,t,m}\xi_{j,t,m}. \tag{2.6}$$

Note that in Definitions 2.5 and 2.6, we assume to have random variables $\xi_{j,t,m}$ and $\lambda_{j,t,m}$ for $(j,t,m) \in \mathbb{N}_0 \times \{F,M\}^d \times \mathbb{Z}^d$, although only those for $(j,t,m) \in J$ are "used" in (2.3)–(2.6). This is done for convenience in the proofs. In our analysis, expectations $\mathbb{E}$ and variances $\mathbb{V}$ are taken with respect to these random variables $\xi_{j,t,m}$ and $\lambda_{j,t,m}$.

Remark 2.7 and Figure 2 show samples of the two priors defined in Definition 2.5 for different sets of parameters. (A Python implementation for generating these samples can be found at https://github.com/AndreasHorst/RWS.) In both figures, we used standard normal variables for $\xi_{j,t,m}$ and Daubechies-10 wavelets for $\Psi_{j,t,m}$. We truncated the basis expansions at $j = 10$. Note that Daubechies shows in [9, p. 226] that this wavelet system has its scaling and wavelet function in C^k for all $k < 2.406$ and, more generally, that the Daubechies-N wavelet system has its scaling and wavelet function in $C^{\mu N}$ with $\mu \approx 0.2075$ for sufficiently large $N \in \mathbb{N}$.

Remark 2.7. In the sequel, we will work with unweighted Besov spaces $B^s_{p,q}(\mathbb{R}^d)$ and $b^s_{p,q}(\mathbb{R}^d)$. However, the entire analysis can be readily transferred to weighted Besov spaces with weight function $w_\sigma : \mathbb{R}^d \to \mathbb{R}$, $w_\sigma(x) := (1 + \|x\|_2^2)^{\sigma/2}$. (For its admissibility, see Lemma A.1.) To this end, observe that for $j \in \mathbb{N}_0$ and $m \in \mathbb{Z}^d$, we have

$$\begin{aligned} w_\sigma(2^{-(j-\mathbb{1}_{j\neq 0})}m) &\asymp (1 + \|m\|_2/2^{j-\mathbb{1}_{j\neq 0}})^\sigma \\ &\asymp (1 + \|m\|_\infty/2^j)^\sigma. \end{aligned}$$

Now, if f follows a Besov prior with parameters $(\alpha, \beta, \gamma, \theta, k)$ and template random variable X, Theorem 2.3 yields

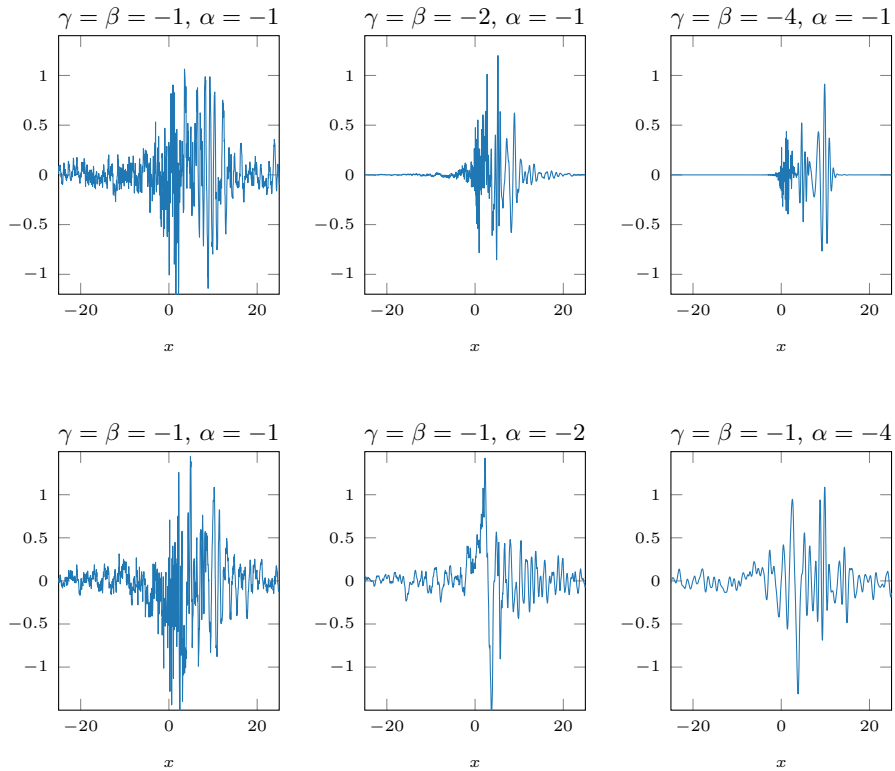


Fig. 1. Samples of our Besov prior for different sets of parameters. The parameters β and γ impact the spatial decay of the function, whereas the smoothness changes with α. In all panels, we set $\theta = 0$.

$$
\begin{aligned}
(2.7) \qquad & \|f\|_{B^s_{p,q}(\mathbb{R}^d, w_\sigma)} \\
& \asymp \left\| \left(2^{j\alpha}(j+1)^\theta (1 + \|m\|_\infty / 2^j)^{\beta + (\gamma-\beta)\mathbb{1}_{j=0}} \xi_{j,t,m}\right)_{(j,t,m)\in J} \right\|_{b^s_{p,q}(\mathbb{R}^d, w_\sigma)} \\
& \asymp \left\| \left(2^{j\alpha}(j+1)^\theta (1 + \|m\|_\infty / 2^j)^{\beta + \sigma + (\gamma+\sigma-(\beta+\sigma))\mathbb{1}_{j=0}} \xi_{j,t,m}\right)_{(j,t,m)\in J} \right\|_{b^s_{p,q}(\mathbb{R}^d)} \\
& \asymp \|g\|_{B^s_{p,q}(\mathbb{R}^d)},
\end{aligned}
$$

where g follows a Besov prior with parameters $(\alpha, \beta + \sigma, \gamma + \sigma, \theta, k)$ and template random variable X. This holds true under the assumption that (2.7) is finite; see also Remark 4.2.

2.4. Terminology for assumptions on the parameters. The terminology introduced in the following definitions will greatly simplify the statements and proofs of our main results.

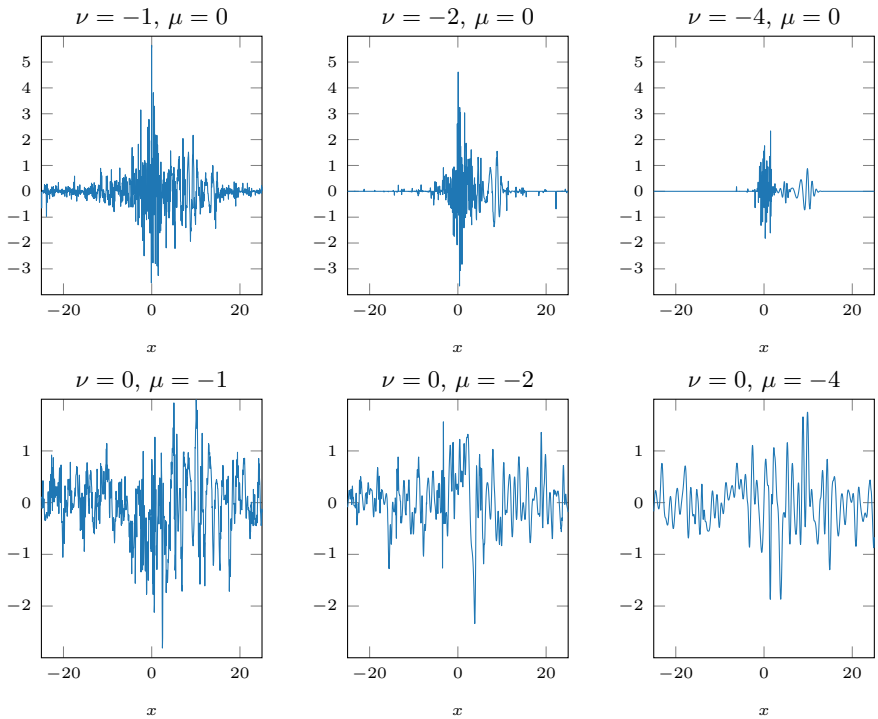


Fig. 2. Samples of our Besov–Bernoulli prior for different sets of parameters. The parameter ν impacts the spatial decay, whereas the smoothness changes with μ. In all panels, we set $\alpha = \gamma = \beta = -0.5$.

DEFINITION 2.8. Let $d \in \mathbb{N}$, $s \in \mathbb{R}$, $p, q \in (0, \infty]$, $\alpha, \beta, \gamma, \theta \in \mathbb{R}$, $\mu, \nu \in (-\infty, 0]$, and $r \in (0, \infty)$. We say that *Property A is satisfied* if

$$\gamma < -d/p, \qquad \beta < -d/p, \qquad s + d/2 + \alpha < 0$$

or

$$\gamma < -d/p, \qquad \beta < -d/p, \qquad s + d/2 + \alpha = 0, \qquad \theta < -1/q.$$

When $p = \infty$, the conditions $\gamma < -d/p$, $\beta < -d/p$ have to be replaced by $\gamma \leq 0$, $\beta \leq 0$. When $q = \infty$, the condition $\theta < -1/q$ has to be replaced by $\theta \leq 0$.

We say that *Property A′ is satisfied* if one of the following sets of conditions is satisfied:

(a) $p \neq \infty$, $q \neq \infty$, $\gamma + \frac{d+\nu}{p} < 0$, $\beta + \frac{d+\nu}{p} < 0$, $s + d/2 + \alpha + \mu/p < 0$;
(b) $p \neq \infty$, $q = \infty$, $\gamma + \frac{d+\nu}{p} < 0$, $\beta + \frac{d+\nu}{p} < 0$, $s + d/2 + \alpha + \mu/p \leq 0$;
(c) $p = \infty$, $q \neq \infty$, $\gamma \leq 0$, $\beta \leq 0$, $s + d/2 + \alpha < 0$;
(d) $p = \infty$, $q = \infty$, $\gamma \leq 0$, $\beta \leq 0$, $s + d/2 + \alpha \leq 0$.

We say that *Property A″ is satisfied* if one of the following sets of conditions is satisfied:

(a) $p \neq \infty$, $\gamma + d/p + \frac{\nu}{\max\{r,p\}} < 0$, $\beta + d/p + \frac{\nu}{\max\{r,p\}} < 0$, $s + d/2 + \alpha + \frac{\mu}{\max\{r,p\}} < 0$;
(b) $p = \infty$, $q < \infty$, $\gamma \leq 0$, $\beta \leq 0$, $s + d/2 + \alpha < 0$;
(c) $p = q = \infty$, $\gamma \leq 0$, $\beta \leq 0$, $s + d/2 + \alpha \leq 0$.

Definition 2.9. Let $d, k \in \mathbb{N}$, $s, \alpha \in \mathbb{R}$, and $p \in (0, \infty]$. We say that *Property B is satisfied* if one of the following sets of conditions is satisfied:

(a) $p \in (0,1)$, $k > \max\{s, d(1/p - 1) - s\}$,
(b) $p \in [1, \infty]$, $k > |s|$.

We say that *Property B′ is satisfied* if one of the following sets of conditions is satisfied:

(a) $p \in (0,1)$, $k > d(1/p - 1) + d/2 + \alpha$,
(b) $p \in [1, \infty]$, $k > d/2 + \alpha$.

Note that Properties B and B′ are always satisfied if k is large enough, depending on d, p, s, and α.

Remark 2.10. In view of our assumptions on the wavelet system (see Theorem 2.2 and (2.1)), Property B can be seen as a compatibility condition between the Besov space $B^s_{p,q}(\mathbb{R}^d)$ and the regularity of the wavelet system. Property B′ is similar, but also incorporates the parameter α that governs the decay/growth of the random wavelet coefficients.

In the results establishing sufficient conditions (Theorems 3.3 to 3.5, 3.8 and 5.1), Property B is used, but only to guarantee that the wavelet system gives rise to a characterization of the Besov space and of the Besov (quasi-)norm; see Theorem 2.3.

Thus, in principle, our sufficient conditions immediately extend to other wavelet systems beyond the setting of compactly supported C^k (Daubechies) wavelets, as long as the wavelet system considered yields a wavelet characterization of Besov spaces similarly to Theorem 2.3. An important example is the Meyer wavelet system [23, Chapter 3.2], in which the wavelets ψ_F and ψ_M belong to $\mathcal{S}(\mathbb{R})$. According to [21, Theorem 3.12], this system in fact provides a wavelet characterization of *all* Besov spaces $B^s_{p,q}(\mathbb{R}^d)$ for *arbitrary* $p, q \in (0, \infty]$ and $s \in \mathbb{R}$. Thus, one could say that this wavelet system *always* satisfies Property B, for arbitrary choices of the parameters.

In the results establishing necessary conditions (Theorems 4.3 and 4.5), both Properties B and B′ are used. Here, Property B is again used to guarantee that the wavelet system gives rise to a characterization of the Besov space $B^s_{p,q}(\mathbb{R}^d)$. Property B′, on the other hand, is used to guarantee that

there exist $\widetilde{s} < s$ and $\sigma < 0$ such that the wavelet system used also characterizes the weighted Besov space $B^{\widetilde{s}}_{p,q}(\mathbb{R}^d, w_\sigma)$ for $w_\sigma(x) = (1 + \|x\|_2^2)^{\sigma/2}$, which is important for the proof, in view of the fact that the wavelet expansion is not necessarily unique if one does not impose *any* a priori decay assumptions on the coefficients; see Remark 4.2.

Thus, our necessity results again extend to essentially arbitrary (not necessarily compactly supported) wavelet systems, as long as they give rise to wavelet characterizations as described in the paragraph above.

3. Sufficiency of Properties A and A′. In this section, we show that Properties A or A′ are sufficient conditions for having $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}] < \infty$ for some $r \in (0, \infty)$ if f follows a Besov prior or a Besov–Bernoulli prior, respectively. (The number r only depends on the concentration properties of the template random variable X.) In case of the Besov prior, we also show the sufficiency of Property A for having $\mathbb{E}[\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)})] < \infty$ for some $c \in (0, \infty)$, where $r \in (0, \infty)$ is again given by the concentration properties of the template random variable X. The assumptions on the template random variable X in Definition 2.5 that we will impose when $p \neq \infty$ concern the finiteness of its moments. On the other hand, our standing assumption on X (mentioned explicitly in the various theorems) is essential boundedness when $p = \infty$, i.e., there exists a real number $R > 0$ such that $|X| \leq R$ almost surely. This assumption makes the analysis relatively easy but also it cannot be neglected if we want to keep Property A as it is; cf. Proposition 5.4 and Remark 5.5.

The main observations for the proof are the following lemmas which show for $p \neq \infty$ that under Property A, the norm $\|f\|_{B^s_{p,q}(\mathbb{R}^d)}$ of a function f following a Besov prior is bounded above by a random quantity Ξ that does not depend on the parameters $(\alpha, \beta, \gamma, \theta, k)$ but only on the template variable X and on the parameters p and d governing the Besov space.

LEMMA 3.1. *Let $d \in \mathbb{N}$, $p \in (0, \infty)$, and $\beta \in \mathbb{R}$ such that $\beta < -d/p$. Let further $(\xi_{j,t,m})_{(j,t,m) \in \mathbb{N}_0 \times \{F,M\}^d \times \mathbb{Z}^d}$ be a family of random variables. Then for all $j \in \mathbb{N}_0$ and $t \in \{F, M\}^d$, we have*

$$Z_{j,t} := \left\| \left((1 + \|m\|_\infty / 2^j)^\beta |\xi_{j,t,m}| \right)_{m \in \mathbb{Z}^d} \right\|_{\ell_p} \lesssim 2^{jd/p} \cdot \Xi^{1/p}, \tag{3.1}$$

where the implied constant only depends on d, β, and p. Here

$$\Xi_t := \sup_{j \in \mathbb{N}_0} \left(\frac{1}{M_j} \sum_{\tau=1}^{M_j} |\xi_{j,t,\phi(\tau)}|^p + \sup_{\ell \geq j} \frac{1}{N_\ell} \sum_{\tau = M_\ell + 1}^{M_{\ell+1}} |\xi_{j,t,\phi(\tau)}|^p \right)$$

for $t \in \{F, M\}^d$, and

$$\Xi := \max_{t \in \{F,M\}^d} \Xi_t, \tag{3.2}$$

where $\phi : \mathbb{N} \to \mathbb{Z}^d$ is a bijection such that $\tau \mapsto \|\phi(\tau)\|_\infty$ is nondecreasing, and for $j \in \mathbb{N}_0$, we set

$$N_j := \#\{m \in \mathbb{Z}^d : 2^j < \|m\|_\infty \le 2^{j+1}\},$$

and

$$M_j := \#\{m \in \mathbb{Z}^d : \|m\|_\infty \le 2^j\} = (2^{j+1}+1)^d.$$

Proof. We have $M_j \asymp 2^{jd}$, and also

$$\begin{aligned} N_j = M_{j+1} - M_j &= (2^{j+2}+1)^d - (2^{j+1}+1)^d \\ &= d \cdot y^{d-1}((2^{j+2}+1) - (2^{j+1}+1)) \asymp 2^{(d-1)j}2^j = 2^{jd} \end{aligned}$$

for some $y \in (2^{j+1}+1, 2^{j+2}+1)$, due to the mean value theorem applied to $x \mapsto x^d$. Next note that

$$\begin{aligned} Z_{j,t}^p &= \sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty/2^j)^{\beta p}|\xi_{j,t,m}|^p \\ &= \sum_{\substack{m\in\mathbb{Z}^d \\ \|m\|_\infty\le 2^j}} (1+\|m\|_\infty/2^j)^{\beta p}|\xi_{j,t,m}|^p + \sum_{\substack{m\in\mathbb{Z}^d \\ \|m\|_\infty> 2^j}} (1+\|m\|_\infty/2^j)^{\beta p}|\xi_{j,t,m}|^p \\ &\asymp \sum_{\substack{m\in\mathbb{Z}^d \\ \|m\|_\infty\le 2^j}} |\xi_{j,t,m}|^p + \sum_{\substack{m\in\mathbb{Z}^d \\ \|m\|_\infty> 2^j}} (\|m\|_\infty/2^j)^{\beta p}|\xi_{j,t,m}|^p \\ &= \sum_{\tau=1}^{M_j} |\xi_{j,t,\phi(\tau)}|^p + \sum_{\ell\ge j}\sum_{\tau=M_\ell+1}^{M_{\ell+1}} (\|\phi(\tau)\|_\infty/2^j)^{\beta p}|\xi_{j,t,\phi(\tau)}|^p \\ &\asymp \sum_{\tau=1}^{M_j} |\xi_{j,t,\phi(\tau)}|^p + 2^{-j\beta p}\sum_{\ell\ge j} 2^{\ell d+\ell\beta p}\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}} |\xi_{j,t,\phi(\tau)}|^p \\ &= M_j S_{j,t} + 2^{-j\beta p}\sum_{\ell\ge j} 2^{\ell d+\ell\beta p} S_{j,t,\ell} \le \Big(M_j + 2^{-j\beta p}\sum_{\ell\ge j} 2^{\ell(d+\beta p)}\Big)\cdot \Xi_t \\ &\lesssim 2^{jd}\Xi_t \le 2^{jd}\Xi \end{aligned}$$

since $d + \beta p < 0$, where

$$S_{j,t} := \frac{1}{M_j}\sum_{\tau=1}^{M_j} |\xi_{j,t,\phi(\tau)}|^p \le \Xi_t \quad \text{and} \quad S_{j,t,\ell} := \frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}} |\xi_{j,t,\phi(\tau)}|^p \le \Xi_t$$

for $\ell \ge j$. ∎

LEMMA 3.2. *Let $d \in \mathbb{N}$, $s, \alpha, \beta, \gamma, \theta \in \mathbb{R}$, $p \in (0,\infty)$, and $q \in (0,\infty]$. Let further $(\xi_{j,t,m})_{(j,t,m)\in\mathbb{N}_0\times\{F,M\}^d\times\mathbb{Z}^d}$ be a family of random variables. Set $\mathbf{a} = (a_{j,t,m})_{(j,t,m)\in J}$, where*

$$a_{j,t,m} := 2^{j\alpha}(j+1)^\theta(1+\|m\|_\infty/2^j)^{\beta+(\gamma-\beta)\mathbb{1}_{j=0}}\xi_{j,t,m}$$

for $(j,t,m) \in J$. If Property A is satisfied, then for Ξ as in (3.2), we have $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} \lesssim \Xi^{1/p}$, where the implied constant only depends on d, p, q, s, α, θ, β, and γ.

Proof. Note that Property A implies $\gamma < -d/p$ and $\beta < -d/p$, since $p < \infty$. Recall that $\xi_m := \xi_{0,(F,\dots,F),m}$ for $m \in \mathbb{Z}^d$, and let

$$\eta_1 := \|((1+\|m\|_\infty)^\gamma |\xi_m|)_{m\in\mathbb{Z}^d}\|_{\ell_p},$$
$$\eta_2 := \|(2^{j(s+d/2-d/p+\alpha)}(j+1)^\theta Z_{j,t})_{j\in\mathbb{N},\, t\in T_j}\|_{\ell_q}$$

with $Z_{j,t}$ as in (3.1). Utilize (3.1) from Lemma 3.1 with γ replacing β, and $j=0$, to conclude that $\eta_1 \lesssim \Xi^{1/p}$. Furthermore, we have

$$\begin{aligned}\eta_2 &= \|(2^{j(s+d/2-d/p+\alpha)}(j+1)^\theta Z_{j,t})_{j\in\mathbb{N},\, t\in T_j}\|_{\ell_q}\\ &\lesssim \|(2^{j(s+d/2+\alpha)}(j+1)^\theta \Xi^{1/p})_{j\in\mathbb{N},\, t\in T_j}\|_{\ell_q}\\ &\lesssim \|(2^{j(s+d/2+\alpha)}(j+1)^\theta)_{j\in\mathbb{N}}\|_{\ell_q}\Xi^{1/p} \lesssim \Xi^{1/p},\end{aligned}$$

thanks to Property A. If $q \neq \infty$, then $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} = (\eta_1^q + \eta_2^q)^{1/q} \lesssim \Xi^{1/p}$. On the other hand, if $q = \infty$, then $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} = \max\{\eta_1, \eta_2\} \lesssim \Xi^{1/p}$. ∎

Now we are ready to prove the sufficiency of Property A for our three target conditions when f follows a Besov prior. We start with the almost sure membership of f in $B^s_{p,q}(\mathbb{R}^d)$.

THEOREM 3.3. *Let $d,k \in \mathbb{N}$, $s,\alpha,\beta,\gamma,\theta \in \mathbb{R}$, and $p,q \in (0,\infty]$. Assume that $\mathbf{a}$ follows a Besov sequence prior and f follows a Besov prior, both with parameters $(\alpha,\beta,\gamma,\theta,k)$ and template random variable X, satisfying*

$$\begin{cases} \exists \varepsilon > 0 : \mathbb{E}[|X|^{(1+\varepsilon)p}] < \infty & \text{if } p \in (0,\infty), \\ \exists R > 0 : |X| \overset{\text{a.s.}}{\leq} R & \text{if } p = \infty. \end{cases} \tag{3.3}$$

If Property A holds, then $\mathbf{a}$ almost surely belongs to $b^s_{p,q}(\mathbb{R}^d)$. Moreover, if Properties A and B are satisfied, then f belongs almost surely to $B^s_{p,q}(\mathbb{R}^d)$.

Proof. STEP 1: $p \neq \infty$. From Lemma 3.2, we obtain $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} \lesssim \Xi^{1/p}$ with Ξ as in (3.2). Next we show that $\Xi \overset{\text{a.s.}}{<} \infty$. To this end, for fixed $t \in \{F,M\}^d$, apply Lemma A.7 to $X_{j,\tau} := |\xi_{j,t,\phi(\tau)}|^p$ and $\sigma := 1$, where again $\phi : \mathbb{N} \to \mathbb{Z}^d$ is a bijection such that $\tau \mapsto \|\phi(\tau)\|_\infty$ is nondecreasing. We obtain $\mathbb{E}[\Xi_t] < \infty$, which implies that $\Xi_t \overset{\text{a.s.}}{<} \infty$, whence also $\Xi = \max_{t\in\{F,M\}^d} \Xi_t \overset{\text{a.s.}}{<} \infty$.

STEP 2: $p = \infty$. Note that $|\xi_{j,t,m}| \overset{\text{a.s.}}{\leq} R$ and thus

$$Z_{j,t} = \|((1+\|m\|_\infty/2^j)^\beta |\xi_{j,t,m}|)_{m\in\mathbb{Z}^d}\|_{\ell_p} \overset{\text{a.s.}}{\leq} R$$

by $\beta \le 0$ (which follows from Property A), and moreover

$$\eta_1 = \|((1+\|m\|_\infty)^\gamma |\xi_m|)_{m\in\mathbb{Z}^d}\|_{\ell_p} \overset{\text{a.s.}}{\le} R$$

by $\gamma \le 0$ (which follows from Property A). Finally,

$$\begin{aligned}\eta_2 &= \|(2^{j(s+d/2+\alpha)}(j+1)^\theta Z_{j,t})_{j\in\mathbb{N},\, t\in T_j}\|_{\ell_q} \\ &\overset{\text{a.s.}}{\le} R\|(2^{j(s+d/2+\alpha)}(j+1)^\theta)_{j\in\mathbb{N},\, t\in T_j}\|_{\ell_q} < \infty,\end{aligned}$$

thanks to the conditions on s, d, α, θ, and q from Property A. This easily implies $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} \overset{\text{a.s.}}{\lesssim} R < \infty$.

Step 3. If Properties A and B are satisfied, then Theorem 2.3 shows that $\|f\|_{B^s_{p,q}(\mathbb{R}^d)} = \|S(\mathbf{a})\|_{B^s_{p,q}(\mathbb{R}^d)} \asymp \|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)}$ and hence $\|f\|_{B^s_{p,q}(\mathbb{R}^d)} < \infty$ almost surely. ∎

Next we show the sufficiency of Property A for the finiteness of the moments of $\|f\|_{B^s_{p,q}(\mathbb{R}^d)}$ when f follows a Besov prior.

Theorem 3.4. *Let $d, k \in \mathbb{N}$, $s, \alpha, \beta, \gamma, \theta \in \mathbb{R}$, $p, q \in (0,\infty]$, and $r \in (0,\infty)$. Assume that $\mathbf{a}$ follows a Besov sequence prior and f follows a Besov prior, both with parameters $(\alpha, \beta, \gamma, \theta, k)$ and template random variable X, satisfying*

$$\begin{cases} \exists \varepsilon > 0 : \mathbb{E}[|X|^{(1+\varepsilon)\max\{r,p\}}] < \infty & \text{if } p \in (0,\infty), \\ \exists R > 0 : |X| \overset{\text{a.s.}}{\le} R & \text{if } p = \infty. \end{cases} \tag{3.4}$$

If Property A holds, then $\mathbb{E}[\|\mathbf{a}\|^r_{b^s_{p,q}(\mathbb{R}^d)}] < \infty$. Moreover, if Properties A and B are satisfied, then $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}] < \infty$.

Proof. Step 1: $p \ne \infty$. From Lemma 3.2 and the monotonicity of the expectation, we conclude $\mathbb{E}[\|\mathbf{a}\|^r_{b^s_{p,q}(\mathbb{R}^d)}] \lesssim \mathbb{E}[\Xi^{r/p}]$ with Ξ as in (3.2). Next we show that $\mathbb{E}[\Xi^{r/p}] < \infty$.

To this end, fix $t \in \{F,M\}^d$ and apply Lemma A.7 to $X_{j,\tau} := |\xi_{j,t,\phi(\tau)}|^p$ and $\sigma := \max\{1, r/p\}$, where again $\phi : \mathbb{N} \to \mathbb{Z}^d$ is a bijection such that $\tau \mapsto \|\phi(\tau)\|_\infty$ is nondecreasing. This lemma is applicable because

$$\mathbb{E}[X_{j,\tau}^{(1+\varepsilon)\sigma}] = \mathbb{E}[|\xi_{j,t,\phi(\tau)}|^{\max\{(1+\varepsilon)p,(1+\varepsilon)r\}}] = \mathbb{E}[|X|^{(1+\varepsilon)\max\{p,r\}}] < \infty.$$

Hence, Lemma A.7 shows that $\mathbb{E}[\Xi_t^\sigma] < \infty$ and thus also $\mathbb{E}[\Xi_t^{r/p}] < \infty$ by Jensen's inequality. It follows that also

$$\mathbb{E}[\Xi^{r/p}] = \mathbb{E}\Big[\max_{t\in\{F,M\}^d} \Xi_t^{r/p}\Big] \le \mathbb{E}\Big[\sum_{t\in\{F,M\}^d} \Xi_t^{r/p}\Big] = \sum_{t\in\{F,M\}^d} \mathbb{E}[\Xi_t^{r/p}] < \infty.$$

Step 2: $p = \infty$. Similar to Step 2 of the proof of Theorem 3.3, we have $\|\mathbf{a}\|^r_{b^s_{p,q}(\mathbb{R}^d)} \overset{\text{a.s.}}{\lesssim} R^r < \infty$.

Step 3. If Properties A and B are satisfied, then Theorem 2.3 shows that $\|f\|_{B^s_{p,q}(\mathbb{R}^d)} = \|S(\mathbf{a})\|_{B^s_{p,q}(\mathbb{R}^d)} \asymp \|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)}$ and hence

$$\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}] = \mathbb{E}[\|S(\mathbf{a})\|^r_{B^s_{p,q}(\mathbb{R}^d)}] \asymp \mathbb{E}[\|\mathbf{a}\|^r_{b^s_{p,q}(\mathbb{R}^d)}] < \infty. \ \blacksquare$$

Next we show the sufficiency of Property A for the finiteness of the moment generating function $\mathbb{E}[\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)})]$ for small c when f follows a Besov prior. If $c \in (-\infty, 0]$, then $\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}) \in (0, 1]$ is bounded, which implies that $\mathbb{E}[\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)})] < \infty$. The following theorem is about the nontrivial case where c is positive.

Theorem 3.5. *Let $d, k \in \mathbb{N}$, $s, \alpha, \beta, \gamma, \theta \in \mathbb{R}$, $p, q \in (0, \infty]$, and $r \in (0, \infty)$. Assume that f follows a Besov prior with parameters $(\alpha, \beta, \gamma, \theta, k)$ and template random variable X with*

$$\begin{cases} \exists C > 0 : \mathbb{E}[\exp(C|X|^{\max\{r,p\}})] < \infty & \text{if } p \in (0, \infty), \\ \exists R > 0 : |X| \overset{\text{a.s.}}{\leq} R & \text{if } p = \infty. \end{cases}$$

If Properties A and B are satisfied, then there exists a constant $c > 0$ with $\mathbb{E}[\exp(c\|f\|^r_{B^s_{p,q}})] < \infty$.

Proof. Step 1: $p \neq \infty$. For Ξ as in (3.2), we have $\Xi^{r/p} \leq 1 + \Xi^{\frac{\max\{r,p\}}{p}}$. Lemma 3.2 gives the existence of a constant $\kappa_1 = \kappa_1(d, s, \alpha, \beta, \gamma, \theta, p, q) > 0$ such that $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} \leq \kappa_1 \Xi^{1/p}$ where $\mathbf{a}$ is as in (2.5). Moreover, Theorem 2.3 yields the existence of a constant $\kappa_2 = \kappa_2(d, s, p, q, k) > 0$ such that $\|f\|_{B^s_{p,q}(\mathbb{R}^d)} = \|S(\mathbf{a})\|_{B^s_{p,q}(\mathbb{R}^d)} \leq \kappa_2 \|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} \leq \kappa_1 \kappa_2 \Xi^{1/p}$ whenever $\Xi < \infty$. Thus, for $c > 0$ and $\kappa := \kappa_1 \kappa_2$, we have

$$\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}) \leq \exp\left(c\kappa^r \Xi^{r/p}\right) \lesssim \exp\left(c\kappa^r \Xi^{\frac{\max\{r,p\}}{p}}\right).$$

The monotonicity of the expectation yields

$$\mathbb{E}[\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)})] \lesssim \mathbb{E}\big[\exp\left(c\kappa^r \Xi^{\frac{\max\{r,p\}}{p}}\right)\big].$$

Next we show that $\mathbb{E}[\exp(c\kappa^r \Xi^{\frac{\max\{r,p\}}{p}})] < \infty$ for sufficiently small $c > 0$. To this end, fix $t \in \{F, M\}^d$ and apply Lemma A.8 to $X_{j,\tau} := |\xi_{j,t,\phi(\tau)}|^p$ and $\sigma := \max\{r, p\}/p$, where again $\phi : \mathbb{N} \to \mathbb{Z}^d$ is a bijection such that $\tau \mapsto \|\phi(\tau)\|_\infty$ is nondecreasing. This lemma is applicable by choosing $c = \frac{C}{2^{\sigma+1}(1+\varepsilon)\kappa^r}$ for some $\varepsilon > 0$ and noting

$$\begin{aligned} \mathbb{E}[\exp(CX^\sigma_{j,\tau})] &= \mathbb{E}[\exp(C|\xi_{j,t,\phi(\tau)}|^{\max\{r,p\}})] \\ &= \mathbb{E}[\exp(C|X|^{\max\{r,p\}})] < \infty. \end{aligned}$$

Lemma A.8 now yields $\mathbb{E}[\exp(c\kappa^r \Xi^\sigma_t)] = \mathbb{E}[\exp(\frac{C}{2^{\sigma+1}(1+\varepsilon)} \Xi^\sigma_t)] < \infty$. Note that since the $\xi_{j,t,m}$ are jointly independent, the random variables Ξ_t,

$t \in \{F,M\}^d$, are independent. It follows that

$$\mathbb{E}[\exp(c\kappa^r \Xi^\sigma)] = \mathbb{E}\Big[\exp\Big(c\kappa^r \max_{t\in\{F,M\}^d} \Xi_t^\sigma\Big)\Big] \le \mathbb{E}\Big[\exp\Big(c\kappa^r \sum_{t\in\{F,M\}^d} \Xi_t^\sigma\Big)\Big]$$
$$= \mathbb{E}\Big[\prod_{t\in\{F,M\}^d} \exp(c\kappa^r \Xi_t^\sigma)\Big] = \prod_{t\in\{F,M\}^d} \mathbb{E}[\exp(c\kappa^r \Xi_t^\sigma)] < \infty.$$

Step 2: $p = \infty$. Similar to Step 2 of the proof of Theorem 3.3, we have $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} \overset{\text{a.s.}}{\lesssim} R$, where $\mathbf{a}$ is as in (2.5). By Theorem 2.3, it follows again that $\|f\|_{B^s_{p,q}(\mathbb{R}^d)} \overset{\text{a.s.}}{\lesssim} R$, i.e., there exists a constant $\kappa > 0$ with $\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)} \le \kappa R^r$ almost surely. Then the monotonicity of the expectation directly gives $\mathbb{E}[\exp(c\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)})] \le \mathbb{E}[\exp(c\kappa R^r)] < \infty$. ∎

We conclude this section by establishing the sufficiency of Property A″ for the finiteness of the moments of $\|f\|_{B^s_{p,q}(\mathbb{R}^d)}$ (and thus for the almost sure membership of f in $B^s_{p,q}(\mathbb{R}^d)$) when f follows a Besov–Bernoulli prior. Let us start with the counterpart of Lemma 3.1.

Lemma 3.6. *Let $d \in \mathbb{N}$, $p, r \in (0,\infty)$, $\beta \in \mathbb{R}$, $\mu, \nu \in (-\infty, 0]$, and $\delta \in (0,1)$ such that $\beta p + d + \frac{\nu p}{\max\{r,p\}}(1-\delta) < 0$. Let $(\xi_{j,t,m})_{(j,t,m)\in\mathbb{N}_0\times\{F,M\}^d\times\mathbb{Z}^d}$ and $(\lambda_{j,t,m})_{(j,t,m)\in\mathbb{N}_0\times\{F,M\}^d\times\mathbb{Z}^d}$ be two families of random variables with $\xi_{j,t,m} \in \mathbb{R}$ and $\lambda_{j,t,m} \in \{0,1\}$ for $(j,t,m) \in \mathbb{N}_0 \times \{F,M\}^d \times \mathbb{Z}^d$. Finally, set*

$$\varrho_{j,m} := 2^{j\mu}(1 + \|m\|_\infty / 2^j)^\nu$$

for $j \in \mathbb{N}_0$ and $m \in \mathbb{Z}^d$. Then for all $j \in \mathbb{N}_0$ and $t \in \{F,M\}^d$, we have

(3.5)
$$\widetilde{Z}_{j,t} := \|((1 + \|m\|_\infty/2^j)^\beta \lambda_{j,t,m}|\xi_{j,t,m}|)_{m\in\mathbb{Z}^d}\|_{\ell_p} \lesssim 2^{jd/p + j\frac{\mu}{\max\{r,p\}}(1-\delta)} \widetilde{\Xi}^{1/p},$$

where the implied constant only depends on d, p, β, μ, ν, r, and δ. Here

$$\widetilde{\Xi}_t := \sup_{j\in\mathbb{N}_0}\Bigg(\frac{1}{M_j}\sum_{\tau=1}^{M_j} \varrho_{j,\phi(\tau)}^{\frac{p}{\max\{r,p\}}(\delta-1)} \lambda_{j,t,\phi(\tau)}|\xi_{j,t,\phi(\tau)}|^p + \sup_{\ell\ge j}\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}} \varrho_{j,\phi(\tau)}^{\frac{p}{\max\{r,p\}}(\delta-1)} \lambda_{j,t,\phi(\tau)}|\xi_{j,t,\phi(\tau)}|^p\Bigg)$$

for $t \in \{F,M\}^d$ and

(3.6)
$$\widetilde{\Xi} := \max_{t\in\{F,M\}^d} \widetilde{\Xi}_t,$$

where $\phi : \mathbb{N} \to \mathbb{Z}^d$ is a bijection such that $\tau \mapsto \|\phi(\tau)\|_\infty$ is nondecreasing, and for $j \in \mathbb{N}_0$,

$$N_j := \#\{m \in \mathbb{Z}^d : 2^j < \|m\|_\infty \le 2^{j+1}\},$$
$$M_j := \#\{m \in \mathbb{Z}^d : \|m\|_\infty \le 2^j\} = (2^{j+1}+1)^d.$$

Proof. Let $j \in \mathbb{N}_0$. If $1 \leq \tau \leq M_j$, then $0 \leq \|\phi(\tau)\|_\infty \leq 2^j$, i.e., $1 \leq 1 + \|\phi(\tau)\|_\infty / 2^j \leq 2$. It follows that $\varrho_{j,\phi(\tau)} \asymp 2^{j\mu}$. If, on the other hand, $\tau > M_j$, then there exists a unique number $\ell \in \mathbb{N}_0$ with $\ell \geq j$ such that $M_\ell + 1 \leq \tau \leq M_{\ell+1}$. Then $2^\ell < \|\phi(\tau)\|_\infty \leq 2^{\ell+1}$ and

$$2^{\ell-j} \leq 1 + 2^{\ell-j} \leq 1 + \|\phi(\tau)\|_\infty / 2^j \leq 1 + 2^{\ell+1-j} \leq 3 \cdot 2^{\ell-j},$$

i.e., $\varrho_{j,\phi(\tau)} \asymp 2^{j\mu} \cdot 2^{(\ell-j)\nu}$. For $j \in \mathbb{N}_0$ and $t \in T_j$, we thus obtain

$$\begin{aligned}
\widetilde{Z}_{j,t}^p &= \sum_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty / 2^j)^{\beta p} \lambda_{j,t,m} |\xi_{j,t,m}|^p \\
&= \sum_{\tau=1}^{M_j} (1 + \|\phi(\tau)\|_\infty / 2^j)^{\beta p} \lambda_{j,t,\phi(\tau)} |\xi_{j,t,\phi(\tau)}|^p \\
&\quad + \sum_{\ell \geq j} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} (1 + \|\phi(\tau)\|_\infty / 2^j)^{\beta p} \lambda_{j,t,\phi(\tau)} |\xi_{j,t,\phi(\tau)}|^p \\
&\lesssim M_j \frac{1}{M_j} \sum_{\tau=1}^{M_j} \lambda_{j,t,\phi(\tau)} |\xi_{j,t,\phi(\tau)}|^p \\
&\quad + \sum_{\ell \geq j} N_\ell \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} 2^{(\ell-j)\beta p} \lambda_{j,t,\phi(\tau)} |\xi_{j,t,\phi(\tau)}|^p.
\end{aligned}$$

Next, we note

$$\begin{aligned}
&M_j \frac{1}{M_j} \sum_{\tau=1}^{M_j} \lambda_{j,t,\phi(\tau)} |\xi_{j,t,\phi(\tau)}|^p \\
&\qquad \asymp 2^{jd} 2^{j\mu \frac{p}{\max\{r,p\}}(1-\delta)} \frac{1}{M_j} \sum_{\tau=1}^{M_j} \varrho_{j,\phi(\tau)}^{\frac{p}{\max\{r,p\}}(\delta-1)} \lambda_{j,t,\phi(\tau)} |\xi_{j,t,\phi(\tau)}|^p \\
&\qquad \lesssim 2^{jd} 2^{j\mu \frac{p}{\max\{r,p\}}(1-\delta)} \widetilde{\Xi}_t.
\end{aligned}$$

Finally, recall from the proof of Lemma 3.1 that $N_\ell \asymp 2^{\ell d}$, and from the assumptions of the present lemma that $\beta p + d + \frac{\nu p}{\max\{r,p\}}(1-\delta) < 0$, and hence

$$\begin{aligned}
&\sum_{\ell \geq j} N_\ell \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} 2^{(\ell-j)\beta p} \lambda_{j,t,\phi(\tau)} |\xi_{j,t,\phi(\tau)}|^p \\
&\quad \asymp \sum_{\ell \geq j} 2^{(\ell-j)\beta p + \ell d} \big(2^{j\mu} 2^{(\ell-j)\nu}\big)^{\frac{p(1-\delta)}{\max\{r,p\}}} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} \varrho_{j,\phi(\tau)}^{\frac{p(\delta-1)}{\max\{r,p\}}} \lambda_{j,t,\phi(\tau)} |\xi_{j,t,\phi(\tau)}|^p \\
&\quad \lesssim \sum_{\ell \geq j} 2^{(\ell-j)\beta p + \ell d} \big(2^{j\mu} 2^{(\ell-j)\nu}\big)^{\frac{p(1-\delta)}{\max\{r,p\}}} \widetilde{\Xi}_t
\end{aligned}$$

$$
\begin{aligned}
&= 2^{-j\beta p + j\mu \frac{p}{\max\{r,p\}}(1-\delta) - j\nu \frac{p(1-\delta)}{\max\{r,p\}}} \sum_{\ell \geq j} 2^{\ell\beta p + \ell d + \ell\nu \frac{p(1-\delta)}{\max\{r,p\}}} \widetilde{\Xi}_t \\
&\lesssim 2^{-j\beta p + j\mu \frac{p}{\max\{r,p\}}(1-\delta) - j\nu \frac{p(1-\delta)}{\max\{r,p\}}} 2^{j(\beta p + d + \nu \frac{p(1-\delta)}{\max\{r,p\}})} \widetilde{\Xi}_t \\
&= 2^{jd} 2^{j\mu \frac{p(1-\delta)}{\max\{r,p\}}} \widetilde{\Xi}_t.
\end{aligned}
$$

Overall, we conclude $\widetilde{Z}_{j,t}^p \lesssim 2^{jd + j\mu \frac{p}{\max\{r,p\}}(1-\delta)} \widetilde{\Xi}_t \leq 2^{jd + j\mu \frac{p}{\max\{r,p\}}(1-\delta)} \widetilde{\Xi}$. ∎

Next, we show an analog of Lemma 3.2.

LEMMA 3.7. *Let $d \in \mathbb{N}$, $s, \alpha, \beta, \gamma \in \mathbb{R}$, $p, r \in (0,\infty)$, $q \in (0,\infty]$, and $\mu, \nu \in (-\infty, 0]$. Let $(\xi_{j,t,m})_{(j,t,m)\in\mathbb{N}_0\times\{F,M\}^d\times\mathbb{Z}^d}$ and $(\lambda_{j,t,m})_{(j,t,m)\in\mathbb{N}_0\times\{F,M\}^d\times\mathbb{Z}^d}$ be two families of random variables with $\xi_{j,t,m} \in \mathbb{R}$ and $\lambda_{j,t,m} \in \{0,1\}$ for $(j,t,m) \in \mathbb{N}_0 \times \{F,M\}^d \times \mathbb{Z}^d$. Set*

$$\varrho_{j,m} := 2^{j\mu}(1 + \|m\|_\infty / 2^j)^\nu$$

and $\mathbf{a} = (a_{j,t,m})_{(j,t,m)\in J}$, where

$$a_{j,t,m} := 2^{j\alpha}(1 + \|m\|_\infty / 2^j)^{\beta + (\gamma-\beta)\mathbb{1}_{j=0}} \lambda_{j,t,m} \xi_{j,t,m}$$

for $(j,t,m) \in J$. If Property A″ is satisfied, then for $\widetilde{\Xi}$ as in (3.6), we have $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} \lesssim \widetilde{\Xi}^{1/p}$, where the implied constant only depends on s, α, μ, ν, β, γ, p, q, d, and r.

Proof. Recall that $\lambda_m := \lambda_{0,(F,\dots,F),m}$ and $\xi_m := \xi_{0,(F,\dots,F),m}$ for $m \in \mathbb{Z}^d$, and let

$$
\begin{aligned}
\widetilde{\eta}_1 &:= \|((1 + \|m\|_\infty)^\gamma \lambda_m |\xi_m|)_{m\in\mathbb{Z}^d}\|_{\ell_p}, \\
\widetilde{\eta}_2 &:= \|(2^{j(s+d/2-d/p+\alpha)} \widetilde{Z}_{j,t})_{j\in\mathbb{N}, t\in T_j}\|_{\ell_q}
\end{aligned}
$$

with $\widetilde{Z}_{j,t}$ as in (3.5).

By Property A″, there exists $\delta \in (0,1)$ such that $\beta p + d + \nu \frac{p}{\max\{r,p\}}(1-\delta) < 0$, $\gamma p + d + \nu \frac{p}{\max\{r,p\}}(1-\delta) < 0$ and $s + d/2 + \alpha + \frac{\mu}{\max\{r,p\}}(1-\delta) < 0$. Lemma 3.6 with γ replacing β and with $j = 0$ directly gives $\widetilde{\eta}_1 \lesssim \widetilde{\Xi}^{1/p}$. Furthermore, again by Lemma 3.6, we have

$$
\begin{aligned}
\widetilde{\eta}_2 &= \|(2^{j(s+d/2-d/p+\alpha)} \widetilde{Z}_{j,t})_{j\in\mathbb{N}, t\in T_j}\|_{\ell_q} \\
&\lesssim \|(2^{j(s+d/2+\alpha+\frac{\mu}{\max\{r,p\}}(1-\delta))} \widetilde{\Xi}^{1/p})_{j\in\mathbb{N}, t\in T_j}\|_{\ell_q} \\
&\lesssim \|(2^{j(s+d/2+\alpha+\frac{\mu}{\max\{r,p\}}(1-\delta))})_{j\in\mathbb{N}}\|_{\ell_q} \widetilde{\Xi}^{1/p} \lesssim \widetilde{\Xi}^{1/p}
\end{aligned}
$$

since $s + d/2 + \alpha + \frac{\mu}{\max\{r,p\}}(1-\delta) < 0$. If $q \neq \infty$, then $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} = (\widetilde{\eta}_1^q + \widetilde{\eta}_2^q)^{1/q} \lesssim \widetilde{\Xi}^{1/p}$. On the other hand, if $q = \infty$, then $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} = \max\{\widetilde{\eta}_1, \widetilde{\eta}_2\} \lesssim \widetilde{\Xi}^{1/p}$. ∎

Now we are ready to show, analogously to Theorem 3.4, that Property A″ is sufficient to guarantee $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}] < \infty$ if f follows a Besov–Bernoulli prior.

THEOREM 3.8. *Let $d, k \in \mathbb{N}$, $s, \alpha, \beta, \gamma \in \mathbb{R}$, $p, q \in (0, \infty]$, $r \in (0, \infty)$, and $\mu, \nu \in (-\infty, 0]$. Assume that $\mathbf{a}$ follows a Besov–Bernoulli sequence prior and f follows a Besov–Bernoulli prior, both with parameters $(\alpha, \beta, \gamma, \mu, \nu, k)$ and template random variable X, satisfying*

$$\begin{cases} \exists \varepsilon > 0 : \mathbb{E}\big[|X|^{(1+\varepsilon)\max\{r,p\}}\big] < \infty & \text{if } p \in (0,\infty), \\ \exists R > 0 : |X| \overset{\text{a.s.}}{\leq} R & \text{if } p = \infty. \end{cases} \tag{3.7}$$

If Property A″ is satisfied, then $\mathbb{E}[\|\mathbf{a}\|^r_{b^s_{p,q}(\mathbb{R}^d)}] < \infty$. If Properties A″ and B are satisfied, then $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}] < \infty$. In particular, f belongs to $B^s_{p,q}(\mathbb{R}^d)$ almost surely.

Proof. STEP 1: $p \neq \infty$. Let $\varepsilon_0 > 0$ be such that $\mathbb{E}[|X|^{(1+\varepsilon_0)\max\{r,p\}}] < \infty$. By Property A″, there exists $\delta \in (0,1)$ such that $\beta p + d + \nu \frac{p}{\max\{r,p\}}(1-\delta) < 0$, $\gamma p + d + \nu \frac{p}{\max\{r,p\}}(1-\delta) < 0$, and $s + d/2 + \alpha + \frac{\mu}{\max\{r,p\}}(1-\delta) < 0$. Next, choose $\varepsilon \in (0, \varepsilon_0)$ such that $\delta + \delta\varepsilon - \varepsilon \geq 0$. From Lemma 3.7 and the monotonicity of the expectation, we obtain $\mathbb{E}[\|\mathbf{a}\|^r_{b^s_{p,q}(\mathbb{R}^d)}] \lesssim \mathbb{E}[\widetilde{\Xi}^{r/p}]$ with $\widetilde{\Xi}$ as in (3.6). Next we show that $\mathbb{E}[\widetilde{\Xi}^{r/p}] \overset{\text{a.s.}}{<} \infty$. To this end, for fixed $t \in \{F, M\}^d$, apply Lemma A.9 to $X_{j,\tau} := |\xi_{j,t,\phi(\tau)}|^p$ and $\lambda_{j,\tau} := \lambda_{j,t,\phi(\tau)}$ and $\sigma := \max\{1, r/p\}$, where again $\phi : \mathbb{N} \to \mathbb{Z}^d$ is a bijection such that $\tau \mapsto \|\phi(\tau)\|_\infty$ is nondecreasing. This lemma is applicable because

$$\mathbb{E}[X_{j,\tau}^{(1+\varepsilon)\sigma}] = \mathbb{E}[|X|^{(1+\varepsilon)\max\{p,r\}}] \lesssim 1 + \mathbb{E}[|X|^{(1+\varepsilon_0)\max\{p,r\}}] < \infty.$$

Hence, Lemma A.9 shows that $\mathbb{E}[\widetilde{\Xi}_t^\sigma] < \infty$ and thus also $\mathbb{E}[\widetilde{\Xi}_t^{r/p}] < \infty$ by Jensen's inequality. It follows that also

$$\begin{aligned} \mathbb{E}[\widetilde{\Xi}^{r/p}] &= \mathbb{E}\Big[\max_{t \in \{F,M\}^d} \widetilde{\Xi}_t^{r/p}\Big] \\ &\leq \mathbb{E}\Big[\sum_{t \in \{F,M\}^d} \widetilde{\Xi}_t^{r/p}\Big] = \sum_{t \in \{F,M\}^d} \mathbb{E}[\widetilde{\Xi}_t^{r/p}] < \infty. \end{aligned}$$

STEP 2: $p = \infty$. Recall that we write $\lambda_m = \lambda_{0,(F,\ldots,F),m}$ and $\xi_m = \xi_{0,(F,\ldots,F),m}$. Note that $\lambda_{j,t,m}|\xi_{j,t,m}| \overset{\text{a.s.}}{\leq} R$ and thus

$$\widetilde{Z}_{j,t} = \|((1 + \|m\|_\infty/2^j)^\beta \lambda_{j,t,m}|\xi_{j,t,m}|)_{m \in \mathbb{Z}^d}\|_{\ell_p} \overset{\text{a.s.}}{\leq} R$$

by $\beta \leq 0$ (which holds by Property A″),

$$\widetilde{\eta}_1 = \|((1 + \|m\|_\infty)^\gamma \lambda_m |\xi_m|)_{m \in \mathbb{Z}^d}\|_{\ell_p} \overset{\text{a.s.}}{\leq} R$$

by $\gamma \le 0$ (which holds by Property A″), and

$$\widetilde{\eta}_2 = \|(2^{j(s+d/2+\alpha)}\widetilde{Z}_{j,t})_{j\in\mathbb{N},\,t\in T_j}\|_{\ell_q} \overset{\text{a.s.}}{\le} R\|(2^{j(s+d/2+\alpha)})_{j\in\mathbb{N},\,t\in T_j}\|_{\ell_q} \lesssim R,$$

since Property A″ implies $s+d/2+\alpha < 0$ in case $q < \infty$ and $s+d/2+\alpha \le 0$ in case $q = \infty$. As in the proof of Lemma 3.7, we then get

$$\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} \lesssim \widetilde{\eta}_1 + \widetilde{\eta}_2 \lesssim R$$

almost surely, which easily implies $\mathbb{E}[\|\mathbf{a}\|^r_{b^s_{p,q}(\mathbb{R}^d)}] < \infty$.

Step 3. If Properties A″ and B are satisfied, Theorem 2.3 shows that

$$\|f\|_{B^s_{p,q}(\mathbb{R}^d)} = \|S(\mathbf{a})\|_{B^s_{p,q}(\mathbb{R}^d)} \asymp \|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)}$$

and hence $\mathbb{E}[\|f\|^r_{B^s_{p,q}(\mathbb{R}^d)}] \asymp \mathbb{E}[\|\mathbf{a}\|^r_{b^s_{p,q}(\mathbb{R}^d)}] < \infty$. ∎

4. Necessity of Properties A and A′. The weakest condition we consider is that $f \in B^s_{p,q}(\mathbb{R}^d)$ almost surely. In this section, we show that Property A is necessary for this when f follows a Besov prior, and that Property A′ is necessary if f follows a Besov–Bernoulli prior. For both priors, we start with the necessity for the finiteness of the Besov sequence norm of the wavelet coefficients of f, and then proceed with the necessity for the finiteness of the Besov norm of f itself.

Theorem 4.1. *Let $d \in \mathbb{N}$, $s, \alpha, \beta, \gamma, \theta \in \mathbb{R}$, and $p, q \in (0, \infty]$. Assume that $\mathbf{a}$ follows a Besov sequence prior with parameters $(\alpha, \beta, \gamma, \theta)$ and template random variable X satisfying*

$$\mathbb{P}(X \ne 0) > 0. \tag{4.1}$$

If $\mathbf{a} \in b^s_{p,q}(\mathbb{R}^d)$ almost surely, then also Property A is satisfied.

Proof. For $j \in \mathbb{N}_0$ and $t \in T_j$, we define the following (random) quantities:

$$\begin{aligned} Z_{j,t} &:= \|((1+\|m\|_\infty/2^j)^\beta|\xi_{j,t,m}|)_{m\in\mathbb{Z}^d}\|_{\ell_p}, \\ \eta_1 &:= \|((1+\|m\|_\infty)^\gamma|\xi_m|)_{m\in\mathbb{Z}^d}\|_{\ell_p}, \\ \eta_2 &:= \|(2^{j(s+d/2-d/p+\alpha)}(j+1)^\theta Z_{j,t})_{j\in\mathbb{N},\,t\in T_j}\|_{\ell_q}. \end{aligned} \tag{4.2}$$

We have $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} \overset{\text{a.s.}}{<} \infty$ if and only if $\eta_1 \overset{\text{a.s.}}{<} \infty$ and $\eta_2 \overset{\text{a.s.}}{<} \infty$.

Since $\mathbb{P}(X \ne 0) > 0$, there exists $R \ge 1$ such that $\mathbb{P}(1/R \le |X| \le R) > 0$. For $(j,t,m) \in J$, define $X_{j,t,m} := \mathbb{1}_{|\xi_{j,t,m}|\ge 1/R}$. We write $X_m := X_{j,t,m}$ when $j = 0$ (and consequently $t = (F, \dots, F)$).

Step 1a: *When $p \ne \infty$, we show that necessarily $\gamma < -d/p$.* Note that $|\xi_m|^p \ge R^{-p}X_m$. We thus deduce from

$$\infty \overset{\text{a.s.}}{>} \eta_1^p \ge R^{-p}\sum_{m\in\mathbb{Z}^d}(1+\|m\|_\infty)^{\gamma p}X_m$$

and Lemmas A.2 and A.3 that

$$\begin{aligned}\infty > \sum_{m\in\mathbb{Z}^d} \mathbb{E}[\min\{1,(1+\|m\|_\infty)^{\gamma p} X_m\}] \\ = \sum_{m\in\mathbb{Z}^d} \mathbb{E}[X_m]\min\{1,(1+\|m\|_\infty)^{\gamma p}\} \\ = \mathbb{E}[X_0]\sum_{m\in\mathbb{Z}^d}\min\{1,(1+\|m\|_\infty)^{\gamma p}\}.\end{aligned}$$

Because of $\mathbb{E}[X_0] > 0$, this implies first of all that $\min\{1,(1+\|m\|_\infty)^{\gamma p}\} \to 0$ as $\|m\|_\infty \to \infty$, and hence $\gamma p < 0$, so that $\min\{1,(1+\|m\|_\infty)^{\gamma p}\} = (1+\|m\|_\infty)^{\gamma p}$. Then Lemma A.4 yields $\gamma < -d/p$.

Step 1b: *When $p = \infty$, we show that necessarily $\gamma \le 0$.* If we assume $\gamma > 0$, then

$$\infty \overset{\text{a.s.}}{>} \eta_1 \ge R^{-1} \sup_{m\in\mathbb{Z}^d} (1+\|m\|_\infty)^{\gamma} X_m$$

implies that almost surely, $X_m = 1$ only holds for finitely many $m \in \mathbb{Z}^d$. Since the X_m are independent, the Borel–Cantelli lemma gives

$$\sum_{m\in\mathbb{Z}^d} \mathbb{P}(X_m = 1) < \infty,$$

which yields the desired contradiction since $\mathbb{P}(X_m = 1) = \mathbb{P}(|\xi_m| \ge 1/R) > 0$ is independent of m.

Step 2a: *When $p \ne \infty$, we show the remaining part of Property A.* Let $\xi^*_{j,t,m} := \xi_{j,t,m}\mathbb{1}_{1/R\le|\xi_{j,t,m}|\le R}$ for $(j,t,m) \in J$ be the truncation of $\xi_{j,t,m}$. Then the random variables $\xi^*_{j,t,m}$ for $(j,t,m) \in J$ are independent and identically distributed, with $\mathbb{E}[|\xi^*_{j,t,m}|^r] < \infty$ for all $r \in (0,\infty)$, and

$$\infty \overset{\text{a.s.}}{>} \eta_2 \ge \|(2^{j(s+d/2-d/p+\alpha)}(j+1)^{\theta} Y_{j,t}^{1/p})_{j\in\mathbb{N},\, t\in T_j}\|_{\ell_q} \tag{4.3}$$

where, for $j \in \mathbb{N}$ and $t \in T_j$, we set

$$Y_{j,t} := \|((1+\|m\|_\infty/2^j)^{\beta}|\xi^*_{j,t,m}|)_{m\in\mathbb{Z}^d}\|^p_{\ell_p}.$$

In particular, we see that almost surely, for $t^* = (M,\dots,M)$ we have

$$\infty > Y_{1,t^*} = \sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty/2)^{\beta p}|\xi^*_{1,t^*,m}|^p,$$

so that Lemmas A.2 and A.3 show for $X^*_{j,t,m} := \mathbb{1}_{1/R\le|\xi_{j,t,m}|\le R}$ that

$$\begin{aligned}\infty > \sum_{m\in\mathbb{Z}^d} \mathbb{E}[\min\{1,(1+\|m\|_\infty/2)^{\beta p}|\xi^*_{1,t^*,m}|^p\}] \\ \gtrsim R^{-p}\sum_{m\in\mathbb{Z}^d} \mathbb{E}[X^*_{1,t^*,m}\min\{1,(1+\|m\|_\infty)^{\beta p}\}]\end{aligned}$$

$$\gtrsim \sum_{m\in\mathbb{Z}^d} \min\{1, (1+\|m\|_\infty)^{\beta p}\}.$$

Using a similar argument to Step 1a, we obtain $\beta < -d/p$.

Now, the fact that the random variables $\xi_{j,t,m}$ are identically distributed and $\beta p < -d$ implies with Lemma A.4 that

$$\begin{aligned}\mathbb{E}[Y_{j,t}] &= \sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty/2^j)^{\beta p}\mathbb{E}[|\xi^*_{j,t,m}|^p] \\ &= \mathbb{E}[|\xi^*_0|^p] \sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty/2^j)^{\beta p} \asymp 2^{jd},\end{aligned} \tag{4.4}$$

uniformly with respect to $j\in\mathbb{N}$ and $t\in T_j$. Here, we wrote $\xi^*_0 := \xi^*_{0,(F,\dots,F),0}$ and used $\mathbb{P}(|\xi^*_0|>0) = \mathbb{P}(1/R \le |X| \le R) > 0$, so that $\mathbb{E}[|\xi^*_0|^p] > 0$.

Now let $\tau \in (1,\infty)$ and let $\tau' = \frac{\tau}{\tau-1}$, the conjugate exponent of τ. Since $\mathbb{E}[|\xi^*_0|^{p\tau}] < \infty$, Lemma A.4 and Hölder's inequality give

(4.5)

$$\begin{aligned}\mathbb{E}[Y_{j,t}^\tau] &= \mathbb{E}\Big[\Big(\sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty/2^j)^{\beta p}|\xi^*_{j,t,m}|^p\Big)^\tau\Big] \\ &= \mathbb{E}\Big[\Big(\sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty/2^j)^{\frac{\beta p}{\tau}}(1+\|m\|_\infty/2^j)^{\frac{\beta p}{\tau'}}|\xi^*_{j,t,m}|^p\Big)^\tau\Big] \\ &\le \mathbb{E}\Big[\Big(\sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty/2^j)^{\beta p}|\xi^*_{j,t,m}|^{p\tau}\Big)\Big(\sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty/2^j)^{\beta p}\Big)^{\tau/\tau'}\Big] \\ &= \mathbb{E}\Big[\sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty/2^j)^{\beta p}|\xi^*_{j,t,m}|^{p\tau}\Big]\Big(\sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty/2^j)^{\beta p}\Big)^{\tau/\tau'} \\ &\asymp \sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty/2^j)^{\beta p}\mathbb{E}[|\xi^*_0|^{p\tau}]2^{jd\tau/\tau'} \lesssim 2^{jd}2^{jd\tau/\tau'} = 2^{jd\tau}.\end{aligned}$$

Now use (4.4) and (4.5) with $\tau = 2$ and invoke Lemma A.5 (Paley–Zygmund inequality) with $X = Y_{j,t}$ and $\sigma = 1/2$. We obtain

$$\mathbb{P}(Y_{j,t} \ge \tfrac{1}{2}\mathbb{E}[Y_{j,t}]) \ge \frac{1}{4}\frac{\mathbb{E}[Y_{j,t}]^2}{\mathbb{E}[Y_{j,t}^2]} \gtrsim 1,$$

which means that there exists a constant $\kappa \in (0,1]$ such that

$$\mathbb{P}(Y_{j,t} \ge \tfrac{1}{2}\mathbb{E}[Y_{j,t}]) \ge \kappa$$

for all $j\in\mathbb{N}$ and $t\in T_j$.

The random variables $\chi_{j,t} := \mathbb{1}_{\{Y_{j,t}\ge\frac{1}{2}\mathbb{E}[Y_{j,t}]\}}$ with $j\in\mathbb{N}$ and $t\in T_j$ are independent and Bernoulli($p_{j,t}$)-distributed for some $p_{j,t} \ge \kappa$. Using (4.4) and (4.3), we obtain

$$\begin{aligned}\infty &\overset{\text{a.s.}}{>} \|(2^{j(s+d/2-d/p+\alpha)}(j+1)^{\theta}Y_{j,t}^{1/p})_{j\in\mathbb{N},\,t\in T_j}\|_{\ell_q}\\ &\geq \big\|(2^{j(s+d/2-d/p+\alpha)}(j+1)^{\theta}\big(\tfrac{1}{2}\mathbb{E}[Y_{j,t}]\big)^{1/p}\chi_{j,t})_{j\in\mathbb{N},\,t\in T_j}\big\|_{\ell_q}\\ &\gtrsim \|(2^{j(s+d/2-d/p+\alpha)}(j+1)^{\theta}(2^{jd})^{1/p}\chi_{j,t})_{j\in\mathbb{N},\,t\in T_j}\|_{\ell_q}\\ &\gtrsim \|(2^{j(s+d/2+\alpha)}(j+1)^{\theta}\chi_{j,t})_{j\in\mathbb{N},\,t\in T_j}\|_{\ell_q}.\end{aligned}$$

Let $t^* = (M,\dots,M)$. If $q=\infty$, note that the events $E_j := \{\chi_{j,t^*}=1\}$ for $j\in\mathbb{N}$ are independent, and we have $\sum_{j\in\mathbb{N}}\mathbb{P}(E_j)\geq\sum_{j\in\mathbb{N}}\kappa=\infty$. The Borel–Cantelli lemma thus shows that almost surely, the (random) set $I := \{j\in\mathbb{N} : \chi_{j,t^*}=1\}$ is infinite and

$$\infty \overset{\text{a.s.}}{>} \|(2^{j(s+d/2+\alpha)}(j+1)^{\theta}\chi_{j,t^*})_{j\in\mathbb{N}}\|_{\ell_q} \geq \sup_{j\in I} 2^{j(s+d/2+\alpha)}(j+1)^{\theta}.$$

This shows that $s+d/2+\alpha\leq 0$ and in case $s+d/2+\alpha=0$ also $\theta\leq 0$. On the other hand, if $q\neq\infty$, it follows from Lemmas A.2 and A.3 that

$$\begin{aligned}\infty &> \sum_{j\in\mathbb{N}}\sum_{t\in T_j}\mathbb{E}[\min\{1, 2^{jq(s+d/2+\alpha)}(j+1)^{\theta q}\chi_{j,t}\}]\\ &= \sum_{j\in\mathbb{N}}\sum_{t\in T_j}\mathbb{E}[\chi_{j,t}]\min\{1, 2^{jq(s+d/2+\alpha)}(j+1)^{\theta q}\}\\ &\geq \kappa\sum_{j\in\mathbb{N}}\min\{1, 2^{jq(s+d/2+\alpha)}(j+1)^{\theta q}\},\end{aligned}$$

i.e., $s+d/2+\alpha\leq 0$ and moreover $\theta q<-1$ if $s+d/2+\alpha=0$.

Step 2b: *When $p=\infty$, we show the remaining part of Property A.* Suppose $\beta>0$. Since $\eta_2 \overset{\text{a.s.}}{<} \infty$, we have

$$\infty \overset{\text{a.s.}}{>} Z_{1,t^*} \geq R^{-1}\sup_{m\in\mathbb{Z}^d}(1+\|m\|_\infty/2)^{\beta}X_{1,t^*,m}.$$

But then, almost surely, $X_{1,t^*,m}=1$ for only finitely many $m\in\mathbb{Z}^d$. The Borel–Cantelli lemma shows that $\sum_{m\in\mathbb{Z}^d}\mathbb{P}(X_{1,t^*,m}=1)<\infty$, which is absurd since $\mathbb{P}(X_{1,t^*,m}=1)=\mathbb{P}(|\xi_{1,t^*,m}|\geq 1/R)>0$ is independent of $m\in\mathbb{Z}^d$. Hence $\beta\leq 0$.

Now that we know $\beta\leq 0$, we conclude

$$\infty \overset{\text{a.s.}}{>} \eta_2 \geq \|(2^{j(s+d/2+\alpha)}(j+1)^{\theta}\xi_{j,t^*,0})_{j\in\mathbb{N}}\|_{\ell_q}.$$

If $q=\infty$, then the Borel–Cantelli lemma implies as in Step 1b that $s+d/2+\alpha\leq 0$ and that $\theta\leq 0$ if $s+d/2+\alpha=0$.

On the other hand, if $q\neq\infty$, we know that with $t^*=(M,\dots,M)$,

$$\begin{aligned}\infty &\overset{\text{a.s.}}{>} \|(2^{j(s+d/2+\alpha)}(j+1)^{\theta}\xi_{j,t^*,0})_{j\in\mathbb{N}}\|_{\ell_q}^q = \sum_{j\in\mathbb{N}}2^{jq(s+d/2+\alpha)}(j+1)^{\theta q}|\xi_{j,t^*,0}|^q\\ &\geq R^{-q}\sum_{j\in\mathbb{N}}2^{jq(s+d/2+\alpha)}(j+1)^{\theta q}X_{j,t^*,0}.\end{aligned}$$

The random variables $X_{j,t^*,0}$ are independent and identically Bernoulli(ϱ)-distributed with some parameter $\varrho \in (0,1]$. Lemmas A.2 and A.3 now give

$$\begin{aligned}\infty &> \sum_{j\in\mathbb{N}} \mathbb{E}[\min\{1, 2^{jq(s+d/2+\alpha)}(j+1)^{\theta q} X_{j,t^*,0}\}] \\ &= \varrho \sum_{j\in\mathbb{N}} \min\{1, 2^{jq(s+d/2+\alpha)}(j+1)^{\theta q}\}.\end{aligned}$$

This first of all implies that $\lim_{j\to\infty} \min\{1, 2^{jq(s+d/2+\alpha)}(j+1)^{\theta q}\} = 0$, and hence $s + d/2 + \alpha \le 0$. If $s + d/2 + \alpha = 0$, it also follows that $\sum_{j\in\mathbb{N}}(j+1)^{\theta q} < \infty$, i.e., $\theta q < -1$. ∎

Remark 4.2. The subtle distinction between the conditions of Theorem 4.1 and Theorem 4.3 below is ignored in most of the literature, where it is simply pretended that if $f = S(\mathbf{a})$, then the conditions $f \in B^s_{p,q}(\mathbb{R}^d)$ and $\mathbf{a} \in b^s_{p,q}(\mathbb{R}^d)$ are equivalent. However, one only knows that

(i) if $\mathbf{a}, \overline{\mathbf{a}} \in b^s_{p,q}(\mathbb{R}^d)$ with $S(\mathbf{a}) = f = S(\overline{\mathbf{a}})$, then $\mathbf{a} = \overline{\mathbf{a}}$,
(ii) if $f \in B^s_{p,q}(\mathbb{R}^d)$, then $f = S(\overline{\mathbf{a}})$ for *some* $\overline{\mathbf{a}} \in b^s_{p,q}(\mathbb{R}^d)$.

Without knowing anything about the decay of the coefficient sequence $\mathbf{a}$, it is not clear that this implies $\mathbf{a} = \overline{\mathbf{a}} \in b^s_{p,q}(\mathbb{R}^d)$.

In fact, [12, Remark 4.7] provides an example of a 1D wavelet system (namely, the system generated by the Haar wavelet) for which there are two distinct coefficient sequences (which even both belong to $b^s_{p,q}(\mathbb{R})$ for $s < 1/p - 1$) representing the same tempered distribution. We expect that a similar phenomenon can also occur for Daubechies wavelets.

Theorem 4.3. *Let $d, k \in \mathbb{N}$, $s, \alpha, \beta, \gamma, \theta \in \mathbb{R}$, and $p, q \in (0,\infty]$. Assume that f follows a Besov prior with parameters $(\alpha, \beta, \gamma, \theta, k)$ and template random variable X with*

$$\begin{cases} \exists \varepsilon > 0 : \mathbb{E}\big[|X|^{(1+\varepsilon)p}\big] < \infty & \text{if } p \in (0,\infty), \\ \exists R > 0 : |X| \overset{\text{a.s.}}{\le} R & \text{if } p = \infty, \end{cases}$$

and

$$\mathbb{P}(X \neq 0) > 0.$$

If $f \in B^s_{p,q}(\mathbb{R}^d)$ almost surely and if Properties B and B′ are satisfied, then also Property A is satisfied.

Proof. From Property B, it follows that $d(1/p-1) - k < s < k$ if $p \in (0,1)$ and $-k < s < k$ otherwise. Taking Property B' into account, we may choose $\widetilde{s} \in \mathbb{R}$ such that $d(1/p-1) - k < \widetilde{s} < s$ and $\widetilde{s} + d/2 + \alpha < 0$ when $p \in (0,1)$. On the other hand, if $p \in [1,\infty]$, we may choose $\widetilde{s} \in \mathbb{R}$ such that $-k < \widetilde{s} < s$ and $\widetilde{s} + d/2 + \alpha < 0$. Note that in each of the two cases Property B is satisfied with $\widetilde{s}$ in place of s. Choose $\sigma < 0$ small enough such

that $\gamma + \sigma < -d/p$ and $\beta + \sigma < -d/p$ when $p \neq \infty$, and $\gamma + \sigma \leq 0$ and $\beta + \sigma \leq 0$ when $p = \infty$. Since (3.3) holds, we conclude from Theorem 3.3 that $\overline{\mathbf{a}} = (\overline{a}_{j,t,m})_{(j,t,m)\in J} \in b^{\widetilde{s}}_{p,q}(\mathbb{R}^d)$ almost surely where $\xi_{j,t,m} \sim X$ are the i.i.d. random variables from (2.3) and

$$\overline{a}_{j,t,m} := 2^{j\alpha}(j+1)^{\theta}(1 + \|m\|_\infty / 2^j)^{\beta+\sigma+(\gamma-\beta)\mathbb{1}_{j=0}} \xi_{j,t,m}.$$

With $w_\sigma : \mathbb{R}^d \to \mathbb{R}$, $w_\sigma(x) = (1+\|x\|_2^2)^{\sigma/2}$, we then see in view of Remark 2.7 that $\mathbf{a} = (a_{j,t,m})_{(j,t,m)\in J} \in b^{\widetilde{s}}_{p,q}(\mathbb{R}^d, w_\sigma)$, where

$$a_{j,t,m} = 2^{j\alpha}(j+1)^{\theta}(1 + \|m\|_\infty / 2^j)^{\beta+(\gamma-\beta)\mathbb{1}_{j=0}} \xi_{j,t,m}.$$

If $f \in B^s_{p,q}(\mathbb{R}^d)$ almost surely, then almost surely $f = S(\widetilde{\mathbf{a}})$ for some $\widetilde{\mathbf{a}} \in b^s_{p,q}(\mathbb{R}^d)$ by Theorem 2.3. From $\sigma < 0$ and $s > \widetilde{s}$, we get the embeddings

$$b^s_{p,q}(\mathbb{R}^d) \hookrightarrow b^s_{p,q}(\mathbb{R}^d, w_\sigma) \hookrightarrow b^{\widetilde{s}}_{p,q}(\mathbb{R}^d, w_\sigma).$$

But this implies $\mathbf{a}, \widetilde{\mathbf{a}} \in b^{\widetilde{s}}_{p,q}(\mathbb{R}^d, w_\sigma)$ almost surely with $S(\mathbf{a}) = f = S(\widetilde{\mathbf{a}})$, from which it follows that $\mathbf{a} = \widetilde{\mathbf{a}} \in b^s_{p,q}(\mathbb{R}^d)$ almost surely by Theorem 2.3. From Theorem 4.1, we conclude that Property A holds. ∎

We now show that Property A′ is necessary for the coefficient sequence in (2.4) to almost surely belong to the Besov sequence space $b^s_{p,q}(\mathbb{R}^d)$. Our a priori assumptions $\mu \geq -d$ and $\nu \geq -d$ in the following theorem are parallel to the assumption $\beta \leq 1$ in [1, Theorem 1].

THEOREM 4.4. *Let $d \in \mathbb{N}$, $s, \alpha, \beta, \gamma \in \mathbb{R}$, $p, q \in (0, \infty]$, and $\mu, \nu \in [-d, 0]$. Assume that $\mathbf{a}$ follows a Besov–Bernoulli sequence prior with parameters $(\alpha, \beta, \gamma, \mu, \nu)$ and template variable X satisfying*

$$\mathbb{P}(X \neq 0) > 0.$$

If $\mathbf{a} \in b^s_{p,q}(\mathbb{R}^d)$ almost surely, then also Property A′ is satisfied.

Proof. For $j \in \mathbb{N}_0$ and $t \in T_j$, we define the following (random) quantities:

$$\begin{aligned}
\widetilde{Z}_{j,t} &:= \|((1 + \|m\|_\infty / 2^j)^{\beta} \lambda_{j,t,m} |\xi_{j,t,m}|)_{m\in\mathbb{Z}^d}\|_{\ell_p}, \\
\widetilde{\eta}_1 &:= \|((1 + \|m\|_\infty)^{\gamma} \lambda_m |\xi_m|)_{m\in\mathbb{Z}^d}\|_{\ell_p}, \\
\widetilde{\eta}_2 &:= \|(2^{j(s+d/2-d/p+\alpha)} \widetilde{Z}_{j,t})_{j\in\mathbb{N},\, t\in T_j}\|_{\ell_q}.
\end{aligned}$$

We have $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} \overset{\text{a.s.}}{<} \infty$ if and only if $\widetilde{\eta_1} \overset{\text{a.s.}}{<} \infty$ and $\widetilde{\eta_2} \overset{\text{a.s.}}{<} \infty$.

Since $\mathbb{P}(X \neq 0) > 0$, there exists $R \geq 1$ such that $\mathbb{P}(1/R \leq |X| \leq R) > 0$. For $(j,t,m) \in \mathbb{N}_0 \times \{F, M\}^d \times \mathbb{Z}^d$, define $X_{j,t,m} := \mathbb{1}_{|\xi_{j,t,m}| \geq 1/R}$. We write $X_m := X_{j,t,m}$, $\lambda_m := \lambda_{j,t,m}$, and $\xi_m := \xi_{j,t,m}$ when $j = 0$ (and consequently $t = (F, \ldots, F)$).

Step 1a: *When $p \neq \infty$, we show that necessarily $\nu + \gamma p < -d$.* Note that $|\xi_m|^p \geq R^{-p} X_m$. We thus deduce from

$$\infty \overset{\text{a.s.}}{>} \widetilde{\eta}_1^p \geq R^{-p} \sum_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty)^{\gamma p} \lambda_m X_m$$

and Lemmas A.2 and A.3 that

$$\begin{aligned}
\infty &> \sum_{m \in \mathbb{Z}^d} \mathbb{E}[\min\{1, (1 + \|m\|_\infty)^{\gamma p} \lambda_m X_m\}] \\
&= \sum_{m \in \mathbb{Z}^d} \mathbb{E}[\lambda_m X_m] \min\{1, (1 + \|m\|_\infty)^{\gamma p}\} \\
&= \mathbb{E}[X_0] \sum_{m \in \mathbb{Z}^d} \mathbb{E}[\lambda_m] \min\{1, (1 + \|m\|_\infty)^{\gamma p}\} \\
&= \mathbb{E}[X_0] \sum_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty)^{\nu} \min\{1, (1 + \|m\|_\infty)^{\gamma p}\} \\
&\asymp \sum_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty)^{\min\{\nu, \nu + \gamma p\}}.
\end{aligned}$$

Then Lemma A.4 yields $\min\{\nu, \nu + \gamma p\} < -d$, or equivalently $\min\{0, \gamma\} + \frac{d+\nu}{p} < 0$. Since $\nu \geq -d$, we obtain $\gamma < 0$, i.e., $0 > \min\{0, \gamma\} + \frac{d+\nu}{p} = \gamma + \frac{d+\nu}{p}$.

Step 1b: *When $p = \infty$, we show that necessarily $\gamma \leq 0$.* Let us assume towards a contradiction that $\gamma > 0$. Then

$$\infty \overset{\text{a.s.}}{>} \widetilde{\eta}_1 \geq R^{-1} \sup_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty)^{\gamma} \lambda_m X_m$$

implies that almost surely, $\lambda_m X_m = 1$ only holds for finitely many $m \in \mathbb{Z}^d$. The Borel–Cantelli lemma gives

$$\begin{aligned}
\infty &> \sum_{m \in \mathbb{Z}^d} \mathbb{P}(\lambda_m X_m = 1) = \sum_{m \in \mathbb{Z}^d} \mathbb{P}(\lambda_m = 1) \mathbb{P}(X_m = 1) \\
&= \sum_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty)^{\nu} \mathbb{P}(X_m = 1),
\end{aligned}$$

which yields the desired contradiction since $\mathbb{P}(X_m = 1) = \mathbb{P}(|\xi_m| \geq 1/R) > 0$ is independent of m, and $\nu \geq -d$; see Lemma A.4.

Step 2a: *When $p \neq \infty$, we show the remaining part of Property A'.* Let $\xi^*_{j,t,m} := \xi_{j,t,m} \mathbb{1}_{1/R \leq |\xi_{j,t,m}| \leq R}$ for $(j,t,m) \in J$ be the truncation of $\xi_{j,t,m}$. Then the random variables $\xi^*_{j,t,m}$ are independent and identically distributed, with $\mathbb{E}[|\xi^*_{j,t,m}|^r] < \infty$ for all $r \in (0, \infty)$, and

$$\infty \overset{\text{a.s.}}{>} \widetilde{\eta}_2 \geq \|(2^{j(s+d/2-d/p+\alpha)} \widetilde{Y}_{j,t}^{1/p})_{j \in \mathbb{N},\, t \in T_j}\|_{\ell_q}$$

where, for $j \in \mathbb{N}$ and $t \in T_j$, we set

$$\widetilde{Y}_{j,t} := \|((1 + \|m\|_\infty / 2^j)^{\beta} \lambda_{j,t,m} |\xi^*_{j,t,m}|)_{m \in \mathbb{Z}^d}\|_{\ell_p}^p.$$

In particular, writing $t^* = (M, \dots, M)$, we see that almost surely, we have

$$\infty > \widetilde{Y}_{1,t^*} = \sum_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty/2)^{\beta p} \lambda_{1,t^*,m} |\xi^*_{1,t^*,m}|^p,$$

so that Lemmas A.2 and A.3 show for $X^*_{j,t,m} := \mathbb{1}_{1/R \leq |\xi_{j,t,m}| \leq R}$ that

$$\begin{aligned}
\infty &> \sum_{m \in \mathbb{Z}^d} \mathbb{E}[\min\{1, (1 + \|m\|_\infty/2)^{\beta p} \lambda_{1,t^*,m} |\xi^*_{1,t^*,m}|^p\}] \\
&\gtrsim R^{-p} \sum_{m \in \mathbb{Z}^d} \mathbb{E}[X^*_{1,t^*,m} \min\{1, (1 + \|m\|_\infty)^{\beta p} \lambda_{1,t^*,m}\}] \\
&= R^{-p} \sum_{m \in \mathbb{Z}^d} \mathbb{E}[X^*_{1,t^*,m}] \mathbb{E}[\min\{1, (1 + \|m\|_\infty)^{\beta p} \lambda_{1,t^*,m}\}] \\
&= R^{-p} \sum_{m \in \mathbb{Z}^d} \mathbb{E}[X^*_{1,t^*,m}] \mathbb{E}[\lambda_{1,t^*,m}] \min\{1, (1 + \|m\|_\infty)^{\beta p}\} \\
&\gtrsim \sum_{m \in \mathbb{Z}^d} \mathbb{E}[\lambda_{1,t^*,m}] \min\{1, (1 + \|m\|_\infty)^{\beta p}\} \\
&\gtrsim \sum_{m \in \mathbb{Z}^d} \min\{(1 + \|m\|_\infty)^{\nu}, (1 + \|m\|_\infty)^{\nu + \beta p}\}.
\end{aligned}$$

Using a similar argument to Step 1a, we obtain $\min\{\nu, \nu + \beta p\} < -d$, which in turn because of $\nu \geq -d$ implies $\beta + \frac{d+\nu}{p} < 0$, and in particular $\beta < 0$. Further, we have

$$\infty \overset{\text{a.s.}}{>} \widetilde{\eta}_2 \geq \|(2^{j(s+d/2-d/p+\alpha)} U_{j,t}^{1/p})_{j \in \mathbb{N},\, t \in T_j}\|_{\ell_q},$$

where, for $j \in \mathbb{N}$ and $t \in T_j$, we set

$$U_{j,t} := \sum_{\substack{m \in \mathbb{Z}^d \\ \|m\|_\infty \leq 2^j}} (1 + \|m\|_\infty/2^j)^{\beta p} \lambda_{j,t,m} |\xi^*_{j,t,m}|^p.$$

In particular, we see that

$$\infty \overset{\text{a.s.}}{>} U_{j,t} \asymp \sum_{\substack{m \in \mathbb{Z}^d \\ \|m\|_\infty \leq 2^j}} \lambda_{j,t,m} |\xi^*_{j,t,m}|^p \asymp \sum_{\substack{m \in \mathbb{Z}^d \\ \|m\|_\infty \leq 2^j}} \lambda_{j,t,m} X^*_{j,t,m}.$$

Since the random variables $X^*_{j,t,m}$ are i.i.d. Bernoulli (ρ)-distributed for some $\rho \in (0, 1]$, we see that $\lambda_{j,t,m} X^*_{j,t,m}$ is Bernoulli$(2^{j\mu}(1 + \|m\|_\infty/2^j)^\nu \rho)$-distributed.

If $\|m\|_\infty \leq 2^j$, we have

$$2^{j\mu}(1 + \|m\|_\infty/2^j)^\nu \rho \geq 2^{j\mu + \nu} \rho.$$

Thus there exists a Bernoulli$(2^{j\mu+\nu}\rho)$-distributed random variable $B_{j,t,m}$ such that $\lambda_{j,t,m} X^*_{j,t,m} \geq B_{j,t,m}$ and the family $(B_{j,t,m})_{(j,t,m)}$ is independent.

We conclude

$$\infty \overset{\text{a.s.}}{>} U_{j,t} \asymp \sum_{\substack{m\in\mathbb{Z}^d\\ \|m\|_\infty\le 2^j}} \lambda_{j,t,m} X^*_{j,t,m} \ge \sum_{\substack{m\in\mathbb{Z}^d\\ \|m\|_\infty\le 2^j}} B_{j,t,m} =: A_{j,t} \tag{4.6}$$

with $A_{j,t} \sim \mathrm{Bin}((1+2^{j+1})^d, 2^{j\mu+\nu}\rho)$ following a binomial distribution. Therefore, Lemma A.6 implies for $\tau \ge 1$ because of $\mu \ge -d$ that

$$\begin{aligned}\mathbb{E}[A_{j,t}^\tau] &\lesssim \max\{(1+2^{j+1})^d 2^{j\mu+\nu}\rho, ((1+2^{j+1})^d 2^{j\mu+\nu}\rho)^\tau\}\\ &\asymp \max\{2^{j(\mu+d)}, 2^{j\tau(\mu+d)}\} = 2^{j\tau(\mu+d)}.\end{aligned}$$

Set $\tau = 2$ and invoke Lemma A.5 (Paley–Zygmund inequality) with $X = A_{j,t}$ and $\sigma = 1/2$ to obtain

$$\mathbb{P}\big(A_{j,t} \ge \tfrac{1}{2}\mathbb{E}[A_{j,t}]\big) \ge \frac{1}{4}\frac{\mathbb{E}[A_{j,t}]^2}{\mathbb{E}[A_{j,t}^2]} \gtrsim 1,$$

which means that there exists a constant $\kappa > 0$ such that

$$\mathbb{P}\big(A_{j,t} \ge \tfrac{1}{2}\mathbb{E}[A_{j,t}]\big) \ge \kappa \tag{4.7}$$

for all $j \in \mathbb{N}$ and $t \in T_j$.

The random variables $\chi_{j,t} := \mathbb{1}_{\{A_{j,t} \ge \frac{1}{2}\mathbb{E}[A_{j,t}]\}}$ with $j \in \mathbb{N}$ and $t \in T_j$ are independent and Bernoulli$(p_{j,t})$-distributed for some $p_{j,t} \ge \kappa$. Recall from (4.6) that $U_{j,t} \gtrsim A_{j,t}$. Hence,

$$\begin{aligned}\infty &\overset{\text{a.s.}}{>} \|(2^{j(s+d/2-d/p+\alpha)} A_{j,t}^{1/p})_{j\in\mathbb{N},\, t\in T_j}\|_{\ell_q}\\ &\ge \big\|(2^{j(s+d/2-d/p+\alpha)} (\tfrac{1}{2}\mathbb{E}[A_{j,t}])^{1/p} \chi_{j,t})_{j\in\mathbb{N},\, t\in T_j}\big\|_{\ell_q}\\ &\gtrsim \|(2^{j(s+d/2-d/p+\alpha)} (2^{j(\mu+d)})^{1/p} \chi_{j,t})_{j\in\mathbb{N},\, t\in T_j}\|_{\ell_q}\\ &\gtrsim \|(2^{j(s+d/2+\alpha+\mu/p)} \chi_{j,t})_{j\in\mathbb{N},\, t\in T_j}\|_{\ell_q}.\end{aligned}$$

If $q = \infty$, let $t^* = (M, \dots, M)$ and note that the events $E_j := \{\chi_{j,t^*} = 1\}$ for $j \in \mathbb{N}$ are independent. Furthermore, $\sum_{j\in\mathbb{N}} \mathbb{P}(E_j) \ge \sum_{j\in\mathbb{N}} \kappa = \infty$. The Borel–Cantelli lemma thus implies that almost surely, the set $I := \{j \in \mathbb{N} : \chi_{j,t^*} = 1\}$ is an infinite set and

$$\infty \overset{\text{a.s.}}{>} \|(2^{j(s+d/2+\alpha+\mu/p)} \chi_{j,t})_{j\in\mathbb{N},\, t\in T_j}\|_{\ell_q} \ge \sup_{j\in I} 2^{j(s+d/2+\alpha+\mu/p)}.$$

This shows that $s + d/2 + \alpha + \mu/p \le 0$. On the other hand, if $q \ne \infty$, it follows from Lemmas A.2 and A.3 that

$$\begin{aligned}&\infty > \sum_{j\in\mathbb{N}} \sum_{t\in T_j} \mathbb{E}[\min\{1, 2^{jq(s+d/2+\alpha+\mu/p)} \chi_{j,t}\}]\\ &= \sum_{j\in\mathbb{N}} \sum_{t\in T_j} \mathbb{E}[\chi_{j,t}] \min\{1, 2^{jq(s+d/2+\alpha+\mu/p)}\} \ge \kappa \sum_{j\in\mathbb{N}} \min\{1, 2^{jq(s+d/2+\alpha+\mu/p)}\},\end{aligned}$$

and thus $s + d/2 + \alpha + \mu/p < 0$.

Step 2b: *When $p = \infty$, we show the remaining part of Property A′.* Suppose towards a contradiction that $\beta > 0$. Since $\widetilde{\eta}_2 \overset{\text{a.s.}}{<} \infty$, with $t^* = (M, \dots, M)$ we have

$$\infty \overset{\text{a.s.}}{>} \widetilde{Z}_{1,t^*} \geq R^{-1} \sup_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty/2)^\beta \lambda_{1,t^*,m} X_{1,t^*,m}.$$

But then, almost surely, $\lambda_{1,t^*,m} X_{1,t^*,m} = 1$ for only finitely many $m \in \mathbb{Z}^d$. Since the random variables $\lambda_{1,t^*,m} X_{1,t^*,m}$ for $m \in \mathbb{Z}^d$ are independent, the Borel–Cantelli lemma thus shows that

$$\begin{aligned} \infty &> \sum_{m \in \mathbb{Z}^d} \mathbb{P}(\lambda_{1,t^*,m} X_{1,t^*,m} = 1) = \sum_{m \in \mathbb{Z}^d} \mathbb{P}(\lambda_{1,t^*,m} = 1) \mathbb{P}(X_{1,t^*,m} = 1) \\ &= \sum_{m \in \mathbb{Z}^d} 2^\mu (1 + \|m\|_\infty/2)^\nu \mathbb{P}(X_{1,t^*,m} = 1), \end{aligned}$$

which is absurd since $\mathbb{P}(X_{1,t^*,m} = 1) = \mathbb{P}(|\xi_{1,t^*,m}| \geq 1/R) > 0$ is independent of $m \in \mathbb{Z}^d$, and $\nu \geq -d$ (see Lemma A.4). Hence $\beta \leq 0$.

Further, with $X^*_{j,t,m} := \mathbb{1}_{1/R \leq |\xi_{j,t,m}| \leq R}$ as before, for all $j \in \mathbb{N}$ and $t \in T_j$ we have

$$\begin{aligned} \widetilde{Z}_{j,t} &\geq \sup_{\substack{m \in \mathbb{Z}^d \\ \|m\|_\infty \leq 2^j}} (1 + \|m\|_\infty/2^j)^\beta \lambda_{j,t,m} |\xi_{j,t,m}| \\ &\gtrsim \sup_{\substack{m \in \mathbb{Z}^d \\ \|m\|_\infty \leq 2^j}} \lambda_{j,t,m} X^*_{j,t,m} =: \widetilde{U}_{j,t} \in \{0, 1\}, \end{aligned}$$

which in turn implies

$$\infty \overset{\text{a.s.}}{>} \widetilde{\eta}_2 \geq \|(2^{j(s+d/2+\alpha)} \widetilde{U}_{j,t^*})_{j \in \mathbb{N}}\|_{\ell_q}, \tag{4.8}$$

where again $t^* = (M, \dots, M)$. Let us note that (4.7) and (4.6) yield the existence of $c > 0$ such that

$$\mathbb{E}[\widetilde{U}_{j,t^*}] = \mathbb{P}(\widetilde{U}_{j,t^*} = 1) \geq \mathbb{P}(A_{j,t^*} \neq 0) \geq \mathbb{P}(A_{j,t^*} \geq c 2^{j(\mu+d)}) \geq \kappa > 0.$$

If $q = \infty$, then the Borel–Cantelli lemma and (4.8) imply as in Step 1b that $s + d/2 + \alpha \leq 0$. On the other hand, if $q \neq \infty$, then (4.8) also implies $\infty \overset{\text{a.s.}}{>} \|(2^{j(s+d/2+\alpha)} \widetilde{U}_{j,t^*})_{j \in \mathbb{N}}\|_{\ell_\infty}$, which in turn yields $s + d/2 + \alpha \leq 0$. Moreover, Lemmas A.2 and A.3 give

$$\infty > \sum_{j \in \mathbb{N}} \mathbb{E}[\min\{1, 2^{jq(s+d/2+\alpha)} \widetilde{U}_{j,t^*}\}] = \sum_{j \in \mathbb{N}} 2^{jq(s+d/2+\alpha)} \mathbb{E}[\widetilde{U}_{j,t^*}].$$

As $\mathbb{E}[\widetilde{U}_{j,t^*}] \geq \kappa$ is independent of $j \in \mathbb{N}_0$, it follows that $s + d/2 + \alpha < 0$. ∎

Theorem 4.5. *Let $d, k \in \mathbb{N}$, $s, \alpha, \beta, \gamma \in \mathbb{R}$, $p, q \in (0, \infty]$, and $\mu, \nu \in [-d, 0]$. Assume that f follows a Besov–Bernoulli prior with parameters*

$(\alpha, \beta, \gamma, \mu, \nu, k)$ and template random variable X with

$$\begin{cases} \exists \varepsilon > 0 : \mathbb{E}\big[|X|^{(1+\varepsilon)p}\big] < \infty & \text{if } p \in (0, \infty), \\ \exists R > 0 : |X| \overset{\text{a.s.}}{\leq} R & \text{if } p = \infty, \end{cases}$$

and

$$\mathbb{P}(X \neq 0) > 0.$$

If $f \in B^s_{p,q}(\mathbb{R}^d)$ almost surely and if Properties B and B′ are satisfied, then also Property A′ is satisfied.

Proof. From Property B, it follows that $d(1/p - 1) - k < s < k$ if $p \in (0, 1)$ and $-k < s < k$ otherwise. Taking Property B′ into account, we may choose $\widetilde{s} \in \mathbb{R}$ such that $d(1/p - 1) - k < \widetilde{s} < s$ and $\widetilde{s} + d/2 + \alpha < 0$ when $p \in (0, 1)$. On the other hand, if $p \in [1, \infty]$, we may choose $\widetilde{s} \in \mathbb{R}$ such that $-k < \widetilde{s} < s$ and $\widetilde{s} + d/2 + \alpha < 0$. Note that then Property B is satisfied with $\widetilde{s}$ in place of s. Choose $\sigma < 0$ small enough such that $\gamma + \sigma + \frac{d+\nu}{p} < 0$ and $\beta + \sigma + \frac{d+\nu}{p} < 0$. Since (3.3) holds, we conclude from Theorem 3.8 with $r = p$ that $\overline{\mathbf{a}} = (\overline{a}_{j,t,m})_{(j,t,m)\in J} \in b^{\tilde{s}}_{p,q}(\mathbb{R}^d)$ almost surely, where $\lambda_{j,t,m}$ and $\xi_{j,t,m}$ are the random variables from (2.4) and

$$\overline{a}_{j,t,m} := 2^{j\alpha}(1 + \|m\|_\infty / 2^j)^{\beta+\sigma+(\gamma-\beta)\mathbb{1}_{j=0}} \lambda_{j,t,m} \xi_{j,t,m}.$$

Here, we use the fact that Property A″ is satisfied (for $\beta+\sigma$ and $\gamma+\sigma$ instead of β, γ and for $\widetilde{s}$ instead of s), since $\mu \leq 0$. With $w_\sigma : \mathbb{R}^d \to \mathbb{R}$, $w_\sigma(x) = (1 + \|x\|_2^2)^{\sigma/2}$, we then see in view of Remark 2.7 that $\mathbf{a} = (a_{j,t,m})_{(j,t,m)\in J} \in b^{\tilde{s}}_{p,q}(\mathbb{R}^d, w_\sigma)$, where

$$a_{j,t,m} = 2^{j\alpha}(1 + \|m\|_\infty / 2^j)^{\beta+(\gamma-\beta)\mathbb{1}_{j=0}} \lambda_{j,t,m} \xi_{j,t,m}.$$

If $f \in B^s_{p,q}(\mathbb{R}^d)$ almost surely, then almost surely $f = S(\widetilde{\mathbf{a}})$ for some $\widetilde{\mathbf{a}} \in b^s_{p,q}(\mathbb{R}^d)$ by Theorem 2.3. From $\sigma < 0$ and $s > \widetilde{s}$, we get the embeddings

$$b^s_{p,q}(\mathbb{R}^d) \hookrightarrow b^s_{p,q}(\mathbb{R}^d, w_\sigma) \hookrightarrow b^{\tilde{s}}_{p,q}(\mathbb{R}^d, w_\sigma).$$

But this implies $\mathbf{a}, \widetilde{\mathbf{a}} \in b^{\tilde{s}}_{p,q}(\mathbb{R}^d, w_\sigma)$ almost surely with $S(\mathbf{a}) = f = S(\widetilde{\mathbf{a}})$, from which it follows that almost surely $\mathbf{a} = \widetilde{\mathbf{a}} \in b^s_{p,q}(\mathbb{R}^d)$ by Theorem 2.3. From Theorem 4.4, we conclude that Property A′ holds. ∎

5. Discussion of the conditions on the random variables. When $p \neq \infty$, the condition (3.3) in Theorem 3.3 can be replaced by a weaker condition. In the following, let $\log_2^+(x) := \max\{0, \log_2(x)\}$.

THEOREM 5.1. *Let $d, k \in \mathbb{N}$, $s, \alpha, \beta, \gamma, \theta \in \mathbb{R}$, and $p, q \in (0, \infty]$. Assume that $\mathbf{a}$ follows a Besov sequence prior and f follows a Besov prior, both with*

parameters $(\alpha, \beta, \gamma, \theta, k)$ *and template random variable* X*, satisfying*

$$\begin{cases} \exists \varepsilon > 0 : \mathbb{E}[|X|^p \cdot \log_2^+(|X|)] < \infty & \text{if } p \in (0, \infty), \\ \exists R > 0 : |X| \overset{\text{a.s.}}{\leq} R & \text{if } p = \infty. \end{cases} \tag{5.1}$$

If Property A holds, then $\mathbf{a}$ *almost surely belongs to* $b^s_{p,q}(\mathbb{R}^d)$*. Moreover, if Properties A and B are satisfied, then* f *belongs almost surely to* $B^s_{p,q}(\mathbb{R}^d)$*.*

Proof. The proof follows the same lines as the one of Theorem 3.3 but we need a new argument for $\Xi \overset{\text{a.s.}}{<} \infty$, with Ξ from Lemma 3.1 when $p \neq \infty$. This is done by invoking Lemma A.10 for $X_{j,\tau} = |\xi_{j,t,\phi(\tau)}|^p$, where again $\phi : \mathbb{N} \to \mathbb{Z}^d$ is a bijection such that $\tau \mapsto \|\phi(\tau)\|_\infty$ is nondecreasing. Regarding the applicability of Lemma A.10, note that

$$\begin{aligned} \mathbb{E}[X_{j,\tau} \cdot \log_2^+(X_{j,\tau})] &= \mathbb{E}[|\xi_{j,t,\phi(\tau)}|^p \cdot \log_2^+(|\xi_{j,t,\phi(\tau)}|^p)] \\ &= p\mathbb{E}[|X|^p \cdot \log_2^+(|X|)]. \ \blacksquare \end{aligned}$$

The preceding theorem raises the question of the sharpness of conditions (3.3) and (5.1). We address the case $p \neq \infty$ first. We only consider the condition $\mathbf{a} = (a_{j,t,m})_{(j,t,m)\in J} \in b^s_{p,q}(\mathbb{R}^d)$ almost surely, with $a_{j,t,m}$ as defined in (2.5), instead of the condition $f \in B^s_{p,q}(\mathbb{R}^d)$ almost surely. See the remark preceding Theorem 4.3 for a discussion of the relation between these two conditions.

Proposition 5.2. *Let* $d \in \mathbb{N}$*,* $s, \alpha, \beta, \gamma, \theta \in \mathbb{R}$*,* $p \in (0, \infty)$*, and* $q \in (0, \infty]$*. Assume that* $\mathbf{a} = (a_{j,t,m})_{(j,t,m)\in J}$ *follows a Besov sequence prior with parameters* $(\alpha, \beta, \gamma, \theta)$ *and template random variable* X *satisfying* $\mathbb{P}(X \neq 0) > 0$*. If* $\mathbf{a} \in b^s_{p,q}(\mathbb{R}^d)$ *almost surely, then* $\gamma < 0$ *and* $\mathbb{E}[|X|^{-d/\gamma}] < \infty$*.*

Proof. Let $\widetilde{\gamma} := \min\{0, \gamma\}$. Since $\mathbf{a} \in b^s_{p,q}(\mathbb{R}^d)$ almost surely, we know that also

$$\|((1 + \|m\|_\infty)^{\widetilde{\gamma}} |\xi_m|)_{m\in\mathbb{Z}^d}\|_{\ell_p} \leq \|((1 + \|m\|_\infty)^{\gamma} |\xi_m|)_{m\in\mathbb{Z}^d}\|_{\ell_p} \overset{\text{a.s.}}{<} \infty.$$

By Lemma A.2, this implies

$$\infty > \sum_{m\in\mathbb{Z}^d} \mathbb{E}[\min\{1, (1 + \|m\|_\infty)^{\widetilde{\gamma} p} |\xi_m|^p\}].$$

If we had $\widetilde{\gamma} = 0$, then this would imply

$$\infty > \sum_{m\in\mathbb{Z}^d} \mathbb{E}[\min\{1, |\xi_m|^p\}] = \sum_{m\in\mathbb{Z}^d} \mathbb{E}[\min\{1, |X|^p\}] = \infty$$

since $\mathbb{P}(\min\{1, |X|^p\} > 0) = \mathbb{P}(X \neq 0) > 0$. Hence $0 \neq \widetilde{\gamma}$ thus $\widetilde{\gamma} = \gamma < 0$. Next, we get

$$
\begin{aligned}
(5.2)\qquad \infty &> \sum_{m\in\mathbb{Z}^d} \mathbb{E}[\min\{1,(1+\|m\|_\infty)^{\gamma p}|\xi_m|^p\}] \\
&\geq \sum_{m\in\mathbb{Z}^d} \mathbb{E}[(1+\|m\|_\infty)^{\gamma p}|\xi_m|^p \mathbb{1}_{|\xi_m|^p\leq(1+\|m\|_\infty)^{-\gamma p}}] \\
&= \sum_{m\in\mathbb{Z}^d} \mathbb{E}[(1+\|m\|_\infty)^{\gamma p}|X|^p \mathbb{1}_{|X|^p\leq(1+\|m\|_\infty)^{-\gamma p}}] \\
&= \mathbb{E}\Big[|X|^p \sum_{m\in\mathbb{Z}^d} (1+\|m\|_\infty)^{\gamma p} \mathbb{1}_{|X|^p\leq(1+\|m\|_\infty)^{-\gamma p}}\Big] \\
&\geq \mathbb{E}\Big[|X|^p \mathbb{1}_{|X|^p\geq 1} \sum_{\substack{m\in\mathbb{Z}^d \\ \|m\|_\infty\geq|X|^{-1/\gamma}-1}} (1+\|m\|_\infty)^{\gamma p}\Big] \\
&\asymp \mathbb{E}[|X|^p \mathbb{1}_{|X|^p\geq 1}(|X|^{-1/\gamma})^{\gamma p+d}] = \mathbb{E}[|X|^{-d/\gamma}\mathbb{1}_{|X|^p\geq 1}].
\end{aligned}
$$

Since $\gamma < 0$, this implies $\mathbb{E}[|X|^{-d/\gamma}] < \infty$. The step $(5.2)_6$ is justified by the following computation, valid for all $N \in \mathbb{N}_0$ (cf. also the proof of Lemma A.4):

$$
\begin{aligned}
(5.3)\qquad \sum_{\substack{m\in\mathbb{Z}^d \\ \|m\|_\infty\geq N}} (1+\|m\|_\infty)^{\gamma p} &= \sum_{\ell\geq N} (1+\ell)^{\gamma p} \#\{m\in\mathbb{Z}^d : \|m\|_\infty = \ell\} \\
&= \sum_{\ell\geq N} (1+\ell)^{\gamma p}((2\ell+1)^d - (2\ell-1)^d) \\
&\asymp \sum_{\ell\geq N} (1+\ell)^{\gamma p+d-1} = \sum_{\ell\geq N} \int_\ell^{\ell+1} (1+\ell)^{\gamma p+d-1}\,\mathrm{d}x \\
&\asymp \sum_{\ell\geq N} \int_\ell^{\ell+1} (1+x)^{\gamma p+d-1}\,\mathrm{d}x = \int_N^\infty (1+x)^{\gamma p+d-1}\,\mathrm{d}x \\
&= \begin{cases} -\dfrac{(N+1)^{\gamma p+d}}{\gamma p+d} & \text{if } \gamma p+d<0, \\ \infty & \text{else.} \end{cases}
\end{aligned}
$$

Now, regarding $(5.2)_6$, this computation shows that if $\gamma p + d \geq 0$, then

$$
\infty > \mathbb{E}\Big[|X|^p \mathbb{1}_{|X|^p\geq 1} \sum_{\substack{m\in\mathbb{Z}^d \\ \|m\|_\infty\geq|X|^{-1/\gamma}-1}} (1+\|m\|_\infty)^{\gamma p}\Big] = \infty \cdot \mathbb{E}[|X|^p \mathbb{1}_{|X|^p\geq 1}],
$$

and hence $\mathbb{E}[|X|^p \mathbb{1}_{|X|^p\geq 1}] = 0$, which shows that $|X| < 1$ almost surely, and so $(5.2)_6$ holds. If otherwise $\gamma p + d < 0$, then the computation in (5.3) justifies $(5.2)_6$ directly. ∎

Remark 5.3. Let $d \in \mathbb{N}$, $s, \alpha, \beta, \gamma, \theta \in \mathbb{R}$, $p \in (0,\infty)$, and $q \in (0,\infty]$. Assume that $\mathbf{a}$ follows a Besov sequence prior with parameters $(\alpha,\beta,\gamma,\theta)$

and template random variable X satisfying $\mathbb{P}(X \neq 0) > 0$. If $\mathbf{a} \in b^s_{p,q}(\mathbb{R}^d)$ almost surely whenever Property A is satisfied, then Proposition 5.2 implies that $\mathbb{E}[|X|^{-d/\gamma}] < \infty$ whenever $\gamma p < -d$, which in turn yields $\mathbb{E}[|X|^\rho] < \infty$ for all $\rho \in (0, p)$. This shows that the conditions (3.3) and (5.1) are quite sharp if $p \neq \infty$.

Next, we address the sharpness of the conditions (3.3) and (5.1) when $p = \infty$.

PROPOSITION 5.4. *Let $d \in \mathbb{N}$, $s, \alpha, \beta, \gamma, \theta \in \mathbb{R}$, $p = \infty$, and $q \in (0, \infty]$. Assume that $\mathbf{a}$ follows a Besov sequence prior with parameters $(\alpha, \beta, \gamma, \theta)$ and template random variable X. Suppose that X is such that $\mathbf{a} \in b^s_{\infty,q}(\mathbb{R}^d)$ almost surely. If $\beta \geq 0$ or $\gamma \geq 0$, then there exists $R > 0$ such that $|X| \leq R$ almost surely.*

Proof. Assume that $\gamma \geq 0$ and that for all $R > 0$, we have $\mathbb{P}(|X| \geq R) > 0$. Then $\sum_{m \in \mathbb{Z}^d} \mathbb{P}(|\xi_m| \geq R) = \infty$. Thus, the Borel–Cantelli lemma shows for every $R \in \mathbb{N}$ that the event A_R on which $\sup_{m \in \mathbb{Z}^d} |\xi_m| \geq R$ has probability $\mathbb{P}(A_R) = 1$. Hence also $\mathbb{P}(A) = 1$ for $A := \bigcap_{R \in \mathbb{R}} A_R$. But on the event A we have $\sup_{m \in \mathbb{Z}^d} |\xi_m| = \infty$, from which $\|\mathbf{a}\|_{b^s_{\infty,q}(\mathbb{R}^d)} = \infty$ follows. Similar arguments apply when $\beta \geq 0$. ∎

The following remark shows that if p is infinite, then the essential boundedness of the template variable also cannot be simply dropped when $s + d/2 + \alpha = 0$.

REMARK 5.5. Let $d \in \mathbb{N}$, $s, \alpha, \beta, \gamma, \theta \in \mathbb{R}$, $p = \infty$, and $q \in (0, \infty]$. Let further $X \sim \mathcal{N}(0, 1)$ be a standard normal random variable, and assume that $\mathbf{a}$ follows a Besov sequence prior with parameters $(\alpha, \beta, \gamma, \theta)$ and template variable X. Suppose that $s + d/2 + \alpha = 0$ and $\theta \geq -1/2 - 1/q$. We will show that, almost surely,

$$\mathbf{a} \notin b^s_{\infty,q}(\mathbb{R}^d). \tag{5.4}$$

This shows that even for $\beta, \gamma \ll 0$, Property A (which in the present setting reduces to $\beta \leq 0$, $\gamma \leq 0$, and $\theta < -1/q$ (for $q < \infty$) or $\theta \leq 0$ (for $q = \infty$)) is in general *not* sufficient to ensure that $\mathbf{a} \in b^s_{p,q}(\mathbb{R}^d)$ almost surely for $p = \infty$ if one drops condition (3.3).

To prove (5.4), fix $t^* = (M, \dots, M)$. It follows from Lemma A.11 that there exists an $\mathbb{N}$-valued random variable N satisfying

$$\sup_{\substack{m \in \mathbb{Z}^d \\ \|m\|_\infty \leq 2^j}} |\xi_{j,t^*,m}| \geq \sqrt{\log((2^{j+1} + 1)^d)} \geq c\sqrt{j}$$

for all $j \geq N$ and some absolute constant $c > 0$. Hence, with

$$\eta_2 := \|(2^{j(s+d/2+\alpha)}(j+1)^\theta Z_{j,t})_{j \in \mathbb{N}, t \in T_j}\|_{\ell_q},$$

and $Z_{j,t}$ as in Lemma 3.1, we have

$$\begin{aligned}\|\mathbf{a}\|^q_{b^s_{\infty,q}(\mathbb{R}^d)} \gtrsim \eta_2^q &\geq \sum_{j\in\mathbb{N}} (j+1)^{\theta q} \sup_{\substack{m\in\mathbb{Z}^d\\ \|m\|_\infty\leq 2^j}} (1+\|m\|_\infty/2^j)^{\beta q} |\xi_{j,t^*,m}|^q \\ &\gtrsim \sum_{j\in\mathbb{N}} (j+1)^{\theta q} \Big(\sup_{\substack{m\in\mathbb{Z}^d\\ \|m\|_\infty\leq 2^j}} |\xi_{j,t^*,m}| \Big)^q \\ &\gtrsim \sum_{j\geq N} j^{(\theta+1/2)q} = \infty,\end{aligned}$$

because $s + d/2 + \alpha = 0$ and $\theta \geq -1/2 - 1/q$.

Similarly to Remark 5.3, we can discuss the sharpness of the condition (3.4) when $p \neq \infty$.

Remark 5.6. Let $d \in \mathbb{N}$, $s, \alpha, \beta, \gamma, \theta \in \mathbb{R}$, $p, r \in (0,\infty)$, and $q \in (0,\infty]$. Let further X be a random variable with $\mathbb{P}(X \neq 0) > 0$, and assume that $\mathbf{a}$ follows a Besov sequence prior with parameters $(\alpha,\beta,\gamma,\theta)$ and template random variable X. Suppose that $\mathbb{E}[\|\mathbf{a}\|^r_{b^s_{p,q}(\mathbb{R}^d)}] < \infty$ whenever Property A is satisfied. On the one hand, this implies that $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} < \infty$ almost surely whenever Property A is satisfied. As seen in Remark 5.3, this gives $\mathbb{E}[|X|^\rho] < \infty$ for all $\rho \in (0,p)$. On the other hand, $\|\mathbf{a}\|_{b^s_{p,q}(\mathbb{R}^d)} \gtrsim |\xi_0|$. Thus, if $\mathbb{E}[\|\mathbf{a}\|^r_{b^s_{p,q}(\mathbb{R}^d)}] < \infty$, then necessarily $\mathbb{E}[|\xi_0|^r] < \infty$. Overall, this shows that $\mathbb{E}[|X|^{\max\{r,\rho\}}] < \infty$ for all $\rho \in (0,p)$. Hence (3.4) is quite sharp if $p \neq \infty$.

Appendix A. Auxiliary statements. The following lemma shows the admissibility of the function $w_\sigma : \mathbb{R}^d \to \mathbb{R}$ given by $w_\sigma(x) := (1+\|x\|_2^2)^{\sigma/2}$ with the notion of admissibility introduced in Definition 2.1.

Lemma A.1. *Let $\sigma \in \mathbb{R}$ and $w_\sigma : \mathbb{R}^d \to \mathbb{R}$, $w_\sigma(x) := (1+\|x\|_2^2)^{\sigma/2}$.*

(a) *For every $\beta \in \mathbb{N}_0^d$, we have*

$$D^\beta (1+\|x\|_2^2)^{\sigma/2} = p_{\sigma,\beta}(x)(1+\|x\|_2^2)^{\sigma/2-\|\beta\|_1}$$

where $p_{\sigma,\beta}$ is a polynomial of degree at most $\|\beta\|_1$. In particular, we have $|D^\beta(1+\|x\|_2^2)^{\sigma/2}| \lesssim (1+\|x\|_2^2)^{\sigma/2}$.

(b) *We have $w_\sigma(x) \lesssim w_{|\sigma|}(x-y) w_\sigma(y)$ for all $x, y \in \mathbb{R}^d$.*

Proof. We show (a) by induction. If $\beta = 0$, then we can choose $p_{\sigma,\beta}(x) = 1$, which is indeed a polynomial of degree at most $\|\beta\|_1 = 0$. Now assume the claim holds true for β. We show that it holds true for $\beta + e_k$ in place of β, where $e_k \in \mathbb{R}^d$ denotes the vector whose entries are 0 except for the kth, which is 1. Note that

$$
\begin{aligned}
&D^{\beta+e_k}(1+\|x\|_2^2)^{\sigma/2}\\
&= D^{e_k}(p_{\sigma,\beta}(x)(1+\|x\|_2^2)^{\sigma/2-\|\beta\|_1})\\
&= (D^{e_k}p_{\sigma,\beta}(x))(1+\|x\|_2^2)^{\sigma/2-\|\beta\|_1} + p_{\sigma,\beta}(x)(D^{e_k}(1+\|x\|_2^2)^{\sigma/2-\|\beta\|_1})\\
&= (D^{e_k}p_{\sigma,\beta}(x))(1+\|x\|_2^2)^{\sigma/2-\|\beta\|_1} + p_{\sigma,\beta}(x)(\tfrac{\sigma}{2}-\|\beta\|_1)2x_k(1+\|x\|_2^2)^{\sigma/2-\|\beta\|_1-1}\\
&= (1+\|x\|_2^2)^{\sigma/2-\|\beta\|_1-1}((D^{e_k}p_{\sigma,\beta}(x))(1+\|x\|_2^2) + p_{\sigma,\beta}(x)(\tfrac{\sigma}{2}-\|\beta\|_1)2x_k),
\end{aligned}
$$

and $(D^{e_k}p_{\sigma,\beta}(x))(1+\|x\|_2^2) + p_{\sigma,\beta}(x)(\sigma/2-\|\beta\|_1)2x_k$ is a polynomial of degree at most $\|\beta+e_k\|_1 = \|\beta\|_1 + 1$.

For the proof of (b), note that the case $\sigma = 0$ is trivial. Further, we observe that

$$1+\|x\|_2 \le 1+\|x-y\|_2+\|y\|_2 \le (1+\|x-y\|_2)(1+\|y\|_2).$$

If $\sigma > 0$, it follows that

$$w_\sigma(x) \lesssim (1+\|x\|_2)^\sigma \le (1+\|x-y\|_2)^\sigma(1+\|y\|_2)^\sigma \lesssim w_\sigma(x-y)w_\sigma(y).$$

On the other hand, if $\sigma < 0$, then

$$
\begin{aligned}
w_\sigma(x) &\asymp (1+\|x\|_2)^\sigma = ((1+\|x\|_2)^{-1})^{-\sigma}\\
&\le ((1+\|x-y\|_2)(1+\|y\|_2)^{-1})^{-\sigma}\\
&= (1+\|x-y\|_2)^{-\sigma}(1+\|y\|_2)^\sigma\\
&\asymp w_{-\sigma}(x-y)(1+\|y\|_2)^\sigma \asymp w_{|\sigma|}(x-y)w_\sigma(y). \ \blacksquare
\end{aligned}
$$

The following lemma is taken from [15, Proposition 5.14]. We use it in the proofs of Theorems 4.1 and 4.4.

LEMMA A.2. *Let $(\xi_k)_{k\in\mathbb{N}}$ be a sequence of independent nonnegative (finite-valued) random variables. Then $\sum_{k\in\mathbb{N}}\xi_k < \infty$ almost surely if and only if*

$$\sum_{k\in\mathbb{N}} \mathbb{E}[\min\{\xi_k, 1\}] < \infty.$$

The following calculations are used in the proofs of Theorems 4.1 and 4.4.

LEMMA A.3.

(a) *Let X be a $\{0,1\}$-valued random variable and $\sigma\in(0,\infty)$. Then*

$$\mathbb{E}[\min\{1,\sigma X\}] = \mathbb{E}[X]\min\{1,\sigma\}.$$

(b) *Let ξ be a (real-valued) random variable, $R,p\in(0,\infty)$, $\sigma\in[0,\infty)$, and $X := \mathbb{1}_{1/R\le|\xi|\le R}$. Then $\min\{1,\sigma|\xi X|^p\} \ge R^{-p}X\min\{1,\sigma\}$ and in particular*

$$\mathbb{E}[\min\{1,\sigma|\xi|^p\}] \ge \mathbb{E}[\min\{1,\sigma|\xi X|^p\}] \ge R^{-p}\mathbb{E}[X]\min\{1,\sigma\}.$$

Proof. (a) In both cases $X = 0$ and $X = 1$, we have $\min\{1,\sigma X\} = X\min\{1,\sigma\}$. The linearity of the expectation yields the assertion.

(b) In both cases $X = 0$ and $X = 1$, we have

$$\min\{1, \sigma|\xi X|^p\} \geq R^{-p} X \min\{1, \sigma\}.$$

Here, we have used the fact that if $X = 1$, then necessarily $R \geq 1$. The linearity and the monotonicity of the expectation yield the assertion. ∎

LEMMA A.4. *Let $d, a \in \mathbb{N}$, $\beta \in \mathbb{R}$, and $p \in (0, \infty)$. We have*

$$\sum_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty / a)^{\beta p} < \infty$$

if and only if $\beta < -d/p$, in which case $\sum_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty/a)^{\beta p} \asymp a^d$, with the implied constants only depending on β, p, and d.

Proof. Note for $\ell \in \mathbb{N}$ that

$$\begin{aligned}
&\#\{m \in \mathbb{Z}^d : \|m\|_\infty = \ell\} \\
&\qquad = \#\{m \in \mathbb{Z}^d : \|m\|_\infty \leq \ell\} - \#\{m \in \mathbb{Z}^d : \|m\|_\infty \leq \ell - 1\} \\
&\qquad = (2\ell + 1)^d - (2(\ell - 1) + 1)^d = (2\ell + 1)^d - (2\ell - 1)^d.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty/a)^{\beta p} &= 1 + \sum_{\ell \in \mathbb{N}} \sum_{\substack{m \in \mathbb{Z}^d \\ \|m\|_\infty = \ell}} (1 + \|m\|_\infty/a)^{\beta p} \\
&= 1 + \sum_{\ell \in \mathbb{N}} ((2\ell + 1)^d - (2\ell - 1)^d)(1 + \ell/a)^{\beta p}.
\end{aligned}$$

Next, note that thanks to the mean value theorem, there exists a number $y = y_\ell \in (2\ell - 1, 2\ell + 1)$ such that $(2\ell + 1)^d - (2\ell - 1)^d = d \cdot y_\ell^{d-1} \cdot 2 \asymp \ell^{d-1}$. This entails the estimate

$$\begin{aligned}
&\sum_{\ell \in \mathbb{N}} ((2\ell + 1)^d - (2\ell - 1)^d) \left(1 + \frac{\ell}{a}\right)^{\beta p} \\
&\quad \asymp \sum_{\ell \in \mathbb{N}} \ell^{d-1} \left(1 + \frac{\ell}{a}\right)^{\beta p} = \sum_{\ell=1}^{a-1} \ell^{d-1} \left(1 + \frac{\ell}{a}\right)^{\beta p} + \sum_{\ell \geq a} \ell^{d-1} \left(1 + \frac{\ell}{a}\right)^{\beta p}.
\end{aligned}$$

In the first sum, we have $1 \leq 1 + \ell/a \leq 2$ and thus $\sum_{\ell=1}^{a-1} \ell^{d-1}(1 + \ell/a)^{\beta p} \asymp \sum_{\ell=1}^{a-1} \ell^{d-1}$. In the second sum, we have $\ell/a \leq 1 + \ell/a \leq 2\ell/a$ and thus

$$\sum_{\ell \geq a} \ell^{d-1}(1 + \ell/a)^{\beta p} \asymp \sum_{\ell \geq a} a^{-\beta p} \ell^{d-1+\beta p}.$$

Next, note for any $\nu \in \mathbb{R}$ and $\ell \in \mathbb{N}$ that if $\ell \leq x \leq \ell + 1$, then $\ell \leq x \leq 2\ell$. Hence $x \asymp \ell$ and also $x^\nu \asymp \ell^\nu$. It follows that

$$\sum_{\ell=1}^{a-1} \ell^\nu = \sum_{\ell=1}^{a-1} \int_\ell^{\ell+1} \ell^\nu \, \mathrm{d}x \asymp \sum_{\ell=1}^{a-1} \int_\ell^{\ell+1} x^\nu \, \mathrm{d}x = \int_1^a x^\nu \, \mathrm{d}x$$

and similarly

$$\sum_{\ell \geq a} \ell^\nu = \sum_{\ell \geq a} \int_\ell^{\ell+1} \ell^\nu \, \mathrm{d}x \asymp \sum_{\ell \geq a} \int_\ell^{\ell+1} x^\nu \, \mathrm{d}x = \int_a^\infty x^\nu \, \mathrm{d}x.$$

Summarizing, we see that

$$\begin{aligned} &\sum_{\ell=1}^{a-1} \ell^{d-1}(1+\ell/a)^{\beta p} + \sum_{\ell \geq a} \ell^{d-1}(1+\ell/a)^{\beta p} \\ &\quad \asymp \sum_{\ell=1}^{a-1} \ell^{d-1} + \sum_{\ell=a}^{\infty} a^{-\beta p} \ell^{d-1+\beta p} \asymp \int_1^a x^{d-1} \, \mathrm{d}x + a^{-\beta p} \int_a^\infty x^{d-1+\beta p} \, \mathrm{d}x, \end{aligned}$$

where the right-hand side is finite if and only if $\beta p < -d$. In this case, we get

$$\int_1^a x^{d-1} \, \mathrm{d}x + a^{-\beta p} \int_a^\infty x^{d-1+\beta p} \, \mathrm{d}x = \frac{a^d - 1}{d} - a^{-\beta p} \frac{a^{d+\beta p}}{d+\beta p} \asymp a^d,$$

so that $\sum_{m \in \mathbb{Z}^d} (1 + \|m\|_\infty / a)^{\beta p} \asymp 1 + a^d \asymp a^d$. ∎

A proof of the Paley–Zygmund inequality stated below can be found in [15, Lemma 5.1]. We use it in the proofs of Theorems 4.1 and 4.4.

Lemma A.5 (Paley–Zygmund inequality). *Let $X \geq 0$ be a random variable with $0 < \mathbb{E}[X^2] < \infty$, and let $\sigma \in (0,1]$. Then*

$$\mathbb{P}(X > \sigma \mathbb{E}[X]) \geq (1-\sigma)^2 \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}.$$

The following lemma provides bounds for the raw moments for the binomial distribution. We use it in the proof of Theorem 4.4. A related bound can be found in [17, (4.5)].

Lemma A.6. *Let $n \in \mathbb{N}$, $\rho \in (0,1)$, $\sigma \in (0,\infty)$, and let $X \sim \mathrm{Bin}(n,\rho)$ be a binomially distributed random variable. Then $\mathbb{E}[X^\sigma] \lesssim \max\{n\rho, (n\rho)^\sigma\}$, where the implied constant only depends on σ.*

Proof. If $\sigma \leq 1$, we trivially have $\mathbb{E}[X^\sigma] \leq \mathbb{E}[X] = n\rho$. Hence we may assume that $\sigma > 1$ in what follows. If $n\rho \leq 1$, let $c := \lceil \sigma \rceil \in \mathbb{N}$. With $S(c,k)$ and $n^{\underline{k}}$ denoting the Stirling numbers of the second kind and the falling factorials, respectively, we deduce from [17, (2.9) and (3.1)] and because of $S(c,0) = 0$ that

$$\mathbb{E}[X^\sigma] \leq \mathbb{E}[X^c] = \sum_{k=1}^c S(c,k) n^{\underline{k}} \rho^k \leq \sum_{k=1}^c S(c,k) (n\rho)^k \leq \sum_{k=1}^c S(c,k) n\rho \lesssim n\rho.$$

On the other hand, if $n\rho > 1$, let $c \in \mathbb{N}$ such that $c \leq \sigma < c+1$. Thus there exists $\theta \in [0,1]$ such that $\frac{1}{\sigma} = \frac{\theta}{c+1} + \frac{1-\theta}{c}$. Using [17, (2.9) and (3.1)], we

obtain

$$\mathbb{E}[X^c] = \sum_{k=1}^{c} S(c,k) n^{\underline{k}} \rho^k \le \sum_{k=1}^{c} S(c,k)(n\rho)^k \le \sum_{k=1}^{c} S(c,k)(n\rho)^c \lesssim (n\rho)^c.$$

Thus $\|X\|_{L_c} = \mathbb{E}[X^c]^{1/c} \lesssim n\rho$, and similarly $\|X\|_{L_{c+1}} = \mathbb{E}[X^{c+1}]^{\frac{1}{c+1}} \lesssim n\rho$.

Now [11, Proposition 6.10] gives

$$\mathbb{E}[X^\sigma] = \|X\|_{L_\sigma}^\sigma \le (\|X\|_{L_c}^{1-\theta} \|X\|_{L_{c+1}}^{\theta})^\sigma \lesssim (n\rho)^\sigma$$

and the proof is complete. ∎

The following lemma is the technical part of the proofs of Theorems 3.3 and 3.4. Its proof is loosely based on Etemadi's proof of the strong law of large numbers in [5, §12].

LEMMA A.7. *Let $(X_{j,\tau})_{j\in\mathbb{N}_0,\tau\in\mathbb{N}}$ be a family of nonnegative i.i.d. random variables, $d \in \mathbb{N}$, $\varepsilon \in (0,\infty)$, and $\sigma \in [1,\infty)$. Suppose that*

$$\mathbb{E}[X_{j,\tau}^{(1+\varepsilon)\sigma}] < \infty.$$

Let

$$M_j := \#\{m \in \mathbb{Z}^d : \|m\|_\infty \le 2^j\} = (2^{j+1}+1)^d,$$
$$N_j := \#\{m \in \mathbb{Z}^d : 2^j < \|m\|_\infty \le 2^{j+1}\} = M_{j+1} - M_j$$

for $j \in \mathbb{N}_0$. Then, for

$$\Xi := \sup_{j\in\mathbb{N}_0} \Big(\frac{1}{M_j} \sum_{\tau=1}^{M_j} X_{j,\tau} + \sup_{\ell\ge j} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} X_{j,\tau} \Big),$$

we have $\mathbb{E}[\Xi^\sigma] < \infty$.

Proof. Define

$$E := \mathbb{E}[X_{0,1}^\sigma],$$
$$\kappa_{j,\tau} := \begin{cases} 2^{jd/3} & \text{if } 1 \le \tau \le M_j, \\ 2^{\ell d/3} & \text{if } M_\ell + 1 \le \tau \le M_{\ell+1} \text{ for some (unique) } \ell \ge j, \end{cases}$$
$$E_{j,\tau} := \mathbb{E}[X_{j,\tau}^\sigma \mathbb{1}_{X_{j,\tau}^\sigma \le \kappa_{j,\tau}}],$$

and note $E_{j,\tau} \le E$. Since $\sigma \ge 1$, Jensen's inequality shows $\big(\frac{1}{N}\sum_{\tau=1}^N a_\tau\big)^\sigma \le \frac{1}{N}\sum_{\tau=1}^N a_\tau^\sigma$ for $N \in \mathbb{N}$ and $a_1, \dots, a_N \ge 0$. Then, from

$$\Xi^\sigma \le \Big(\sup_{j\in\mathbb{N}_0} \frac{1}{M_j} \sum_{\tau=1}^{M_j} X_{j,\tau} + \sup_{j\in\mathbb{N}_0} \sup_{\ell\ge j} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} X_{j,\tau} \Big)^\sigma$$
$$\lesssim \Big(\sup_{j\in\mathbb{N}_0} \frac{1}{M_j} \sum_{\tau=1}^{M_j} X_{j,\tau} \Big)^\sigma + \Big(\sup_{j\in\mathbb{N}_0} \sup_{\ell\ge j} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} X_{j,\tau} \Big)^\sigma$$

$$
\begin{aligned}
&= \sup_{j\in\mathbb{N}_0}\bigg(\frac{1}{M_j}\sum_{\tau=1}^{M_j} X_{j,\tau}\bigg)^{\sigma} + \sup_{j\in\mathbb{N}_0}\sup_{\ell\geq j}\bigg(\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}} X_{j,\tau}\bigg)^{\sigma}\\
&\leq \sup_{j\in\mathbb{N}_0}\frac{1}{M_j}\sum_{\tau=1}^{M_j} X_{j,\tau}^{\sigma} + \sup_{j\in\mathbb{N}_0}\sup_{\ell\geq j}\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}} X_{j,\tau}^{\sigma}\\
&\leq \sup_{j\in\mathbb{N}_0}\frac{1}{M_j}\sum_{\tau=1}^{M_j}(X_{j,\tau}^{\sigma}-E) + \sup_{j\in\mathbb{N}_0}\sup_{\ell\geq j}\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(X_{j,\tau}^{\sigma}-E) + 2E\\
&\leq \sup_{j\in\mathbb{N}_0}\frac{1}{M_j}\sum_{\tau=1}^{M_j}(X_{j,\tau}^{\sigma}-E_{j,\tau}) + \sup_{j\in\mathbb{N}_0}\sup_{\ell\geq j}\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(X_{j,\tau}^{\sigma}-E_{j,\tau}) + 2E\\
&\leq \sum_{j\in\mathbb{N}_0}\bigg|\frac{1}{M_j}\sum_{\tau=1}^{M_j}(X_{j,\tau}^{\sigma}-E_{j,\tau})\bigg| + \sum_{j\in\mathbb{N}_0}\sum_{\ell\geq j}\bigg|\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(X_{j,\tau}^{\sigma}-E_{j,\tau})\bigg| + 2E,
\end{aligned}
$$

we obtain

$$
\begin{aligned}
\mathbb{E}[\Xi^{\sigma}] \lesssim 2E &+ \sum_{j\in\mathbb{N}_0}\mathbb{E}\bigg[\bigg|\frac{1}{M_j}\sum_{\tau=1}^{M_j}(X_{j,\tau}^{\sigma}-E_{j,\tau})\bigg|\bigg]\\
&+ \sum_{j\in\mathbb{N}_0}\sum_{\ell\geq j}\mathbb{E}\bigg[\bigg|\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(X_{j,\tau}^{\sigma}-E_{j,\tau})\bigg|\bigg].
\end{aligned}
$$

In order to conclude that $\mathbb{E}[\Xi^{\sigma}] < \infty$, we show that the two infinite sums in the last expression are finite. Let $X_* := X_{0,1}$. First, we see that

$$
\begin{aligned}
&\mathbb{E}\bigg[\bigg|\frac{1}{M_j}\sum_{\tau=1}^{M_j}(X_{j,\tau}^{\sigma}-E_{j,\tau})\bigg|\bigg]\\
&= \mathbb{E}\bigg[\bigg|\frac{1}{M_j}\sum_{\tau=1}^{M_j}(X_{j,\tau}^{\sigma}\mathbb{1}_{X_{j,\tau}^{\sigma}\leq\kappa_{j,\tau}}-E_{j,\tau}) + \frac{1}{M_j}\sum_{\tau=1}^{M_j}X_{j,\tau}^{\sigma}\mathbb{1}_{X_{j,\tau}^{\sigma}>\kappa_{j,\tau}}\bigg|\bigg]\\
&\leq \mathbb{E}\bigg[\bigg|\frac{1}{M_j}\sum_{\tau=1}^{M_j}(X_{j,\tau}^{\sigma}\mathbb{1}_{X_{j,\tau}^{\sigma}\leq\kappa_{j,\tau}}-E_{j,\tau})\bigg|\bigg] + \frac{1}{M_j}\sum_{\tau=1}^{M_j}\mathbb{E}[|X_{j,\tau}^{\sigma}\mathbb{1}_{X_{j,\tau}^{\sigma}>\kappa_{j,\tau}}|]\\
&\leq \sqrt{\mathbb{E}\bigg[\bigg(\frac{1}{M_j}\sum_{\tau=1}^{M_j}(X_{j,\tau}^{\sigma}\mathbb{1}_{X_{j,\tau}^{\sigma}\leq\kappa_{j,\tau}}-E_{j,\tau})\bigg)^2\bigg]} + \frac{1}{M_j}\sum_{\tau=1}^{M_j}\kappa_{j,\tau}^{-\varepsilon}\mathbb{E}[|X_{j,\tau}^{\sigma}|^{1+\varepsilon}]\\
&= \sqrt{\frac{1}{M_j^2}\sum_{\tau=1}^{M_j}\mathbb{V}[X_{j,\tau}^{\sigma}\mathbb{1}_{X_{j,\tau}^{\sigma}\leq\kappa_{j,\tau}}]} + 2^{-jd\varepsilon/3}\mathbb{E}[|X_*^{\sigma}|^{1+\varepsilon}]
\end{aligned}
$$

$$\leq \sqrt{\frac{1}{M_j^2}\sum_{\tau=1}^{M_j}\kappa_{j,\tau}^2} + 2^{-jd\varepsilon/3}\mathbb{E}[|X_*^\sigma|^{1+\varepsilon}]$$
$$\leq \sqrt{\frac{2^{2jd/3}}{M_j}} + 2^{-jd\varepsilon/3}\mathbb{E}[|X_*^\sigma|^{1+\varepsilon}] \lesssim 2^{-jd/6} + 2^{-jd\varepsilon/3}.$$

Here we have used the fact that $\mathbb{E}[|Z|] \leq \sqrt{\mathbb{E}[Z^2]}$ by Jensen's inequality and that if $Z_1, \dots, Z_n$ are independent random variables, then

$$\mathbb{E}\Big[\Big(\sum_{j=1}^n (Z_j - \mathbb{E}[Z_j])\Big)^2\Big] = \mathbb{V}\Big[\sum_{j=1}^n Z_j\Big] = \sum_{j=1}^n V[Z_j],$$

and furthermore that $\mathbb{V}[Z] = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2 \leq \mathbb{E}[Z^2]$. Similarly, we find that

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(X_{j,\tau}^\sigma - E_{j,\tau})\Big|\Big] \\
&\quad= \mathbb{E}\Big[\Big|\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(X_{j,\tau}^\sigma \mathbb{1}_{X_{j,\tau}^\sigma \leq \kappa_{j,\tau}} - E_{j,\tau}) + \frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}} X_{j,\tau}^\sigma \mathbb{1}_{X_{j,\tau}^\sigma > \kappa_{j,\tau}}\Big|\Big] \\
&\quad\leq \mathbb{E}\Big[\Big|\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(X_{j,\tau}^\sigma \mathbb{1}_{X_{j,\tau}^\sigma \leq \kappa_{j,\tau}} - E_{j,\tau})\Big|\Big] + \frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}\mathbb{E}[|X_{j,\tau}^\sigma \mathbb{1}_{X_{j,\tau}^\sigma > \kappa_{j,\tau}}|] \\
&\quad\leq \sqrt{\mathbb{E}\Big[\Big(\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(X_{j,\tau}^\sigma \mathbb{1}_{X_{j,\tau}^\sigma \leq \kappa_{j,\tau}} - E_{j,\tau})\Big)^2\Big]} + \frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}\kappa_{j,\tau}^{-\varepsilon}\,\mathbb{E}[|X_{j,\tau}^\sigma|^{1+\varepsilon}] \\
&\quad\asymp \sqrt{\frac{1}{N_\ell^2}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}\mathbb{V}[X_{j,\tau}^\sigma \mathbb{1}_{X_{j,\tau}^\sigma \leq \kappa_{j,\tau}}]} + 2^{-\ell d\varepsilon/3}\mathbb{E}[|X_*^\sigma|^{1+\varepsilon}] \\
&\quad\leq \sqrt{\frac{1}{N_\ell^2}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}\kappa_{j,\tau}^2} + 2^{-\ell d\varepsilon/3}\mathbb{E}[|X_*^\sigma|^{1+\varepsilon}] \\
&\quad\leq \sqrt{\frac{2^{2\ell d/3}}{N_\ell}} + 2^{-\ell d\varepsilon/3}\mathbb{E}[|X_*^\sigma|^{1+\varepsilon}] \lesssim 2^{-\ell d/6} + 2^{-\ell d\varepsilon/3}.
\end{aligned}$$

Here, we have used $M_j \asymp N_j \asymp 2^{jd}$ (see the proof of Lemma 3.1).

It follows that

$$\sum_{j\in\mathbb{N}_0}\mathbb{E}\Big[\Big|\frac{1}{M_j}\sum_{\tau=1}^{M_j}(X_{j,\tau}^\sigma - E_{j,\tau})\Big|\Big] \lesssim \sum_{j\in\mathbb{N}_0}(2^{-jd/6} + 2^{-jd\varepsilon/3}) < \infty$$

and

$$\sum_{j\in\mathbb{N}_0}\sum_{\ell\geq j}\mathbb{E}\Big[\Big|\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(X_{j,\tau}^\sigma - E_{j,\tau})\Big|\Big] \lesssim \sum_{j\in\mathbb{N}_0}\sum_{\ell\geq j}(2^{-\ell d/6}+2^{-\ell d\varepsilon/3}) \lesssim \sum_{j\in\mathbb{N}_0}(2^{-jd/6}+2^{-jd\varepsilon/3}) < \infty.$$

This completes the proof. ∎

The following lemma is the technical part of the proof of Theorem 3.5.

LEMMA A.8. *Let $(X_{j,\tau})_{j\in\mathbb{N}_0,\tau\in\mathbb{N}}$ be a family of nonnegative i.i.d. random variables, $d\in\mathbb{N}$, $C,\varepsilon\in(0,\infty)$, and $\sigma\in[1,\infty)$. Suppose that*

$$\mathbb{E}[\exp(CX_{j,\tau}^\sigma)] < \infty.$$

Let

$$M_j := \#\{m\in\mathbb{Z}^d : \|m\|_\infty \leq 2^j\} = (2^{j+1}+1)^d,$$
$$N_j := \#\{m\in\mathbb{Z}^d : 2^j < \|m\|_\infty \leq 2^{j+1}\} = M_{j+1}-M_j$$

for $j\in\mathbb{N}_0$. Then, for

$$\Xi := \sup_{j\in\mathbb{N}_0}\Big(\frac{1}{M_j}\sum_{\tau=1}^{M_j}X_{j,\tau} + \sup_{\ell\geq j}\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}X_{j,\tau}\Big),$$

we have $\mathbb{E}\big[\exp\big(\frac{C}{2^{\sigma+1}(1+\varepsilon)}\Xi^\sigma\big)\big] < \infty$.

Proof. First note that $(a+b)^\sigma \leq (2\max\{a,b\})^\sigma \leq 2^\sigma(a^\sigma+b^\sigma)$ and

$$\exp(a+b) = \exp(a)\exp(b) \leq \frac{\exp(a)^2}{2}+\frac{\exp(b)^2}{2} \leq \exp(2a)+\exp(2b)$$

for all $a,b\in\mathbb{R}$. Together with Jensen's inequality, we deduce for

$$Y_j := \frac{1}{M_j}\sum_{\tau=1}^{M_j}X_{j,\tau} \quad\text{and}\quad Y_{j,\ell} := \frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}X_{j,\tau}$$

that

$$\begin{aligned}
&\exp\Big(\frac{C}{2^{\sigma+1}(1+\varepsilon)}\Xi^\sigma\Big)\\
&\quad\leq \exp\Big(\frac{C}{2^{\sigma+1}(1+\varepsilon)}\Big(\sup_{j\in\mathbb{N}_0}Y_j + \sup_{j\in\mathbb{N}_0}\sup_{\ell\geq j}Y_{j,\ell}\Big)^\sigma\Big)\\
&\quad\leq \exp\Big(\frac{C}{2^{\sigma+1}(1+\varepsilon)}2^\sigma\Big(\sup_{j\in\mathbb{N}_0}Y_j\Big)^\sigma + \frac{C}{2^{\sigma+1}(1+\varepsilon)}2^\sigma\Big(\sup_{j\in\mathbb{N}_0}\sup_{\ell\geq j}Y_{j,\ell}\Big)^\sigma\Big)\\
&\quad\leq \exp\Big(\frac{C}{1+\varepsilon}\sup_{j\in\mathbb{N}_0}Y_j^\sigma\Big) + \exp\Big(\frac{C}{1+\varepsilon}\sup_{j\in\mathbb{N}_0}\sup_{\ell\geq j}Y_{j,\ell}^\sigma\Big)
\end{aligned}$$

$$= \sup_{j\in\mathbb{N}_0} \exp\Big(\frac{C}{1+\varepsilon} Y_j^\sigma\Big) + \sup_{j\in\mathbb{N}_0} \sup_{\ell\geq j} \exp\Big(\frac{C}{1+\varepsilon} Y_{j,\ell}^\sigma\Big)$$
$$\leq \sup_{j\in\mathbb{N}_0} \frac{1}{M_j} \sum_{\tau=1}^{M_j} \exp\Big(\frac{C}{1+\varepsilon} X_{j,\tau}^\sigma\Big) + \sup_{j\in\mathbb{N}_0} \sup_{\ell\geq j} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} \exp\Big(\frac{C}{1+\varepsilon} X_{j,\tau}^\sigma\Big).$$

Let $\tilde{X}_{j,\tau} := \exp(\frac{C}{1+\varepsilon} X_{j,\tau}^\sigma)$. Then $\mathbb{E}[\tilde{X}_{j,\tau}^{1+\varepsilon}] = \mathbb{E}[\exp(CX_{j,\tau}^\sigma)] < \infty$, and the monotonicity of the expectation yields

$$\mathbb{E}\Big[\exp\Big(\frac{C}{2^{\sigma+1}(1+\varepsilon)} \Xi^\sigma\Big)\Big]$$
$$\leq \mathbb{E}\Big[\sup_{j\in\mathbb{N}_0} \frac{1}{M_j} \sum_{\tau=1}^{M_j} \tilde{X}_{j,\tau} + \sup_{j\in\mathbb{N}_0} \sup_{\ell\geq j} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} \tilde{X}_{j,\tau}\Big].$$

Now Lemma A.7 applied to $(\tilde{X}_{j,\tau})_{j\in\mathbb{N}_0,\tau\in\mathbb{N}}$ and $\sigma = 1$ shows that the right-hand side is finite. ∎

The following lemma is the technical part of the proof of Theorem 3.8.

LEMMA A.9. *Let $(X_{j,\tau})_{j\in\mathbb{N}_0,\tau\in\mathbb{N}}$ be a family of nonnegative i.i.d. random variables, $\varepsilon, \delta \in (0,\infty)$, $\mu, \nu \in (-\infty, 0]$, and $\sigma \in [1,\infty)$. Fix $d \in \mathbb{N}$ and let*

$$M_j := \#\{m \in \mathbb{Z}^d : \|m\|_\infty \leq 2^j\} = (2^{j+1}+1)^d,$$
$$N_j := \#\{m \in \mathbb{Z}^d : 2^j < \|m\|_\infty \leq 2^{j+1}\} = M_{j+1} - M_j$$

for $j \in \mathbb{N}_0$. Let further $(\lambda_{j,\tau})_{j\in\mathbb{N}_0,\tau\in\mathbb{N}}$ be a family of independent random variables, independent of $(X_{j,\tau})_{j\in\mathbb{N}_0,\tau\in\mathbb{N}}$ and such that $\lambda_{j,\tau}$ is Bernoulli distributed with success probability

$$\varrho_{j,\tau} := 2^{j\mu}(1 + \|\phi(\tau)\|_\infty / 2^j)^\nu,$$

where $\phi : \mathbb{N} \to \mathbb{Z}^d$ is a bijection such that $\tau \mapsto \|\phi(\tau)\|_\infty$ is nondecreasing. If

$$\mathbb{E}[X_{j,\tau}^{(1+\varepsilon)\sigma}] < \infty \quad \text{and} \quad \delta + \delta\varepsilon - \varepsilon \geq 0, \tag{A.1}$$

then

$$\tilde{\Xi} := \sup_{j\in\mathbb{N}_0} \Big(\frac{1}{M_j} \sum_{\tau=1}^{M_j} \varrho_{j,\tau}^{\frac{\delta-1}{\sigma}} \lambda_{j,\tau} X_{j,\tau} + \sup_{\ell\geq j} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} \varrho_{j,\tau}^{\frac{\delta-1}{\sigma}} \lambda_{j,\tau} X_{j,\tau} \Big)$$

satisfies $\mathbb{E}[\tilde{\Xi}^\sigma] < \infty$.

Proof. Let $E := \mathbb{E}[X_{0,1}^\sigma]$ and, for $j \in \mathbb{N}_0$ and $\tau \in \mathbb{N}$, let

$$Y_{j,\tau} := \varrho_{j,\tau}^{\frac{\delta-1}{\sigma}} \lambda_{j,\tau} X_{j,\tau},$$
$$\kappa_{j,\tau} := \begin{cases} 2^{jd/3} & \text{if } 1 \leq \tau \leq M_j, \\ 2^{\ell d/3} & \text{if } M_\ell + 1 \leq \tau \leq M_{\ell+1} \text{ for some (unique) } \ell \geq j, \end{cases}$$
$$E_{j,\tau} := \mathbb{E}[Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma \leq \kappa_{j,\tau}}].$$

Let $j \in \mathbb{N}_0$. If $1 \leq \tau \leq M_j$, then $0 \leq \|\phi(\tau)\|_\infty \leq 2^j$, i.e., $1 \leq 1 + \|\phi(\tau)\|_\infty / 2^j \leq 2$. It follows that $\varrho_{j,\tau} \asymp 2^{j\mu}$, and

$$\begin{aligned}\mathbb{E}[Y_{j,\tau}^{(1+\varepsilon)\sigma}] &= \varrho_{j,\tau}^{(1+\varepsilon)(\delta-1)} \mathbb{E}[\lambda_{j,\tau}] \mathbb{E}[X_{j,\tau}^{(1+\varepsilon)\sigma}] \\ &\lesssim \varrho_{j,\tau}^{(1+\varepsilon)(\delta-1)+1} \\ &= \varrho_{j,\tau}^{\delta+\delta\varepsilon-\varepsilon} \asymp 2^{j\mu(\delta+\delta\varepsilon-\varepsilon)} \leq 1\end{aligned}$$

since $\mu \leq 0$ and $\delta + \delta\varepsilon - \varepsilon \geq 0$.

If, on the other hand, $\tau > M_j$, then there exists a unique natural number $\ell \geq j$ such that $M_\ell + 1 \leq \tau \leq M_{\ell+1}$. Then $2^\ell < \|\phi(\tau)\|_\infty \leq 2^{\ell+1}$ and

$$2^{\ell-j} \leq 1 + 2^{\ell-j} \leq 1 + \frac{\|\phi(\tau)\|_\infty}{2^j} \leq 1 + 2^{\ell+1-j} \leq 3 \cdot 2^{\ell-j},$$

i.e., $\varrho_{j,\tau} \asymp 2^{j\mu} \cdot 2^{(\ell-j)\nu}$. It follows as above that

$$\mathbb{E}[Y_{j,\tau}^{(1+\varepsilon)\sigma}] \lesssim \varrho_{j,\tau}^{\delta+\delta\varepsilon-\varepsilon} \asymp (2^{j\mu} \cdot 2^{(\ell-j)\nu})^{\delta+\delta\varepsilon-\varepsilon} \leq 1.$$

Further note, because of $\varrho_{j,\tau} \leq 1$ and $\delta \geq 0$, that

$$E \geq \varrho_{j,\tau}^{\delta} \mathbb{E}[X_{j,\tau}^{\sigma}] = \varrho_{j,\tau}^{\delta-1} \mathbb{E}[\lambda_{j,\tau}] \mathbb{E}[X_{j,\tau}^{\sigma}] = \mathbb{E}[Y_{j,\tau}^{\sigma}] \geq E_{j,\tau}.$$

Note that $\sigma \geq 1$, so Jensen's inequality implies $(\frac{1}{N} \sum_{\tau=1}^N a_\tau)^\sigma \leq \frac{1}{N} \sum_{\tau=1}^N a_\tau^\sigma$ for $N \in \mathbb{N}$ and $a_1, \ldots, a_N \geq 0$. Hence,

$$\begin{aligned}\tilde{\Xi}^\sigma &\leq \left(\sup_{j \in \mathbb{N}_0} \frac{1}{M_j} \sum_{\tau=1}^{M_j} Y_{j,\tau} + \sup_{j \in \mathbb{N}_0} \sup_{\ell \geq j} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} Y_{j,\tau} \right)^\sigma \\ &\lesssim \left(\sup_{j \in \mathbb{N}_0} \frac{1}{M_j} \sum_{\tau=1}^{M_j} Y_{j,\tau} \right)^\sigma + \left(\sup_{j \in \mathbb{N}_0} \sup_{\ell \geq j} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} Y_{j,\tau} \right)^\sigma \\ &= \sup_{j \in \mathbb{N}_0} \left(\frac{1}{M_j} \sum_{\tau=1}^{M_j} Y_{j,\tau} \right)^\sigma + \sup_{j \in \mathbb{N}_0} \sup_{\ell \geq j} \left(\frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} Y_{j,\tau} \right)^\sigma \\ &\leq \sup_{j \in \mathbb{N}_0} \frac{1}{M_j} \sum_{\tau=1}^{M_j} Y_{j,\tau}^\sigma + \sup_{j \in \mathbb{N}_0} \sup_{\ell \geq j} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} Y_{j,\tau}^\sigma \\ &\leq \sup_{j \in \mathbb{N}_0} \frac{1}{M_j} \sum_{\tau=1}^{M_j} (Y_{j,\tau}^\sigma - E) + \sup_{j \in \mathbb{N}_0} \sup_{\ell \geq j} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} (Y_{j,\tau}^\sigma - E) + 2E \\ &\leq \sup_{j \in \mathbb{N}_0} \frac{1}{M_j} \sum_{\tau=1}^{M_j} (Y_{j,\tau}^\sigma - E_{j,\tau}) + \sup_{j \in \mathbb{N}_0} \sup_{\ell \geq j} \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} (Y_{j,\tau}^\sigma - E_{j,\tau}) + 2E \\ &\leq \sum_{j \in \mathbb{N}_0} \left| \frac{1}{M_j} \sum_{\tau=1}^{M_j} (Y_{j,\tau}^\sigma - E_{j,\tau}) \right| + \sum_{j \in \mathbb{N}_0} \sum_{\ell \geq j} \left| \frac{1}{N_\ell} \sum_{\tau=M_\ell+1}^{M_{\ell+1}} (Y_{j,\tau}^\sigma - E_{j,\tau}) \right| + 2E,\end{aligned}$$

from which we obtain

$$\mathbb{E}[\tilde{\Xi}^\sigma] \lesssim 2E + \sum_{j\in\mathbb{N}_0} \mathbb{E}\Big[\Big|\frac{1}{M_j}\sum_{\tau=1}^{M_j}(Y_{j,\tau}^\sigma - E_{j,\tau})\Big|\Big] + \sum_{j\in\mathbb{N}_0}\sum_{\ell\geq j}\mathbb{E}\Big[\Big|\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(Y_{j,\tau}^\sigma - E_{j,\tau})\Big|\Big].$$

In order to show that $\mathbb{E}[\tilde{\Xi}^\sigma] < \infty$, it suffices to show that the two infinite series on the right-hand side above are finite. To this end, recall from the start of the proof that $\mathbb{E}[Y_{j,\tau}^{(1+\varepsilon)\sigma}] \lesssim 1$. Hence,

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\frac{1}{M_j}\sum_{\tau=1}^{M_j}(Y_{j,\tau}^\sigma - E_{j,\tau})\Big|\Big] \\
&\quad= \mathbb{E}\Big[\Big|\frac{1}{M_j}\sum_{\tau=1}^{M_j}(Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma\leq\kappa_{j,\tau}} - E_{j,\tau}) + \frac{1}{M_j}\sum_{\tau=1}^{M_j} Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma>\kappa_{j,\tau}}\Big|\Big] \\
&\quad\leq \mathbb{E}\Big[\Big|\frac{1}{M_j}\sum_{\tau=1}^{M_j}(Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma\leq\kappa_{j,\tau}} - E_{j,\tau})\Big|\Big] + \frac{1}{M_j}\sum_{\tau=1}^{M_j}\mathbb{E}[|Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma>\kappa_{j,\tau}}|] \\
&\quad\leq \sqrt{\mathbb{E}\Big[\Big(\frac{1}{M_j}\sum_{\tau=1}^{M_j}(Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma\leq\kappa_{j,\tau}} - E_{j,\tau})\Big)^2\Big]} + \frac{1}{M_j}\sum_{\tau=1}^{M_j}\kappa_{j,\tau}^{-\varepsilon}\mathbb{E}[|Y_{j,\tau}^\sigma|^{1+\varepsilon}] \\
&\quad\lesssim \sqrt{\frac{1}{M_j^2}\sum_{\tau=1}^{M_j}\mathbb{V}[Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma\leq\kappa_{j,\tau}}]} + 2^{-jd\varepsilon/3} \\
&\quad\leq \sqrt{\frac{2^{2jd/3}}{M_j}} + 2^{-jd\varepsilon/3} \lesssim 2^{-jd/6} + 2^{-jd\varepsilon/3}.
\end{aligned}$$

Similarly, we find that

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(Y_{j,\tau}^\sigma - E_{j,\tau})\Big|\Big] \\
&= \mathbb{E}\Big[\Big|\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma\leq\kappa_{j,\tau}} - E_{j,\tau}) + \frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}} Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma>\kappa_{j,\tau}}\Big|\Big] \\
&\leq \mathbb{E}\Big[\Big|\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma\leq\kappa_{j,\tau}} - E_{j,\tau})\Big|\Big] + \frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}\mathbb{E}[|Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma>\kappa_{j,\tau}}|] \\
&\leq \sqrt{\mathbb{E}\Big[\Big(\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma\leq\kappa_{j,\tau}} - E_{j,\tau})\Big)^2\Big]} + \frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}\kappa_{j,\tau}^{-\varepsilon}\mathbb{E}[|Y_{j,\tau}^\sigma|^{1+\varepsilon}]
\end{aligned}$$

$$\lesssim \sqrt{\frac{1}{N_\ell^2}\sum_{\tau=M_\ell+1}^{M_{\ell+1}} \mathbb{V}[Y_{j,\tau}^\sigma \mathbb{1}_{Y_{j,\tau}^\sigma \le \kappa_{j,\tau}}]} + 2^{-\ell d\varepsilon/3} \le \sqrt{\frac{2^{2\ell d/3}}{N_\ell}} + 2^{-\ell d\varepsilon/3}$$
$$\lesssim 2^{-\ell d/6} + 2^{-\ell d\varepsilon/3}.$$

Here, we used the fact that $M_\ell \asymp N_\ell \asymp 2^{\ell d}$; see the proof of Lemma 3.1.

It follows that

$$\sum_{j\in\mathbb{N}_0} \mathbb{E}\Big[\Big|\frac{1}{M_j}\sum_{\tau=1}^{M_j}(Y_{j,\tau}^\sigma - E_{j,\tau})\Big|\Big] \lesssim \sum_{j\in\mathbb{N}_0}(2^{-jd/6} + 2^{-jd\varepsilon/3}) < \infty$$

and

$$\sum_{j\in\mathbb{N}_0}\sum_{\ell\ge j}\mathbb{E}\Big[\Big|\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}(Y_{j,\tau}^\sigma - E_{j,\tau})\Big|\Big] \lesssim \sum_{j\in\mathbb{N}_0}\sum_{\ell\ge j}(2^{-\ell d/6} + 2^{-\ell d\varepsilon/3})$$
$$\lesssim \sum_{j\in\mathbb{N}_0}(2^{-jd/6} + 2^{-jd\varepsilon/3}) < \infty. \ \blacksquare$$

The following lemma is the technical part of the proof of Theorem 5.1.

LEMMA A.10. *Let $(X_{j,\tau})_{j\in\mathbb{N}_0,\tau\in\mathbb{N}}$ be a family of nonnegative i.i.d. random variables. Suppose that*

$$\mathbb{E}[X_{j,\tau}\cdot\log_2^+(X_{j,\tau})] < \infty. \tag{A.2}$$

Fix $d\in\mathbb{N}$ and let

$$M_j := \#\{m\in\mathbb{Z}^d : \|m\|_\infty \le 2^j\} = (2^{j+1}+1)^d,$$
$$N_j := \#\{m\in\mathbb{Z}^d : 2^j < \|m\|_\infty \le 2^{j+1}\} = M_{j+1} - M_j$$

for $j\in\mathbb{N}_0$. Then the (random) quantity

$$\Xi := \sup_{j\in\mathbb{N}_0}\Big(\frac{1}{M_j}\sum_{\tau=1}^{M_j} X_{j,\tau} + \sup_{\ell\ge j}\frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}} X_{j,\tau}\Big)$$

is almost surely finite.

Proof. STEP 1. We prove the desired properties for the truncations

$$Y_{j,\tau} = \begin{cases} X_{j,\tau}\mathbb{1}_{X_{j,\tau}\le M_j} & \text{if } 1\le\tau\le M_j,\\ X_{j,\tau}\mathbb{1}_{X_{j,\tau}\le M_\ell} & \text{if } M_\ell+1\le\tau\le M_{\ell+1} \text{ for some (unique) } \ell\ge j\end{cases}$$

in place of $X_{j,\tau}$, i.e., we show for $E := \mathbb{E}[X_{j,\tau}]$ and arbitrary $C>0$ that

$$\sum_{j\in\mathbb{N}_0}\mathbb{P}(S_j^* \ge C+E) < \infty \quad\text{and}\quad \sum_{j\in\mathbb{N}_0}\sum_{\ell\ge j}\mathbb{P}(S_{j,\ell}^* \ge C+E) < \infty,$$

where

$$S_j^* := \frac{1}{M_j}\sum_{\tau=1}^{M_j} Y_{j,\tau} \quad\text{and}\quad S_{j,\ell}^* := \frac{1}{N_\ell}\sum_{\tau=M_\ell+1}^{M_{\ell+1}} Y_{j,\tau}.$$

This then implies by the Borel–Cantelli lemma that

$$\Xi^* := \sup_{j\in\mathbb{N}_0}\Big(S_j^* + \sup_{\ell\geq j} S_{j,\ell}^*\Big) < \infty$$

almost surely.

Let $X_* := X_{0,1}$ and recall that $X_{j,\tau} \overset{\text{i.i.d.}}{\sim} X_*$. Using $E \geq \mathbb{E}[Y_{j,\tau}]$, Chebyshev's inequality, and the independence of the $Y_{j,\tau}$, we obtain

$$\begin{aligned}
&\mathbb{P}(S_j^* \geq C + E)\\
&= \mathbb{P}\Big(\frac{1}{M_j}\sum_{\tau=1}^{M_j}(Y_{j,\tau} - E) \geq C\Big) \leq \mathbb{P}\Big(\frac{1}{M_j}\sum_{\tau=1}^{M_j}(Y_{j,\tau} - \mathbb{E}[Y_{j,\tau}]) \geq C\Big)\\
&\leq \frac{1}{C^2}\mathbb{V}\Big[\frac{1}{M_j}\sum_{\tau=1}^{M_j}(Y_{j,\tau} - \mathbb{E}[Y_{j,\tau}])\Big] = \frac{1}{C^2M_j^2}\sum_{\tau=1}^{M_j}\mathbb{V}[Y_{j,\tau} - \mathbb{E}[Y_{j,\tau}]]\\
&= \frac{1}{C^2M_j^2}\sum_{\tau=1}^{M_j}\mathbb{V}[Y_{j,\tau}] = \frac{1}{C^2M_j^2}\sum_{\tau=1}^{M_j}(\mathbb{E}[Y_{j,\tau}^2] - \mathbb{E}[Y_{j,\tau}]^2)\\
&\leq \frac{1}{C^2M_j^2}\sum_{\tau=1}^{M_j}\mathbb{E}[Y_{j,\tau}^2] = \frac{1}{C^2M_j^2}\sum_{\tau=1}^{M_j}\mathbb{E}[X_{j,\tau}^2\mathbb{1}_{X_{j,\tau}\leq M_j}]\\
&= \frac{1}{C^2M_j^2}\sum_{\tau=1}^{M_j}\mathbb{E}[X_*^2\mathbb{1}_{X_*\leq M_j}] = \frac{1}{C^2M_j}\mathbb{E}[X_*^2\mathbb{1}_{X_*\leq M_j}].
\end{aligned}$$

Thus, by the monotone convergence theorem,

$$\begin{aligned}
\sum_{j\in\mathbb{N}_0}\mathbb{P}(S_j^* \geq C + E) &\leq \frac{1}{C^2}\sum_{j\in\mathbb{N}_0}\frac{1}{M_j}\mathbb{E}[X_*^2\mathbb{1}_{X_*\leq M_j}]\\
&= \frac{1}{C^2}\mathbb{E}\Big[X_*^2\sum_{j\in\mathbb{N}_0}\frac{1}{M_j}\mathbb{1}_{X_*\leq M_j}\Big]\\
&= \frac{1}{C^2}\mathbb{E}\Big[X_*^2\mathbb{1}_{X_*>0}\sum_{j\in\mathbb{N}_0}\frac{1}{M_j}\mathbb{1}_{X_*\leq M_j}\Big]\\
&\leq \frac{1}{C^2}\mathbb{E}\Big[X_*^2\mathbb{1}_{X_*>0}\sum_{j\in\mathbb{N}_0}2^{-jd}\mathbb{1}_{X_*\leq 2^{(j+2)d}}\Big]\\
&\leq \frac{1}{C^2}\mathbb{E}\Big[X_*^2\mathbb{1}_{X_*>0}\sum_{j\geq\lceil\frac{\log_2(X_*)}{d}-2\rceil}2^{-jd}\Big].
\end{aligned} \tag{A.3}$$

Note that we are allowed to insert the indicator function $\mathbb{1}_{X_*>0}$ in $(\text{A.3})_3$ because by assumption $X_{j,\tau}$ are nonnegative random variables and because the factor X_*^2 vanishes if $X_* = 0$. In $(\text{A.3})_4$, we have used $2^{jd} \leq M_j \leq 2^{(j+2)d}$, whence $\frac{1}{M_j}\mathbb{1}_{X_*\leq M_j} \leq 2^{-jd}\mathbb{1}_{X_*\leq 2^{(j+2)d}}$. For $(\text{A.3})_5$, note that $X_* \leq 2^{(j+2)d}$ is equivalent to $j \geq \lceil\frac{\log_2(X_*)}{d} - 2\rceil$.

Next, perform the index shift $\ell = j + 2$ to see that

$$\sum_{j \geq \lceil \frac{\log_2(X_*)}{d} - 2 \rceil} 2^{-jd} = \sum_{\ell \geq \lceil \frac{\log_2(X_*)}{d} \rceil} 2^{-(\ell-2)d} \lesssim \sum_{\ell \geq \lceil \frac{\log_2(X_*)}{d} \rceil} 2^{-\ell d}$$
$$= \frac{2^d}{2^d - 1} 2^{-d \lceil \frac{\log_2(X_*)}{d} \rceil} \lesssim 2^{-d \lceil \frac{\log_2(X_*)}{d} \rceil}.$$

Thus, finiteness of the series on the left-hand side of (A.3) follows by noting

$$\mathbb{E}\left[X_*^2 \mathbb{1}_{X_* > 0} \frac{1}{2^{d \lceil \frac{\log_2(X_*)}{d} \rceil}} \right] \leq \mathbb{E}\left[X_*^2 \mathbb{1}_{X_* > 0} \frac{1}{X_*} \right] \leq \mathbb{E}[X_*] < \infty. \tag{A.4}$$

Here we have used the inequality $x \log_2^+(x) \geq x$ for $x \geq 2$, which shows that (A.2) implies $\mathbb{E}[X_*] < \infty$.

To estimate the second series, note that by $E \geq \mathbb{E}[Y_{j,\tau}]$ and Chebyshev's inequality

$$\begin{aligned}
\mathbb{P}(S_{j,\ell}^* \geq C + E) &= \mathbb{P}\Big(\frac{1}{N_\ell} \sum_{\tau = M_\ell + 1}^{M_{\ell+1}} (Y_{j,\tau} - E) \geq C \Big) \\
&\leq \mathbb{P}\Big(\frac{1}{N_\ell} \sum_{\tau = M_\ell + 1}^{M_{\ell+1}} (Y_{j,\tau} - \mathbb{E}[Y_{j,\tau}]) \geq C \Big) \\
&\leq \frac{1}{C^2} \mathbb{V}\Big[\frac{1}{N_\ell} \sum_{\tau = M_\ell + 1}^{M_{\ell+1}} (Y_{j,\tau} - \mathbb{E}[Y_{j,\tau}]) \Big] \\
&= \frac{1}{C^2 N_\ell^2} \sum_{\tau = M_\ell + 1}^{M_{\ell+1}} \mathbb{V}[Y_{j,\tau} - \mathbb{E}[Y_{j,\tau}]] \leq \frac{1}{C^2 N_\ell^2} \sum_{\tau = M_\ell + 1}^{M_{\ell+1}} \mathbb{E}[Y_{j,\tau}^2].
\end{aligned}$$

Next, since $M_\ell \asymp N_\ell \asymp 2^{\ell d}$ (see the proof of Lemma 3.1), there exist $\kappa_1, \kappa_2, \kappa_3 > 0$ only depending on d such that if $X_* \leq M_\ell$ and $M_\ell + 1 \leq \tau \leq M_{\ell+1}$, then $\kappa_1 \tau \leq N_\ell \leq \kappa_2 \tau$, and $\kappa_3 \tau \leq M_\ell \leq \tau$, as well as $X_* \leq \tau$. We deduce that

$$\begin{aligned}
\sum_{\ell \geq j} \mathbb{P}(S_{j,\ell}^* \geq C + E) &\leq \frac{1}{C^2} \sum_{\ell \geq j} \frac{1}{N_\ell^2} \sum_{\tau = M_\ell + 1}^{M_{\ell+1}} \mathbb{E}[Y_{j,\tau}^2] \\
&= \frac{1}{C^2} \sum_{\ell \geq j} \sum_{\tau = M_\ell + 1}^{M_{\ell+1}} \mathbb{E}\Big[\frac{X_*^2}{N_\ell^2} \mathbb{1}_{X_* \leq M_\ell} \Big] \\
&\leq \frac{1}{C^2} \sum_{\ell \geq j} \sum_{\tau = M_\ell + 1}^{M_{\ell+1}} \mathbb{E}\Big[\frac{X_*^2}{\kappa_1^2 \tau^2} \mathbb{1}_{X_* \leq \tau} \Big] \\
&\lesssim \sum_{\tau \geq M_j + 1} \mathbb{E}\Big[\frac{X_*^2}{\tau^2} \mathbb{1}_{X_* \leq \tau} \Big] = \mathbb{E}\Big[X_*^2 \sum_{\tau \geq M_j + 1} \frac{\mathbb{1}_{X_* \leq \tau}}{\tau^2} \Big] =: a_j.
\end{aligned}$$

For $N \geq 3$, there exists $n = n(N) \in \mathbb{N}$ such that $2^n + 1 \leq N \leq 2^{n+1}$, whence

$$\begin{aligned}(\text{A.5}) \qquad \sum_{\tau \geq N} \frac{\log_2(\tau)}{\tau^2} &\leq \sum_{\ell \geq n} \sum_{\tau=2^\ell+1}^{2^{\ell+1}} \frac{\log_2(\tau)}{\tau^2} \leq \sum_{\ell \geq n} \sum_{\tau=2^\ell+1}^{2^{\ell+1}} \frac{\ell+1}{(2^\ell)^2} \\ &= \sum_{\ell \geq n} \frac{\ell+1}{2^\ell} = \frac{n+1}{2^n} \sum_{\ell \geq n} \frac{\ell+1}{(n+1)2^{\ell-n}} \\ &\leq \frac{n+1}{2^n} \sum_{s \in \mathbb{N}_0} \frac{s+1}{2^s} \lesssim \frac{n+1}{2^n} \lesssim \frac{1+\log_2(N)}{N}.\end{aligned}$$

Here we have used $\frac{\ell+1}{n+1} = \frac{\ell-n+n+1}{n+1} \leq 1 + \ell - n$.

Also, it is not hard to check that $M_j + 1 \leq \tau$ is equivalent to $\frac{\log_2(\tau-1)}{d} \geq \log_2(2^{j+1}+1)$. Using the monotonicity of the logarithm, this implies $\frac{\log_2(\tau)}{d} \geq j + 1 \geq j$. Hence

$$\begin{aligned}&\sum_{j \in \mathbb{N}_0} \sum_{\ell \geq j} \mathbb{P}(S^*_{j,\ell} \geq C + E) \lesssim \sum_{j \in \mathbb{N}_0} a_j \\ &\quad = \mathbb{E}\Big[X_*^2 \sum_{j \in \mathbb{N}_0} \sum_{\tau \geq M_j+1} \frac{\mathbb{1}_{X_* \leq \tau}}{\tau^2}\Big] = \mathbb{E}\Big[X_*^2 \sum_{\tau \geq 3^d} \sum_{j \in \mathbb{N}_0} \mathbb{1}_{M_j+1 \leq \tau} \frac{\mathbb{1}_{X_* \leq \tau}}{\tau^2}\Big] \\ &\quad \leq \mathbb{E}\Big[X_*^2 \sum_{\tau \geq 3^d} \frac{\mathbb{1}_{X_* \leq \tau}}{\tau^2} \sum_{j=1}^{\lfloor \frac{\log_2(\tau)}{d} \rfloor} 1\Big] \leq \mathbb{E}\Big[X_*^2 \sum_{\tau \geq \max\{3^d, \lceil X_* \rceil\}} \frac{\log_2(\tau)}{d \cdot \tau^2}\Big] \\ &\quad \overset{(\text{A.5})}{\lesssim} \mathbb{E}\Big[X_*^2 \frac{1 + \log_2^+(\max\{3^d, \lceil X_* \rceil\})}{\max\{3^d, \lceil X_* \rceil\}}\Big] \lesssim \mathbb{E}\Big[X_*^2 \frac{1 + \log_2^+(X_*)}{X_*}\Big] \\ &\quad = \mathbb{E}[X_* \cdot (1 + \log_2^+(X_*))] = \mathbb{E}[X_*] + \mathbb{E}[X_* \log_2^+(X_*)] < \infty,\end{aligned}$$

as in (A.4).

Step 2. Next we show $\sum_{j \in \mathbb{N}_0} \sum_{\tau \in \mathbb{N}} \mathbb{P}(X_{j,\tau} \neq Y_{j,\tau}) < \infty$. The Borel–Cantelli lemma then implies $\mathbb{P}(X_{j,\tau} \neq Y_{j,\tau}$ for infinitely many $(j,\tau)) = 0$. That is, if all random variables are defined on the common probability space $(\Omega, \mathbb{P})$, the event A of all $\omega \in \Omega$ for which $X_{j,\tau} = Y_{j,\tau}$ happens for all but finitely many (j,τ) has probability 1. By Step 1, we know that

$$\Xi^* := \sup_{j \in \mathbb{N}_0} \Big(S^*_j + \sup_{\ell \geq j} S^*_{j,\ell}\Big) < \infty$$

holds in an almost sure event B. It is then easy to see that $\Xi < \infty$ in the almost sure event $A \cap B$.

In order to show that $\sum_{j \in \mathbb{N}_0} \sum_{\tau \in \mathbb{N}} \mathbb{P}(X_{j,\tau} \neq Y_{j,\tau}) < \infty$, first note that $X_{j,\tau} > M_\ell$ and $M_\ell + 1 \leq \tau \leq M_{\ell+1}$ imply $\kappa_4 X_{j,\tau} \geq \frac{M_{\ell+1}}{M_\ell} X_{j,\tau} > M_{\ell+1} \geq \tau$ for some $\kappa_4 > 0$ which only depends on d (which is irrelevant for the convergence of the series). The random variables $\kappa_4 X_{j,\tau}$ all have the same distribution

as the random variable $Y := \kappa_4 X_* = \kappa_4 X_{0,1}$. Hence,

$$\begin{aligned}(\text{A.6}) \qquad &\sum_{j\in\mathbb{N}_0}\sum_{\tau\in\mathbb{N}} \mathbb{P}(X_{j,\tau} \neq Y_{j,\tau}) \\ &\leq \sum_{j\in\mathbb{N}_0}\Big(\sum_{\tau=1}^{M_j}\mathbb{P}(X_{j,\tau} > M_j) + \sum_{\ell\geq j}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}\mathbb{P}(X_{j,\tau} > M_\ell)\Big) \\ &\leq \sum_{j\in\mathbb{N}_0}\Big(\sum_{\tau=1}^{M_j}\mathbb{P}(X_* > M_j) + \sum_{\ell\geq j}\sum_{\tau=M_\ell+1}^{M_{\ell+1}}\mathbb{P}(Y > \tau)\Big) \\ &= \sum_{j\in\mathbb{N}_0}\Big(M_j\mathbb{P}(X_* > M_j) + \sum_{\tau\geq M_j+1}\mathbb{P}(Y > \tau)\Big).\end{aligned}$$

Let us show that the two series in $(\text{A.6})_4$ are finite. First, observe thanks to the monotone convergence theorem that

$$\begin{aligned}\sum_{j\in\mathbb{N}_0} M_j\mathbb{P}(X_* > M_j) &= \sum_{j\in\mathbb{N}_0} M_j\mathbb{E}[\mathbb{1}_{X_*>M_j}] = \mathbb{E}\Big[\sum_{j\in\mathbb{N}_0} M_j\mathbb{1}_{X_*>M_j}\Big] \\ &\leq \mathbb{E}\Big[\sum_{j=1}^{\lfloor\frac{\log_2(X_*)}{d}\rfloor} M_j\Big] \lesssim \mathbb{E}\Big[\sum_{j=1}^{\lfloor\frac{\log_2(X_*)}{d}\rfloor} 2^{jd}\Big] \\ &\lesssim \mathbb{E}[2^{d\lfloor\frac{\log_2(X_*)}{d}\rfloor}] \leq \mathbb{E}[X_*] < \infty.\end{aligned}$$

Here we have used the fact that $X_* > M_j$ implies $X_* > 2^{jd}$, which in turn yields $j \leq \log_2(X_*)/d$. Second, we have

$$\begin{aligned}\sum_{j\in\mathbb{N}_0}\sum_{\tau\geq M_j+1}\mathbb{P}(Y > \tau) &= \sum_{j\in\mathbb{N}_0}\sum_{\tau\geq 3^d}\mathbb{1}_{\tau\geq M_j+1}\mathbb{P}(Y > \tau) \\ &\leq \sum_{\tau\geq 3^d}\Big[\mathbb{P}(Y > \tau)\sum_{j=1}^{\lfloor\log_2(\tau)/d\rfloor} 1\Big] = \mathbb{E}\Big[\sum_{\tau\geq 3^d}\mathbb{1}_{Y>\tau}\Big\lfloor\frac{\log_2(\tau)}{d}\Big\rfloor\Big] \\ &\leq \mathbb{E}\Big[\mathbb{1}_{Y>3}\sum_{\tau=3}^{\lfloor Y\rfloor}\log_2(\tau)\Big] \leq \mathbb{E}\Big[\mathbb{1}_{Y>3}\sum_{\tau=3}^{\lfloor Y\rfloor}\log_2^+(\lfloor Y\rfloor)\Big] \\ &\leq \mathbb{E}\Big[\lfloor Y\rfloor\log_2^+(\lfloor Y\rfloor)\Big] \leq \mathbb{E}[Y\log_2^+(Y)] \\ &\leq \kappa_4\mathbb{E}[X_*(\log_2^+(\kappa_4) + \log_2^+(X_*))] \\ &= \kappa_4\log_2^+(\kappa_4)\mathbb{E}[X_*] + \kappa_4\mathbb{E}[X_*\log_2^+(X_*)] < \infty,\end{aligned}$$

as in (A.4). ∎

The following auxiliary statement is used in the proof of Remark 5.5.

LEMMA A.11. *For every $n \in \mathbb{N}$, let $X^{(n)} = (X_1^{(n)}, \dots, X_n^{(n)})$ be a random vector whose entries $X_1^{(n)}, \dots, X_n^{(n)} \sim \mathcal{N}(0,1)$ are i.i.d. standard normal*

random variables. Then the following event has probability 1*:*

$$\|X^{(n)}\|_\infty \geq \sqrt{\log(n)} \quad \textit{for all but finitely many } n \in \mathbb{N}.$$

We emphasize that we do not assume that the random vectors $X^{(1)}, X^{(2)}, \dots$ are independent. It is only assumed that the components of $X^{(n)}$ for fixed $n \in \mathbb{N}$ are jointly independent.

Proof of Lemma A.11. We start by estimating the tail probability of $\|X^{(n)}\|_\infty$. Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers. Then

$$\mathbb{P}(\|X^{(n)}\|_\infty \leq a_n) = \mathbb{P}(|X_1^{(n)}| \leq a_n)^n = \operatorname{erf}(a_n/\sqrt{2})^n,$$

where $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2}\,\mathrm{d}t$ is the error function. To estimate the quantity $\mathbb{P}(\|X^{(n)}\|_\infty \leq a_n)$ further, we first need to establish an inequality for the error function. An inequality for the so called Mills ratio for standard normal random variables [2, 7.1.13] implies for $x \geq 0$ that

$$0 \leq \exp(x^2)\frac{\sqrt{\pi}}{2}(1 - \operatorname{erf}(x)) > \frac{1}{x + \sqrt{x^2+2}}.$$

This in turn implies that for $x \geq 0$, we have

$$0 \leq \operatorname{erf}(x) < 1 - \frac{2}{\sqrt{\pi}} \frac{\exp(-x^2)}{x + \sqrt{x^2+2}}.$$

Using this inequality, we get

$$\begin{aligned}\operatorname{erf}(a_n/\sqrt{2}) &< 1 - \frac{2}{\sqrt{\pi}} \frac{\exp(-a_n^2/2)}{a_n/\sqrt{2} + \sqrt{a_n^2/2+2}} \\ &\leq 1 - \frac{2}{\sqrt{\pi}} \frac{\exp(-a_n^2/2)}{\sqrt{2}(a_n+1)},\end{aligned}$$

where the last inequality follows by the fact that $(a_n/\sqrt{2}+\sqrt{2})^2 \geq a_n^2/2+2$.

We now choose $a_n := \sqrt{2\log(2n) - \delta_n}$ with $0 < \delta_n < 2\log(2n)$, which gives the estimate

$$\begin{aligned}\operatorname{erf}(a_n/\sqrt{2}) &< 1 - \frac{2\exp(-\log(2n) + \delta_n/2)}{\sqrt{2\pi}(\sqrt{2\log(2n) - \delta_n} + 1)} \\ &= 1 - \frac{1}{\sqrt{2\pi}n(\sqrt{2\log(2n) - \delta_n} + 1)} \exp(\delta_n/2) \\ &\leq 1 - \frac{1}{\sqrt{2\pi}n(\sqrt{2\log(2n)} + 1)} \exp(\delta_n/2) =: 1 + x.\end{aligned}$$

Since $x \geq -1$, we can use the monotonicity of x^n on the non-negative numbers and the elementary estimate $1 + x \leq e^x$ to conclude

$$\operatorname{erf}(a_n/\sqrt{2})^n \leq (1+x)^n \leq \exp(xn) = \exp\Big(-\frac{\exp(\delta_n/2)}{\sqrt{2\pi}(\sqrt{2\log(2n)}+1)}\Big).$$

Choosing $\delta_n = K \log(\log(n))$ with $K > 0$ and noting that $\delta_n < 2\log(2n)$ for $n \geq n_0 = n_0(K)$ yields

$$\mathbb{P}(\|X^{(n)}\|_\infty \leq a_n) \leq \exp\left(-\frac{\log(n)^{K/2}}{\sqrt{2\pi}(\sqrt{2\log(2n)}+1)}\right). \tag{A.7}$$

Note that $\sqrt{2\log(2n)}+1 \leq \sqrt{2\log(n)}+\sqrt{2\log(2)}+1 \leq C\sqrt{\log(n)}$ for some absolute constant $C > 0$ and $n \geq 3$. Therefore,

$$\frac{\log(n)^{K/2}}{\sqrt{2\pi}(\sqrt{2\log(2n)}+1)} \geq \frac{1}{\sqrt{2\pi}C}\log(n)^{(K-1)/2} \geq 2\log(n)$$

for fixed $K > 3$ and $n \geq n_1 = n_1(K) \geq n_0(K)$. Combined with (A.7) this implies

$$\mathbb{P}(\|X^{(n)}\|_\infty \leq a_n) \leq \exp(-2\log(n)) = n^{-2}.$$

Thus

$$\sum_{n=1}^{\infty} \mathbb{P}(\|X^{(n)}\|_\infty \leq a_n) \leq \sum_{n=1}^{n_1-1} 1 + \sum_{n \geq n_1} \frac{1}{n^2} < \infty.$$

The Borel–Cantelli lemma thus shows that almost surely, there are only finitely many $n \in \mathbb{N}$ with $\|X^{(n)}\|_\infty \leq a_n$. Further note that

$$a_n = \sqrt{2\log(2n) - K\log(\log(n))} \geq \sqrt{\log(n)}$$

for all but finitely many $n \in \mathbb{N}$. Hence, we see that almost surely, we have $\|X^{(n)}\|_\infty \geq \sqrt{\log(n)}$ for all but finitely many $n \in \mathbb{N}$. ∎

Acknowledgements. The authors thank Jakob Lemvig for raising the question considered in this paper, and Andrei Caragea, Dorothee Haroske, and Dominik Stöger for their helpful comments.

References

[1] F. Abramovich, T. Sapatinas, and B. W. Silverman, *Wavelet thresholding via a Bayesian approach*, J. R. Statist. Soc. Ser. B Statist. Methodol. 60 (1998), 725–749.

[2] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Applied Mathematics Ser. 55, U.S. Government Printing Office, Washington, DC, 1964.

[3] J.-M. Aubry and S. Jaffard, *Random wavelet series*, Comm. Math. Phys. 227 (2002), 483–514.

[4] S. Aziznejad and J. Fageot, *Wavelet analysis of the Besov regularity of Lévy white noise*, Electron. J. Probab. 25 (2020), art. 158, 38 pp.

[5] H. Bauer, *Probability Theory*, De Gruyter Stud. Math. 23, de Gruyter, Berlin, 1996.

[6] N. Bochkina, *Besov regularity of functions with sparse random wavelet coefficients*, arXiv:1310.3720v2 (2013).

[7] P. A. Cioica, S. Dahlke, N. Döhring, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling, *Adaptive wavelet methods for SPDEs*, in: Extraction of Quantifiable Information from Complex Systems, Lecture Notes in Comput. Sci. Engrg. 102, Springer, Cham, 2014, 83–107.
[8] M. Dashti, S. Harris, and A. Stuart, *Besov priors for Bayesian inverse problems*, Inverse Problems Imaging 6 (2012), 183–200.
[9] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Reg. Conf. Ser. Appl. Math. 61, SIAM, Philadelphia, PA, 1992.
[10] C. Esser, S. Jaffard, and B. Vedel, *Regularity properties of random wavelet series*, Theory Probab. Math. Statist. 110 (2024), 31–53.
[11] G. B. Folland, *Real analysis*, 2nd ed., Wiley, New York, 1999.
[12] G. Garrigós, A. Seeger, and T. Ullrich, *Haar frame characterizations of Besov–Sobolev spaces and optimal embeddings into their dyadic counterparts*, J. Fourier Anal. Appl. 29 (2023), art. 39, 51 pp.
[13] P. Grohs, A. Klotz, and F. Voigtlaender, *Phase transitions in rate distortion theory and deep learning*, Found. Comput. Math. 23 (2023), 329–392.
[14] D. Haroske and H. Triebel, *Entropy numbers in weighted function spaces and eigenvalue distributions of some degenerate pseudodifferential operators. I*, Math. Nachr. 167 (1994), 131–156.
[15] O. Kallenberg, *Foundations of Modern Probability*, 3rd ed., Probab. Theory Stochastic Modelling 99, Springer, Cham, 2021.
[16] H. Kekkonen, M. Lassas, E. Saksman, and S. Siltanen, *Random tree Besov priors—towards fractal imaging*, Inverse Problems Imaging 17 (2023), 507–531.
[17] A. Knoblauch, *Closed-form expressions for the moments of the binomial probability distribution*, SIAM J. Appl. Math. 69 (2008), 197–204.
[18] S. Lan, M. Pasha, S. Li, and W. Shen, *Spatiotemporal Besov priors for Bayesian inverse problems*, J. Amer. Statist. Assoc. 121 (2026), 1601–1615.
[19] M. Lassas, E. Saksman, and S. Siltanen, *Discretization-invariant Bayesian inversion and Besov space priors*, Inverse Problems Imaging 3 (2009), 87–122.
[20] H. Triebel, *Higher Analysis*, Hochschulbücher für Mathematik, Johann Ambrosius Barth Verlag, Leipzig, 1992.
[21] H. Triebel, *Theory of Function Spaces. III*, Monographs Math. 100, Birkhäuser, Basel, 2006.
[22] H. Triebel, *Function Spaces and Wavelets on Domains*, EMS Tracts in Math. 7, Eur. Math. Soc., Zürich, 2008.
[23] P. Wojtaszczyk, *A Mathematical Introduction to Wavelets*, London Math. Soc. Student Texts 37, Cambridge Univ. Press, Cambridge, 1997.

Andreas Horst
Department of Applied Mathematics and Computer Science
Technical University of Denmark
2800 Kgs. Lyngby, Denmark
E-mail: ahor@dtu.dk

Thomas Jahn, Felix Voigtlaender
Mathematical Institute for Machine Learning and Data Science (MIDS)
Katholische Universität Eichstätt–Ingolstadt
85049 Ingolstadt, Germany
E-mail: thomas.jahn@ku.de
felix.voigtlaender@ku.de