Multiscale functional inequalities in probability: constructive approach

Mitia Duerinckx, Antoine Gloria

WEIGHTED FUNCTIONAL INEQUALITIES: CONCENTRATION PROPERTIES

MITIA DUERINCKX AND ANTOINE GLORIA

Abstract. Consider an ergodic stationary random field A on the ambient space $\mathbb{R}^d$. We are interested in the concentration of measure phenomenon for nonlinear functions $X(A)$ in terms of assumptions on A. In mathematical physics, this phenomenon is often associated with functional inequalities like spectral gap or logarithmic Sobolev inequality. These inequalities are however only known to hold for a restricted class of laws (like product measures, Gaussian measures with integrable covariance, or more general Gibbs measures with nicely behaved Hamiltonians). In this contribution, we introduce a more general class of functional inequalities (which we call weighted functional inequalities) that strictly contains standard functional inequalities, and we study their concentration properties. As an application, we prove specific concentration results for averages of approximately local functions of the field A, which constitutes the main stochastic ingredient to the quenched large-scale regularity theory for random elliptic operators by the second author, Neukamm, and Otto. In a companion article, we develop a constructive approach to weighted functional inequalities based on product structures in higher-dimensional spaces, which allows us to treat all the examples of heterogeneous materials encountered in stochastic homogenization in the applied sciences.

Contents

1. Introduction 1
2. Weighted functional inequalities 3
2.1. Functional inequalities 3
2.2. Decay of correlations 5
2.3. Ergodicity and mixing 9
3. Moment bounds and concentration properties 12
3.1. Control of higher moments 12
3.2. Concentration properties 16
4. Application to spatial averages of the random field 21
4.1. Scaling of spatial averages 21
4.2. Concentration properties 22
Acknowledgements 28
References 28

1. Introduction

Functional inequalities like spectral gap, covariance, or logarithmic Sobolev inequalities are powerful tools to prove nonlinear concentration of measure properties and central limit theorem scalings. Besides their well-known applications in mathematical physics (e.g. for

the study of interacting particle systems like the Ising model, or for interface models), the sensitivity calculus provided by such inequalities is a quite convenient tool that was recently used to establish quantitative stochastic homogenization results, starting with the inspiring unpublished work by Naddaf and Spencer [22], and followed by [15, 16, 12, 18, 13].

These functional inequalities have nevertheless two main limitations for stochastic homogenization. On the one hand, whereas only few examples are known to satisfy them (besides product measures, Gaussian measures, and more general Gibbs measures with nicely behaved Hamiltonians), these inequalities are not robust with respect to various simple constructions: for instance, a Poisson point process satisfies a spectral gap, but the random field corresponding to the Voronoi tessellation of a Poisson point process does not. On the other hand, these functional inequalities require random fields to have an integrable covariance, which prevents one from considering fields with heavier tails.

This article constitutes the first of a series of works dedicated to *weighted* functional inequalities. In view of the application to quantitative stochastic homogenization, we shall consider a stationary random field A on $\mathbb{R}^d$, and study random variables that are nonlinear functions thereof. In this contribution, we introduce the notion of *weighted* functional inequalities and study the concentration properties that they imply. In the companion article [9], we develop a constructive approach to prove the validity of weighted functional inequalities for various examples of random fields considered in the literature (such as the Voronoi tessellation of a Poisson point process mentioned earlier, as well as Gaussian fields with non-integrable correlations, and the random parking point process). In particular, this allows us to address all the examples of [24], a reference textbook on random heterogeneous structures for materials science, which brings the use of functional inequalities (in their weighted versions) in stochastic homogenization to the state-of-the-art of materials science. In the third and last contribution [10], we turn to fluctuations and more precisely to weighted second-order Poincaré inequalities.

In Section 2, we start by recalling the standard definition of functional inequalities for (continuum) stationary random fields on $\mathbb{R}^d$ (in which case there is no canonical choice for the derivative with respect to the field), we introduce the notion of weighted functional inequalities, and quickly establish the relation between the weight and the decay of correlations, as well as the relation to standard notions of mixing. In Section 3, we investigate the concentration properties implied by the weighted functional inequalities, which are in particular shown to be stronger than those implied by the corresponding α-mixing. In the last section of this article, we specialize the analysis to spatial averages of (possibly nonlinear approximately local transformations of) the random field itself, in a form that is needed to establish sharp integrability estimates on the validity of the quenched large-scale regularity for random elliptic systems in [14] (see also [3, 2]).

Notation.

- d is the dimension of the ambient space $\mathbb{R}^d$;
- C denotes various positive constants that only depend on the dimension d and possibly on other controlled quantities; we write $\lesssim$ and $\gtrsim$ for $\le$ and $\ge$ up to such multiplicative constants C; we use the notation $\simeq$ if both relations $\lesssim$ and $\gtrsim$ hold; we add a subscript in order to indicate the dependence of the multiplicative constants on other parameters;

- the notation $a \ll b$ (or equivalently $b \gg a$) stands for $a \leq \frac{1}{C} b$ for some large enough constant $C \simeq 1$;
- $Q^k := [-1/2, 1/2)^k$ denotes the unit cube centered at 0 in dimension k, and for all $x \in \mathbb{R}^d$ and $r > 0$ we set $Q^k(x) := x + Q^k$, $Q^k_r := rQ^k$ and $Q^k_r(x) := x + rQ^k$; when $k = d$ or when there is no confusion possible on the meant dimension, we drop the superscript k;
- we use similar notation for balls, replacing Q^k by B^k (the unit ball in dimension k);
- the Euclidean distance between subsets of $\mathbb{R}^d$ is denoted by $d(\cdot, \cdot)$;
- $\mathcal{B}(\mathbb{R}^k)$ denotes the Borel σ-algebra on $\mathbb{R}^k$;
- $\mathbb{E}[\cdot]$ denotes the expectation, $\mathrm{Var}[\cdot]$ the variance, and $\mathrm{Cov}[\cdot;\cdot]$ the covariance in the underlying probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and the notation $\mathbb{E}[\cdot\|\cdot]$ stands for the conditional expectation;
- for all $a, b \in \mathbb{R}$, we set $a \wedge b := \min\{a, b\}$, $a \vee b := \max\{a, b\}$, and $a_+ := a \vee 0$;
- $\lceil a \rceil$ denotes the smallest integer larger or equal to a.

2. Weighted functional inequalities

2.1. **Functional inequalities.** Let $A : \mathbb{R}^d \times \Omega \to \mathbb{R}$ be a jointly measurable random field on $\mathbb{R}^d$, constructed on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. A spectral gap in probability for A is a functional inequality which allows one to control the variance of any function $X(A)$ in terms of its local dependence on A, i.e. in terms of some "derivative" of $X(A)$ with respect to local restrictions of A.

Let us be more precise. A map $\tilde{\partial} : \mathcal{B}(\mathbb{R}^d) \times \mathrm{Mes}(\Omega; \mathbb{R}) \to \mathrm{Mes}(\Omega; [0, \infty])$ is called a *(wide-sense) derivative with respect to* A if, for all $\sigma(A)$-measurable random variables $X(A), Y(A)$, all $\lambda, \mu \in \mathbb{R}$, and all Borel subsets $S \subset \mathbb{R}^d$,

(i) the random variable $\tilde{\partial}_{A,S} X(A)$ is $\sigma(A)$-measurable, and it vanishes a.s. whenever $X(A)$ is $\sigma(A|_{\mathbb{R}^d \setminus S})$-measurable;

(ii) we have

$$\left|\tilde{\partial}_{A,S}(\lambda X(A) + \mu Y(A))\right| \leq |\lambda|\, \tilde{\partial}_{A,S} X(A) + |\mu|\, \tilde{\partial}_{A,S} Y(A);$$

(iii) for all $R > 0$ the maps $\mathbb{R}^d \times \Omega \to [0, \infty] : (x, \omega) \mapsto \big(\tilde{\partial}_{A,B_R(x)} X(A)\big)(\omega)$ and $\mathbb{R}_+ \times \mathbb{R}^d \times \Omega \to [0, \infty] : (r, x, \omega) \mapsto \big(\tilde{\partial}_{A,B_r(x)} X(A)\big)(\omega)$ are measurable.

We then call $\tilde{\partial}_{A,S} X(A)$ a *(wide-sense) derivative of* $X(A)$ *with respect to* A *on* S, which we think of as a quantification of the functional dependence of $X(A)$ with respect to the restriction $A|_S$ of A on S. Given such a (wide-sense) derivative $\tilde{\partial}$ (see below for typical choices), we may recall the definition of the following standard functional inequalities.

Definition 2.1. We say that A satisfies the *(standard) spectral gap* ($\tilde{\partial}$-SG) *with radius* $R > 0$ *and constant* $C < \infty$ if for all $\sigma(A)$-measurable random variable $X(A)$ we have

$$\mathrm{Var}[X(A)] \leq C \int_{\mathbb{R}^d} \mathbb{E}\left[\left(\tilde{\partial}_{A,B_R(x)} X(A)\right)^2\right] dx; \tag{2.1}$$

it satisfies the *(standard) covariance inequality* ($\tilde{\partial}$-CI) *with radius* $R > 0$ *and constant* $C < \infty$ if for all $\sigma(A)$-measurable random variables $X(A)$ and $Y(A)$ we have

$$\mathrm{Cov}[X(A); Y(A)] \leq C \int_{\mathbb{R}^d} \mathbb{E}\left[\left(\tilde{\partial}_{A,B_R(x)} X(A)\right)^2\right]^{\frac{1}{2}} \mathbb{E}\left[\left(\tilde{\partial}_{A,B_R(x)} Y(A)\right)^2\right]^{\frac{1}{2}} dx; \tag{2.2}$$

it satisfies the *(standard) logarithmic Sobolev inequality* ($\tilde\partial$-LSI) *with radius* $R>0$ *and constant* $C<\infty$ if for all $\sigma(A)$-measurable random variable $Z(A)$ we have

$$\mathrm{Ent}\big[Z(A)^2\big]:=\mathbb{E}\left[Z(A)^2\log Z(A)^2\right]-\mathbb{E}\left[Z(A)^2\right]\log\mathbb{E}\left[Z(A)^2\right]$$
$$\leq C\int_{\mathbb{R}^d}\mathbb{E}\left[\Big(\tilde\partial_{A,B_R(x)}Z(A)\Big)^2\right]dx. \quad (2.3)$$

□

Recall that ($\tilde\partial$-CI) and ($\tilde\partial$-LSI) both imply ($\tilde\partial$-SG). The spectral gap (2.1) indeed follows from the covariance inequality (2.2) for the choice $Y=X$, while it follows from the logarithmic Sobolev inequality (2.3) for the choice $Z=1+\varepsilon X$ in the limit $\varepsilon\downarrow 0$.

In the continuum setting that we consider in this contribution, there is no canonical choice of a derivative with respect to the field A, and we describe below three such possible notions. We start with the derivative most commonly used in the literature (see e.g. [21]).

- As in the discrete setting, the so-called *Glauber derivative* $\partial^{\rm G}$ is defined as follows, letting A' denote an i.i.d. copy of A, and denoting by $\mathbb{E}'[\cdot]$ the expectation with respect to A' only,

$$\partial^{\rm G}_{A,S}X(A):=\mathbb{E}'\big[\big(X(A)-X(A')\big)^2\,\big\|\,A'|_{\mathbb{R}^d\setminus S}=A|_{\mathbb{R}^d\setminus S}\big]^{\frac12}, \quad (2.4)$$

  or equivalently, expanding the square,

$$\partial^{\rm G}_{A,S}X(A)=\Big(X(A)^2-2X(A)\,\mathbb{E}\big[X(A)\,\big\|\,A|_{\mathbb{R}^d\setminus S}\big]+\mathbb{E}\big[X(A)^2\,\big\|\,A|_{\mathbb{R}^d\setminus S}\big]\Big)^{\frac12}.$$

- The *oscillation* $\partial^{\rm osc}$, as used for instance in [17, 18], is formally defined by

$$\partial^{\rm osc}_{A,S}\,X(A)\ :=\operatorname*{sup\,ess}_{A,S}X(A)-\operatorname*{inf\,ess}_{A,S}X(A)$$
$$\text{“=”}\ \operatorname{sup\,ess}\Big\{X(\tilde A):\tilde A\in\mathrm{Mes}(\mathbb{R}^d;\mathbb{R}),\ \tilde A|_{\mathbb{R}^d\setminus S}=A|_{\mathbb{R}^d\setminus S}\Big\}$$
$$-\operatorname{inf\,ess}\Big\{X(\tilde A):\tilde A\in\mathrm{Mes}(\mathbb{R}^d;\mathbb{R}),\ \tilde A|_{\mathbb{R}^d\setminus S}=A|_{\mathbb{R}^d\setminus S}\Big\}, \quad (2.5)$$

  where the essential supremum and infimum are taken with respect to the measure induced by the field A on the space $\mathrm{Mes}(\mathbb{R}^d;\mathbb{R})$ (endowed with the cylindrical σ-algebra). This definition (2.5) of $\partial^{\rm osc}_{A,S}X(A)$ is not measurable in general, and we rather define

$$\partial^{\rm osc}_{A,S}\,X(A):=\mathcal{M}[X\|A|_{\mathbb{R}^d\setminus S}]+\mathcal{M}[-X\|A|_{\mathbb{R}^d\setminus S}]$$

  in terms of the conditional essential supremum $\mathcal{M}[\cdot\|A_{\mathbb{R}^d\setminus S}]$ given $\sigma(A|_{\mathbb{R}^d\setminus S})$, as introduced in [6] (using a Radon-Nikodym theorem in L^∞ due to [5]). Alternatively, we may follow [17, 18] and simply define $\partial^{\rm osc}_{A,S}X(A)$ as the measurable envelope of (2.5).
- The (integrated) *functional (or Malliavin) derivative* $\partial^{\rm fct}$, as used in the first version of [14] and in [11], is the closest generalization of the usual partial derivatives commonly used in the discrete setting. Let us denote by $M\subset\mathrm{L}^\infty(\mathbb{R}^d)$ some open set such that the random field A takes its values in M. Given a $\sigma(A)$-measurable random variable $X(A)$, and given an extension $\tilde X:M\to\mathbb{R}$, its Fréchet derivative

$\partial\tilde{X}(A)/\partial A \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ is defined for any compactly supported perturbation $\delta A \in \mathrm{L}^\infty(\mathbb{R}^d)$ by

$$\lim_{t\to 0} \frac{\tilde{X}(A+t\delta A)-\tilde{X}(A)}{t} = \int_{\mathbb{R}^d} \delta A(x) \frac{\partial \tilde{X}(A)}{\partial A}(x)\,dx,$$

if the limit exists. Since we are interested in the local averages of this derivative, we rather define for all bounded Borel subset $S \subset \mathbb{R}^d$,

$$\partial^{\mathrm{fct}}_{A,S} X(A) = \int_S \Big| \frac{\partial \tilde{X}(A)}{\partial A}(x) \Big| dx.$$

This derivative is additive with respect to the set S: for all disjoint Borel subsets $S_1, S_2 \subset \mathbb{R}^d$ we have $\partial^{\mathrm{fct}}_{A,S_1\cup S_2} X(A) = \partial^{\mathrm{fct}}_{A,S_1} X(A) + \partial^{\mathrm{fct}}_{A,S_2} X(A)$.

It is clear by definition that the oscillation dominates the Glauber derivative. Henceforth we use the notation $\tilde{\partial}$ for any of the above-defined (wide-sense) derivatives with respect to the random field A.

Satisfied for Gaussian random fields with integrable correlations and for product structures (see e.g. [9]), the standard functional inequalities (SG), (LSI), and (CI) are restrictive in the sense that they only hold for fields with sufficiently fast decaying correlations, which excludes many examples of practical interest (typically to stochastic homogenization, cf. [24]). One possible explanation why the standard spectral gap is particularly restrictive is that the RHS in (2.1) only takes into account functional dependences at distance at most R. The definition below relaxes the standard spectral gap by explicitly taking into account dependences upon derivatives with respect to A restricted on arbitrarily large sets, according to some given weight.

Definition 2.2. Given an integrable function $\pi : \mathbb{R}_+ \to \mathbb{R}_+$, we say that A satisfies the *weighted spectral gap* ($\tilde{\partial}$-WSG) *with weight* π if for all $\sigma(A)$-measurable random variable $X(A)$ we have

$$\mathrm{Var}\left[X(A)\right] \le \mathbb{E}\left[\int_0^\infty \int_{\mathbb{R}^d} \Big(\tilde{\partial}_{A,B_{\ell+1}(x)} X(A)\Big)^2 dx\,(\ell+1)^{-d}\pi(\ell)\,d\ell\right]. \tag{2.6}$$

Likewise, we define the corresponding weighted covariance inequality ($\tilde{\partial}$-WCI) and the weighted logarithmic Sobolev inequality ($\tilde{\partial}$-WLSI). □

Note that, as for standard functional inequalities, ($\tilde{\partial}$-WCI) and ($\tilde{\partial}$-WLSI) both imply ($\tilde{\partial}$-WSG). The standard functional inequalities of Definition 2.1 are recovered by taking a compactly supported weight π.

2.2. **Decay of correlations.** In this subsection we quantify the relation between the decay of correlations of the random field and the weight π in the corresponding weighted inequalities, extending the well-known result that the standard spectral gap and covariance inequality imply the integrability of the covariance and the finiteness of the range of dependence, respectively. Note in particular that ($\tilde{\partial}$-WCI) gives much more information than ($\tilde{\partial}$-WSG) on the covariance function. As shown in the companion article [9, Corollary 3.1], this result is sharp: in the Gaussian case each of the necessary conditions below is (essentially) sufficient.

Proposition 2.3. *Let A be a jointly measurable stationary random field on $\mathbb{R}^d$ with $\mathbb{E}\left[|A|^2\right] < \infty$, and let $\mathcal{C}(x) := \mathrm{Cov}\left[A(0); A(x)\right]$ denote its covariance function. When using the derivative $\tilde\partial = \partial^{\mathrm{osc}}$, further assume that A is bounded (except in item (iii)).*

(i) If A satisfies ($\tilde\partial$-SG), and if the covariance function $\mathcal{C}$ is nonnegative, then $\mathcal{C}$ is integrable.

(ii) If A satisfies ($\tilde\partial$-WSG) with weight π, and if the covariance function $\mathcal{C}$ is nonnegative, then $\mathcal{C}$ is integrable whenever $\int_0^\infty \ell^d \pi(\ell) d\ell < \infty$. More generally, $\mathcal{C}$ satisfies

$$\int_{\mathbb{R}^d} (1+|x|)^{-\alpha} \mathcal{C}(x) dx \le C_\alpha \begin{cases} \int_0^\infty (\ell+1)^{d-\alpha} \pi(\ell) d\ell, & \text{if } 0 \le \alpha < d; \\ \int_0^\infty \log^2(2+\ell) \pi(\ell) d\ell, & \text{if } \alpha = d; \\ \int_0^\infty \pi(\ell) d\ell, & \text{if } \alpha > d. \end{cases}$$

(iii) If A satisfies ($\tilde\partial$-CI) with radius $R+\varepsilon$ for all $\varepsilon > 0$, then the range of dependence of A is bounded by $2R$ (that is, for all Borel subsets $S, T \subset \mathbb{R}^d$ the restrictions $A|_S$ and $A|_T$ are independent whenever $d(S,T) > 2R$).

(iv) If A satisfies ($\tilde\partial$-WCI) with weight π, then the covariance function satisfies for all $x \in \mathbb{R}^d$,

$$|\mathcal{C}(x)| \le C \int_{\frac12(|x|-2)\vee 0}^\infty \pi(\ell) d\ell. \qquad \square$$

Proof. We split the proof into four steps.

Step 1. Proof of (i).
Let the field A satisfy ($\tilde\partial$-SG) with radius R. For any $L \ge 1$, the standard spectral gap applied to the $\sigma(A)$-measurable random variable $X(A) = \int_{B_L} A$ (which is well-defined by measurability and moment bounds on A) yields

$$\mathrm{Var}\left[\int_{B_L} A\right] \le C\mathbb{E}\left[\int_{\mathbb{R}^d} \left(\tilde\partial_{A,B_R(x)} \int_{B_L} A\right)^2 dx\right].$$

For each choice of the derivative $\tilde\partial$ (further assuming that A is bounded in the case $\tilde\partial = \partial^{\mathrm{osc}}$), we have

$$\mathbb{E}\left[\left(\tilde\partial_{A,B_R(x)} \int_{B_L} A\right)^2\right] \le C|B_R(x) \cap B_L|^2 \le C_R \mathbb{1}_{|x|\le R+L}.$$

Hence, for $L \ge 1$,

$$\int_{B_L}\int_{B_L} \mathrm{Cov}\left[A(x); A(y)\right] dxdy = \mathrm{Var}\left[\int_{B_L} A\right] \le C_R|B_{R+L}| \le C_R|B_L|.$$

Therefore, if $\mathcal{C}$ is nonnegative, we deduce

$$\int_{B_L} \mathcal{C} \lesssim \int_{B_L} \fint_{B_L} \mathcal{C}(x-y) dydx = \int_{B_L} \fint_{B_L} \mathrm{Cov}\left[A(x); A(y)\right] dydx \le C_R.$$

Letting $L \uparrow \infty$, we conclude that $\mathcal{C}$ is integrable.

Step 2. Proof of (ii).
Let the field A satisfy ($\tilde{\partial}$-WSG) with weight π, and assume that $\mathcal{C}$ is nonnegative. Repeating the argument of Step 1, we deduce for all $L \geq 1$,

$$\begin{aligned} L^d \int_{B_L} \mathcal{C}(x)dx &\lesssim \mathbb{E}\left[\Big(\int_{B_L}(A(x)-\mathbb{E}[A])dx\Big)^2\right] \\ &\leq \int_0^\infty \int_{\mathbb{R}^d} |B_{\ell+1}(x)\cap B_L|^2 dx\,(\ell+1)^{-d}\pi(\ell)d\ell \\ &\lesssim \int_0^L L^d(\ell+1)^d\pi(\ell)d\ell + \int_L^\infty L^{2d}\pi(\ell)d\ell \\ &\lesssim L^d\int_0^\infty (\ell+1)^d\pi(\ell)d\ell, \end{aligned}$$

which shows that $\mathcal{C}$ is integrable if $\int_0^\infty(\ell+1)^d\pi(\ell)d\ell<\infty$.
Let now $\alpha>0$ be fixed, and let $\gamma := \frac{1}{2}(d+\alpha)$. Assume that $\alpha \neq d$ (the case $\alpha = d$ can be treated similarly and yields the logarithmic correction). For all $L \geq 1$, the weighted spectral gap applied to the $\sigma(A)$-measurable random variable $X(A) = \int_{B_L}(1+|y|)^{-\gamma}A(y)dy$ yields

$$\begin{aligned} &\mathrm{Var}\left[\int_{B_L}(1+|y|)^{-\gamma}A(y)dy\right] \\ &\leq \mathbb{E}\left[\int_0^\infty\int_{\mathbb{R}^d}\Big(\tilde{\partial}_{A,B_{\ell+1}(x)}\int_{B_L}(1+|y|)^{-\gamma}A(y)dy\Big)^2 dx\,(\ell+1)^{-d}\pi(\ell)d\ell\right] \\ &\leq C\int_0^\infty\int_{\mathbb{R}^d}\Big(\int_{B_L\cap B_{\ell+1}(x)}(1+|y|)^{-\gamma}dy\Big)^2 dx\,(\ell+1)^{-d}\pi(\ell)d\ell. \end{aligned}$$

Hence,

$$\begin{aligned} &\int_{B_{2L}}\Big(\int_{B_L(-x)}(1+|x+y|)^{-\gamma}(1+|y|)^{-\gamma}dy\Big)\mathcal{C}(x)dx \\ &\qquad = \mathrm{Var}\left[\int_{B_L}(1+|y|)^{-\gamma}A(y)dy\right] \leq C_\alpha\int_0^\infty(\ell+1)^{(d-\alpha)\vee 0}\pi(\ell)d\ell, \end{aligned}$$

which yields the claim by passing to the limit $L\uparrow\infty$.

Step 3. Proof of (iii).
Let the field A satisfy ($\tilde{\partial}$-CI) with radius $R+\varepsilon$ for any $\varepsilon>0$. Given two Borel subsets $S,T\subset\mathbb{R}^d$ with $d(S,T)>2R$, choosing $\varepsilon := \frac{1}{3}(d(S,T)-2R)$, and noting that the sets $S+B_{R+\varepsilon}$ and $T+B_{R+\varepsilon}$ are disjoint, the covariance inequality ($\tilde{\partial}$-CI) with radius $R+\varepsilon$ implies for any $G\in\sigma(A|_S)$ and $H\in\sigma(A|_T)$,

$$\begin{aligned} &|\mathrm{Cov}\left[\mathbb{1}_G;\mathbb{1}_H\right]| \\ &\quad\leq C_\varepsilon\int_{(S+B_{R+\varepsilon})\cap(T+B_{R+\varepsilon})}\mathbb{E}\left[\Big(\tilde{\partial}_{A,B_{R+\varepsilon}(x)}\mathbb{1}_G\Big)^2\right]^{\frac{1}{2}}\mathbb{E}\left[\Big(\tilde{\partial}_{A,B_{R+\varepsilon}(x)}\mathbb{1}_H\Big)^2\right]^{\frac{1}{2}}dx \;=0. \end{aligned}$$

This shows that the σ-algebras $\sigma(A|_S)$ and $\sigma(A|_T)$ are independent.

Step 4. Proof of (iv).
Let the field A satisfy ($\tilde\partial$-WCI) with weight π. For all $x\in\mathbb{R}^d$ and all $\varepsilon>0$, the covariance inequality applied to the $\sigma(A)$-measurable random variables $⨍_{B_\varepsilon(x)}A$ and $⨍_{B_\varepsilon}A$ yields

$$\begin{aligned}\Big|⨍_{B_\varepsilon(x)}⨍_{B_\varepsilon}\mathcal{C}(y-z)dydz\Big| &= \Big|\mathrm{Cov}\Big[⨍_{B_\varepsilon(x)}A;⨍_{B_\varepsilon}A\Big]\Big|\\ &\le \int_0^\infty\int_{\mathbb{R}^d}\mathbb{E}\Big[\Big(\tilde\partial_{A,B_{\ell+1}(y)}⨍_{B_\varepsilon(x)}A\Big)^2\Big]^{\frac12}\mathbb{E}\Big[\Big(\tilde\partial_{A,B_{\ell+1}(y)}⨍_{B_\varepsilon}A\Big)^2\Big]^{\frac12}dy\,(\ell+1)^{-d}\pi(\ell)d\ell\\ &\le \int_0^\infty\int_{\mathbb{R}^d}\varepsilon^{-d}|B_\varepsilon(x)\cap B_{\ell+1}(y)|\varepsilon^{-d}|B_\varepsilon\cap B_{\ell+1}(y)|dy\,(\ell+1)^{-d}\pi(\ell)d\ell.\end{aligned}$$

Letting $\varepsilon\downarrow0$ and using the continuity of the function $\mathcal{C}$ (as a consequence of the stochastic continuity of the field A, which follows from its joint measurability), we deduce the claim: for all $x\in\mathbb{R}^d$,

$$|\mathcal{C}(x)|\le C\int_0^\infty|B_{\ell+1}(x)\cap B_{\ell+1}|\,(\ell+1)^{-d}\pi(\ell)d\ell\le C\int_{\frac12(|x|-2)\vee0}^\infty\pi(\ell)d\ell.\qquad\square$$

As the above proposition shows, if the weight π satisfies $\int_0^\infty(\ell+1)^d\pi(\ell)d\ell<\infty$, both ($\tilde\partial$-SG) and ($\tilde\partial$-WSG) with weight π imply that $\mathcal{C}$ is integrable. The following proposition establishes that (∂^{fct}-SG) and (∂^{fct}-WSG) are actually equivalent for such weights π. This result does not hold if ∂^{fct} is replaced by another derivative or if SG is replaced by CI.

Proposition 2.4. *Let A satisfy (∂^{fct}-WSG) (resp. (∂^{fct}-WLSI)) with some weight π. If $\int_0^\infty(\ell+1)^d\pi(\ell)d\ell<\infty$, then A satisfies (∂^{fct}-SG) (resp. (∂^{fct}-LSI)) with any radius $R>0$.* $\square$

Proof. Let $\varepsilon\in(0,1)$ be fixed. Let $X(A)$ be some $\sigma(A)$-measurable random variable. Cover the cube $Q_\ell(x)$ with the cubes $Q_\varepsilon(z_i^{x,\ell})$, $i=1,\dots,\lceil r/\varepsilon\rceil^d$, where $z_i^{x,\ell}\in\varepsilon\mathbb{Z}^d$ is an enumeration of $Q_{\varepsilon\lceil\ell/\varepsilon\rceil}(x)\cap\varepsilon\mathbb{Z}^d$. We then estimate

$$\begin{aligned}\Big(\int_{Q_\ell(x)}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2 &\le \Big(\sum_{i=1}^{\lceil\ell/\varepsilon\rceil^d}\int_{Q_\varepsilon(z_i^{x,\ell})}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2\\ &\le (1+\ell/\varepsilon)^d\sum_{i=1}^{\lceil\ell/\varepsilon\rceil^d}\Big(\int_{Q_\varepsilon(z_i^{x,\ell})}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2.\end{aligned}$$

For all $\ell>0$, and $y\in\varepsilon\mathbb{Z}^d$, there are at most $\lceil\ell\rceil^d$ possible values of $x\in\mathbb{Z}^d$ such that $y\in\{z_i^{x,\ell}:i=1,\dots,\ell^d\}$, so that we obtain

$$\begin{aligned}\sum_{x\in\mathbb{Z}^d}\Big(\int_{Q_\ell(x)}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2 &\le \varepsilon^{-d}(\ell+1)^d\sum_{x\in\mathbb{Z}^d}\sum_{i=1}^{\lceil\ell/\varepsilon\rceil^d}\Big(\int_{Q_\varepsilon(z_i^{x,\ell})}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2\\ &\le \varepsilon^{-d}(\ell+1)^{2d}\sum_{y\in\varepsilon\mathbb{Z}^d}\Big(\int_{Q_\varepsilon(y)}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2,\end{aligned}$$

which directly yields, bounding integrals on cubes by integral on balls, and sums by integrals,

$$\begin{aligned}\int_{\mathbb{R}^d}\Big(\fint_{B_{\ell+1}(x)}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2dx &\le (\ell+1)^{-2d}\sum_{x\in\mathbb{Z}^d}\Big(\int_{Q_{1+\sqrt{d}+\ell}(x)}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2\\ &\lesssim \varepsilon^{-d}\sum_{y\in\varepsilon\mathbb{Z}^d/\sqrt{d}}\Big(\int_{Q_{\varepsilon/\sqrt{d}}(y)}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2\\ &\lesssim \int_{\mathbb{R}^d}\Big(\fint_{B_\varepsilon(y)}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2dy. \end{aligned}\tag{2.7}$$

If A satisfies ($\partial^{\rm fct}$-WSG) with weight π, we deduce from the above inequality that for all $\varepsilon\in(0,1)$,

$$\begin{aligned}\mathrm{Var}\left[X(A)\right] &\le \int_0^\infty\int_{\mathbb{R}^d}\mathbb{E}\left[\Big(\int_{B_{\ell+1}(x)}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2\right]dx\,(\ell+1)^{-d}\pi(\ell)d\ell\\ &\lesssim \varepsilon^{-2d}\Big(\int_0^\infty(\ell+1)^d\pi(\ell)d\ell\Big)\int_{\mathbb{R}^d}\mathbb{E}\left[\Big(\int_{B_\varepsilon(x)}\Big|\frac{\partial X(A)}{\partial A}\Big|\Big)^2\right]dx,\end{aligned}$$

which shows that the field A also satisfies ($\partial^{\rm fct}$-SG) if $\int_0^\infty(\ell+1)^d\pi(\ell)d\ell<\infty$. □

2.3. **Ergodicity and mixing.** In the previous subsection we established the link between weighted functional inequality and the decay of the covariance function. We now turn to ergodicity properties, and further investigate the relation between weighted spectral gaps and standard mixing conditions.

Let us first recall some terminology. The random field A is said to be *strongly mixing* if for all $\sigma(A)$-measurable random variable $X(A)$ and all Borel subsets $E,E'\subset\mathbb{R}$ we have

$$\mathbb{P}\left[X(A)\in E,\,X(A(\cdot+x))\in E'\right]\overset{|x|\to\infty}{\longrightarrow}\mathbb{P}\left[X(A)\in E\right]\mathbb{P}\left[X(A)\in E'\right].$$

This qualitative property can be quantified into *strong mixing conditions*. A classical way to measure the dependence between two sub-σ-algebras $\mathcal{G}_1,\mathcal{G}_2\subset\mathcal{A}$ is the following *α-mixing coefficient*, first introduced by Rosenblatt [23],

$$\alpha(\mathcal{G}_1,\mathcal{G}_2):=\sup\big\{|\mathbb{P}[G_1\cap G_2]-\mathbb{P}[G_1]\mathbb{P}[G_2]|\,:\,G_1\in\mathcal{G}_1,G_2\in\mathcal{G}_2\big\}.$$

Applied to the random field A, this leads to the following measure of mixing: For all diameters $D\in(0,\infty]$ and distances $R>0$, we set

$$\begin{aligned}\tilde\alpha(R,D;A):=\sup\big\{\alpha(\sigma(A|_{S_1}),\sigma(A|_{S_2}))\,:\,&S_1,S_2\in\mathcal{B}(\mathbb{R}^d),\,d(S_1,S_2)\ge R,\\ &\mathrm{diam}(S_1),\mathrm{diam}(S_2)\le D\big\}.\end{aligned}\tag{2.8}$$

We say that the field A is *α-mixing* if for all diameter $D\in(0,\infty)$ we have $\tilde\alpha(R,D;A)\overset{R\uparrow\infty}{\longrightarrow}0$. Note that α-mixing is the weakest of the usual strong mixing conditions (see e.g. [8]), although it is in general strictly stronger than qualitative strong mixing.

The following result makes explicit the connection between weighted spectral gaps and α-mixing properties. Note that this result is essentially sharp: on the one hand, in the Gaussian case, as shown in the companion article [9, Corollary 3.1], each of the necessary conditions in (i), (ii), and (iv) below is (essentially) sufficient, and on the other hand

the R-scaling in the estimate in (iii) can be checked to be sharp at least in some specific examples.

Proposition 2.5. *Let A be a jointly measurable stationary random field on $\mathbb{R}^d$.*

(i) If A satisfies ($\tilde\partial$-WSG) with integrable weight π, then A is ergodic.
(ii) If A satisfies ($\tilde\partial$-WCI) with integrable weight π, then A is strongly mixing.
(iii) If A satisfies ($\tilde\partial$-WCI) with weight π and with derivative $\tilde\partial=\partial^{\rm G}$ or $\partial^{\rm osc}$, then A is α-mixing with coefficient $\tilde\alpha(R,D;A)\lesssim(1+\frac{D}{R})^d\int_{R-1}^\infty\pi(\ell)d\ell$.
(iv) If A satisfies ($\tilde\partial$-CI) with radius $R+\varepsilon>0$ for all $\varepsilon>0$, then $\tilde\alpha(r,\infty;A)=0$ for all $r>2R$. □

Remark 2.6. Item (iii) is expected to fail in general for the derivative $\tilde\partial=\partial^{\rm fct}$. Indeed, as shown in the companion paper [9, Corollary 3.1], if A is a stationary Gaussian random field with covariance function $\mathcal{C}$ satisfying $|\mathcal{C}(x)|\simeq(1+|x|)^{-\alpha}$ for all x, for some $\alpha>0$, then the field A satisfies ($\partial^{\rm fct}$-WCI) with weight $\pi(r)\simeq(1+r)^{-\alpha-1}$. Therefore, if item (iii) above was true with $\tilde\partial=\partial^{\rm fct}$, we would deduce in this Gaussian example $\tilde\alpha(R,D;A)\lesssim(1+(D/R)^d)R^{-\alpha}$, which is however expected to fail (the correct scaling is rather expected to be $R^{d-\alpha}$ for $\alpha>d$, cf. [8, Corollary 2 of Section 2.1.1] or [19, Corollary p.195]). □

Proof of Proposition 2.5. Item (iv) follows from Proposition 2.3. We split the rest of the proof into three steps.

Step 1. Proof of (i).
Let the field A satisfy ($\tilde\partial$-WSG) with weight π. To prove ergodicity, it suffices to show that for all integrable $\sigma(A)$-measurable random variables $X(A)$ we have

$$\lim_{L\uparrow\infty}\mathbb{E}\Big[\Big|\fint_{B_L}X(A(x+\cdot))dx-\mathbb{E}\left[X(A)\right]\Big|\Big]=0.$$

By an approximation argument in $\mathrm{L}^2(\Omega)$, we may assume that $X(A)$ is bounded and is $\sigma(A|_{B_R})$-measurable for some $R>0$. The spectral gap ($\tilde\partial$-WSG) applied to the $\sigma(A)$-measurable random variable $\fint_{B_L}X(A(\cdot+x))dx$ yields

$$\begin{aligned}S_L:=\mathbb{E}\Big[\Big|\fint_{B_L}X(A(x+\cdot))dx-\mathbb{E}\left[X(A)\right]\Big|\Big]^2&\le\mathrm{Var}\Big[\fint_{B_L}X(A(x+\cdot))dx\Big]\\&\le\mathbb{E}\Big[\int_0^\infty\int_{\mathbb{R}^d}\Big(\fint_{B_L}\tilde\partial_{A,B_{\ell+1}(y)}X(A(x+\cdot))dx\Big)^2dy\,(\ell+1)^{-d}\pi(\ell)d\ell\Big],\end{aligned}$$

and therefore

$$\begin{aligned}S_L\le\mathbb{E}\Big[\int_0^\infty\int_{\mathbb{R}^d}\fint_{B_L}\fint_{B_L}\tilde\partial_{A,B_{\ell+1}(y)}X(A(x+\cdot))\,\tilde\partial_{A,B_{\ell+1}(y)}X(A(x'+\cdot))dxdx'dy\\\times(\ell+1)^{-d}\pi(\ell)d\ell\Big].\end{aligned}$$

By assumption, $\tilde\partial_{A,B_{\ell+1}(y)}X(A(x+\cdot))=0$ whenever $B_R(x)\cap B_{\ell+1}(y)=\varnothing$, i.e. whenever $|x-y|>R+\ell+1$. For the choices $\tilde\partial=\partial^{\rm osc}$ and $\partial^{\rm G}$, we also have $\tilde\partial_{A,B_{\ell+1}(y)}X(A(x+\cdot))\le$

$2\|X\|_{\mathrm{L}^\infty}$, so that the above yields

$$\begin{aligned}
S_L &\le 4\|X\|_{\mathrm{L}^\infty}^2 \int_0^\infty \int_{\mathbb{R}^d} \fint_{B_L}\fint_{B_L} \mathbb{1}_{|x-y|\le R+\ell+1}\mathbb{1}_{|x'-y|\le R+\ell+1} dx dx' dy\,(\ell+1)^{-d}\pi(\ell)d\ell \\
&= 4\|X\|_{\mathrm{L}^\infty}^2 L^{-2d}\int_0^\infty \Big(\int_{B_L}\int_{B_{R+\ell+1}(x)} |B_L\cap B_{R+\ell+1}(y)|dydx\Big)(\ell+1)^{-d}\pi(\ell)d\ell \\
&\le 4\|X\|_{\mathrm{L}^\infty}^2 \int_0^\infty (R+\ell+1)^d\Big(\frac{R+\ell}{L}\wedge 1\Big)^d(\ell+1)^{-d}\pi(\ell)d\ell,
\end{aligned}$$

where the RHS obviously goes to 0 as $L\uparrow\infty$ whenever $\int_0^\infty \pi(\ell)d\ell<\infty$. This proves ergodicity for the choices $\tilde\partial=\partial^{\mathrm{osc}}$ and ∂^{G}.

It remains to treat the case $\tilde\partial=\partial^{\mathrm{fct}}$. An additional approximation argument is then needed in order to restrict attention to those random variables $X(A)$ such that the derivative $\tilde\partial_{A,B_{\ell+1}(x)}X(A)$ is pointwise bounded. The stochastic continuity of the field A (which follows from its joint measurability) ensures that the $\sigma(A|_{B_R})$-measurable random variable $X(A)$ is actually $\sigma(A|_{\mathbb{Q}^d\cap B_R})$-measurable. A standard approximation argument then allows to construct a sequence $(x_n)_n\subset B_R$ and a sequence $(X_n(A))_n$ of random variables such that $X_n(A)$ is $\sigma((A(x_k))_{k=1}^n)$-measurable and converges to $X(A)$ in $\mathrm{L}^2(\Omega)$. By definition, we may write $X_n(A)=f_n(A(x_1),\dots,A(x_n))$ for some Borel function $f_n:(\mathbb{R}^k)^n\to\mathbb{R}$. Another standard approximation argument now allows to replace the Borel maps f_n's by smooth functions. We end up with a sequence that approximates $X(A)$ in $\mathrm{L}^2(\Omega)$, and such that the elements have pointwise bounded $\tilde\partial$-derivative. For these approximations, the conclusion follows as before.

Step 2. Proof of (ii).
Let the field A satisfy ($\tilde\partial$-WCI) with weight π. To prove strong mixing, it suffices to show that for all bounded $\sigma(A)$-measurable random variables $X(A)$ and $Y(A)$ we have $\mathrm{Cov}\,[X(A);Y(A(x+\cdot))]\to 0$ as $|x|\to\infty$ (since the desired property then follows by choosing the random variables $X(A),Y(A)$ to be any pair of indicator functions). Again, a standard approximation argument allows one to consider bounded $\sigma(A|_{B_R})$-measurable random variables $X(A),Y(A)$ for some $R>0$. Given $x\in\mathbb{R}^d$, apply the covariance inequality ($\tilde\partial$-WCI) to $X(A)$ and $Y(A(\cdot+x))$ to obtain

$$\begin{aligned}
&\big|\mathrm{Cov}\,[X(A);Y(A(x+\cdot))]\,\big| \\
&\le \int_0^\infty\int_{\mathbb{R}^d} \mathbb{E}\Big[\Big(\tilde\partial_{A,B_{\ell+1}(y)}X(A)\Big)^2\Big]^{\frac12}\mathbb{E}\Big[\Big(\tilde\partial_{A,B_{\ell+1}(y)}Y(A(x+\cdot))\Big)^2\Big]^{\frac12} dy\,(\ell+1)^{-d}\pi(\ell)d\ell.
\end{aligned}$$

By assumption, $\tilde\partial_{A,B_{\ell+1}(y)}X(A)=0$ whenever $B_R\cap B_{\ell+1}(y)=\varnothing$, i.e. whenever $|y|>R+\ell+1$. For the choices $\tilde\partial=\partial^{\mathrm{osc}}$ and ∂^{G}, we have in addition $\tilde\partial_{A,B_{\ell+1}(y)}X(A)\le 2\|X\|_{\mathrm{L}^\infty}$, so that the above directly yields

$$\begin{aligned}
&\big|\mathrm{Cov}\,[X(A);Y(A(x+\cdot))]\,\big| \\
&\quad\le\ 4\|X\|_{\mathrm{L}^\infty}\|Y\|_{\mathrm{L}^\infty}\int_0^\infty\int_{\mathbb{R}^d}\mathbb{1}_{|y|\le R+\ell+1}\mathbb{1}_{|x-y|\le R+\ell+1}\,dy\,(\ell+1)^{-d}\pi(\ell)d\ell \\
&\quad\lesssim\ \|X\|_{\mathrm{L}^\infty}\|Y\|_{\mathrm{L}^\infty}\int_0^\infty (R+\ell+1)^d(\ell+1)^{-d}\pi(\ell)d\ell
\end{aligned}$$

where the RHS goes to 0 as $|x| \to \infty$ whenever $\int_0^\infty \pi(\ell) d\ell < \infty$. This proves strong mixing for the choices $\tilde\partial = \partial^{\rm osc}$ and $\partial^{\rm G}$. In the case $\tilde\partial = \partial^{\rm fct}$, an additional approximation argument is needed as in Step 1 in order to restrict to random variables $X(A)$ such that $\tilde\partial_{A,B_{\ell+1}(y)} X(A)$ is pointwise bounded.

Step 3. Proof of (iii).
Let the field A satisfy ($\tilde\partial$-WCI) with weight π, and with derivative $\tilde\partial = \partial^{\rm osc}$ or $\partial^{\rm G}$. Given Borel subsets $S, T \subset \mathbb{R}^d$ with diameter $\le D$ and with $d(S,T) \ge 2R$, the covariance inequality ($\tilde\partial$-WCI) for this choice of derivatives yields for all bounded random variables $X(A)$ and $Y(A)$, respectively $\sigma(A|_S)$-measurable and $\sigma(A|_T)$-measurable,

$$\begin{aligned}
&\big|{\rm Cov}\,[X(A);Y(A)]\,\big| \\
&\quad\le \int_0^\infty \int_{\mathbb{R}^d} \mathbb{E}\Big[\Big(\tilde\partial_{A,B_{\ell+1}(x)} X(A)\Big)^2\Big]^{\frac12} \mathbb{E}\Big[\Big(\tilde\partial_{A,B_{\ell+1}(x)} Y(A)\Big)^2\Big]^{\frac12} dx\,(\ell+1)^{-d}\pi(\ell)d\ell \\
&\quad\le 4\|X(A)\|_{\rm L^\infty}\|Y(A)\|_{\rm L^\infty} \int_0^\infty \big|(S+B_{\ell+1})\cap(T+B_{\ell+1})\big|\,(\ell+1)^{-d}\pi(\ell)d\ell \\
&\quad\lesssim \|X(A)\|_{\rm L^\infty}\|Y(A)\|_{\rm L^\infty} \int_{R-1}^\infty (\ell+D+1)^d(\ell+1)^{-d}\pi(\ell)d\ell \\
&\quad\le \|X(A)\|_{\rm L^\infty}\|Y(A)\|_{\rm L^\infty}\Big(1+\frac{D}{R}\Big)^d \int_{R-1}^\infty \pi(\ell)d\ell,
\end{aligned}$$

from which the claim follows by choosing for $X(A), Y(A)$ any pair of indicator functions. □

3. Moment bounds and concentration properties

In this section, we investigate the concentration properties that are implied by weighted spectral gaps, according to both the choice of the derivative and the decay of the weight. Although the results are new, the proofs rely mainly on standard Herbst-type arguments.

3.1. **Control of higher moments.**

As for standard functional inequalities, weighted functional inequalities allow one to control higher moments of random variables. Note that these properties depend crucially on the choice of the derivative.

Proposition 3.1. *Assume that the random field A satisfies ($\tilde\partial$-WSG) with integrable weight $\pi : \mathbb{R}_+ \to \mathbb{R}_+$. Then there exists $C < \infty$ (depending only on π and d) such that for all $1 \le p < \infty$ and all $\sigma(A)$-measurable random variables $X(A)$ we have*

(i) if $\tilde\partial = \partial^{\rm G}$ or $\partial^{\rm fct}$,

$$\begin{aligned}
&\mathbb{E}\Big[\big(X(A)-\mathbb{E}\,[X(A)]\,\big)^{2p}\Big] \\
&\qquad\le (Cp^2)^p\,\mathbb{E}\bigg[\bigg(\int_0^\infty\int_{\mathbb{R}^d}\Big(\tilde\partial_{A,B_{\ell+1}(x)}X(A)\Big)^2 dx\,(\ell+1)^{-d}\pi(\ell)d\ell\bigg)^p\bigg],
\end{aligned}$$

where the multiplicative factor $(Cp^2)^p$ can be upgraded to $(Cp)^p$ if the field A further satisfies ($\tilde\partial$-WLSI);

(ii) if $\tilde\partial = \partial^{\rm osc}$,

$$\mathbb{E}\left[\big(X(A)-\mathbb{E}\left[X(A)\right]\big)^{2p}\right] \leq (Cp^2)^p\,\mathbb{E}\left[\int_0^\infty\bigg(\int_{\mathbb{R}^d}\Big(\tilde\partial_{A,B_{2(\ell+1)}(x)}X(A)\Big)^2dx\bigg)^p(\ell+1)^{-dp}\pi(\ell)d\ell\right]. \quad \square$$

Proof. Let $X(A)$ be $\sigma(A)$-measurable. We may assume without loss of generality that $\mathbb{E}\left[X(A)\right]=0$. We split the proof into two steps.

Step 1. Proof of (i) and (ii) for ($\tilde\partial$-WSG).
Applying the spectral gap ($\tilde\partial$-WSG) to the $\sigma(A)$-measurable random variable $|X(A)|^p$ yields

$$\mathbb{E}\left[X(A)^{2p}\right] \leq \mathbb{E}\left[|X(A)|^p\right]^2 + \mathbb{E}\left[\int_0^\infty\int_{\mathbb{R}^d}\Big(\tilde\partial_{A,B_{\ell+1}(x)}\big(|X(A)|^p\big)\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\right]. \quad (3.1)$$

For $p>2$, Hölder's and Young's inequalities with exponents $(\frac{2(p-1)}{p-2},\frac{2(p-1)}{p})$ and $(\frac{p-1}{p-2},p-1)$, respectively, imply for all $\delta>0$,

$$\begin{aligned}\mathbb{E}\left[|X(A)|^p\right]^2=\mathbb{E}\left[|X(A)|^{p\frac{p-2}{p-1}}|X(A)|^{\frac{p}{p-1}}\right]^2 &\leq \mathbb{E}\left[X(A)^{2p}\right]^{\frac{p-2}{p-1}}\mathbb{E}\left[X(A)^2\right]^{\frac{p}{p-1}}\\ &\leq \frac{p-2}{p-1}\delta\,\mathbb{E}\left[X(A)^{2p}\right]+\frac{1}{p-1}\delta^{2-p}\,\mathbb{E}\left[X(A)^2\right]^p.\end{aligned}$$

while for $p\leq 2$ Jensen's inequality simply yields $\mathbb{E}\left[|X(A)|^p\right]^2\leq\mathbb{E}\left[X(A)^2\right]^p$. Injecting these estimates into (3.1) for some $\delta\gtrsim 1$ small enough, we conclude for all $1\leq p<\infty$,

$$\mathbb{E}\left[X(A)^{2p}\right]\leq p^{-1}C^p\,\mathbb{E}\left[X(A)^2\right]^p + C\,\mathbb{E}\left[\int_0^\infty\int_{\mathbb{R}^d}\Big(\tilde\partial_{A,B_{\ell+1}(x)}\big(|X(A)|^p\big)\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\right].$$

Since $\mathbb{E}\left[X(A)^2\right]=\mathrm{Var}\left[X(A)\right]$ follows from the centering assumption, the first RHS term is estimated by the spectral gap ($\tilde\partial$-WSG). Further using Jensen's inequality, this leads to

$$\mathbb{E}\left[X(A)^{2p}\right]\leq p^{-1}C^p\,\mathbb{E}\left[\bigg(\int_0^\infty\int_{\mathbb{R}^d}\Big(\tilde\partial_{A,B_{\ell+1}(x)}X(A)\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\bigg)^p\right] + C\,\mathbb{E}\left[\int_0^\infty\int_{\mathbb{R}^d}\Big(\tilde\partial_{A,B_{\ell+1}(x)}\big(|X(A)|^p\big)\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\right]. \quad (3.2)$$

We split the rest of this step into three further substeps, and treat separately $\partial^{\rm fct}$, $\partial^{\rm G}$, and $\partial^{\rm osc}$.

Substep 1.1. Proof of (i) for $\tilde\partial=\partial^{\rm fct}$.
By the Leibniz rule, $\partial^{\rm fct}_{A,S}(|X(A)|^p)=p|X(A)|^{p-1}\partial^{\rm fct}_{A,S}X(A)$, so that Hölder's inequality with exponents $(\frac{p}{p-1},p)$ yields

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^\infty\int_{\mathbb{R}^d}\Big(\partial^{\rm fct}_{A,B_{\ell+1}(x)}\big(|X(A)|^p\big)\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\bigg]\\
&\quad\le\ p^2\,\mathbb{E}\bigg[X(A)^{2(p-1)}\int_0^\infty\int_{\mathbb{R}^d}\Big(\partial^{\rm fct}_{A,B_{\ell+1}(x)}X(A)\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\bigg]\\
&\quad\le\ p^2\,\mathbb{E}\big[X(A)^{2p}\big]^{1-\frac1p}\,\mathbb{E}\bigg[\bigg(\int_0^\infty\int_{\mathbb{R}^d}\Big(\partial^{\rm fct}_{A,B_{\ell+1}(x)}X(A)\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\bigg)^p\bigg]^{\frac1p}. \quad (3.3)
\end{aligned}$$

Combined with (3.2) and Young's inequality with exponents $(\frac{p}{p-1},p)$ to absorb the factor $\mathbb{E}\big[X(A)^{2p}\big]$ into the LHS, the conclusion of item (i) follows with the prefactor $(Cp^2)^p$.

Substep 1.2. Proof of (i) for $\tilde\partial=\partial^{\rm G}$.
The inequality $||a|^p-|b|^p|\le p|a-b|(|a|^{p-1}+|b|^{p-1})$ for all $a,b\in\mathbb{R}$ easily implies, by definition of the Glauber derivative (2.4),

$$\begin{aligned}
\mathbb{E}\bigg[\Big(\partial^{\rm G}_{A,S}\big(|X(A)|^p\big)\Big)^2\bigg]&=\mathbb{E}\Big[\mathbb{E}'\big[\big(|X(A')|^p-|X(A)|^p\big)^2\,\big\|\,A'|_{\mathbb{R}^d\setminus S}=A|_{\mathbb{R}^d\setminus S}\big]\Big]\\
&\le\ 2p^2\mathbb{E}\Big[\mathbb{E}'\big[\big(X(A)^{2(p-1)}+X(A')^{2(p-1)}\big)\big(X(A')-X(A)\big)^2\,\big\|\,A'|_{\mathbb{R}^d\setminus S}=A|_{\mathbb{R}^d\setminus S}\big]\Big]\\
&=\ 4p^2\mathbb{E}\bigg[X(A)^{2(p-1)}\Big(\partial^{\rm G}_{A,S}X(A)\Big)^2\bigg],
\end{aligned}$$

and we are now back to the situation of Substep 1.1.

Substep 1.3. Proof of (ii).
Again, the inequality $||a|^p-|b|^p|\le p|a-b|(|a|^{p-1}+|b|^{p-1})$ for all $a,b\in\mathbb{R}$ implies

$$\begin{aligned}
\partial^{\rm osc}_{A,S}\,|X(A)|^p&\le 2p\Big(\sup_{A,S}|X(A)|^{p-1}\Big)\,\partial^{\rm osc}_{A,S}\,X(A)\\
&\le 2p\Big(|X(A)|+\partial^{\rm osc}_{A,S}\,X(A)\Big)^{p-1}\,\partial^{\rm osc}_{A,S}\,X(A). \quad (3.4)
\end{aligned}$$

We then make use of the following inequality that holds for some constant $C\simeq1$ large enough (independent of p): for all $a,b\ge0$, $(a+b)^{p-1}\le 2a^{p-1}+(Cp)^pb^{p-1}$. This allows one to rewrite (3.4) in the form

$$\partial^{\rm osc}_{A,S}\,|X(A)|^p\le 4p|X(A)|^{p-1}\,\partial^{\rm osc}_{A,S}\,X(A)+(Cp)^p(\partial^{\rm osc}_{A,S}\,X(A))^p. \quad (3.5)$$

Arguing as in Substep 1.1, we obtain by Hölder's inequality,

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^\infty\int_{\mathbb{R}^d}\Big(\,\partial^{\rm osc}_{A,B_{\ell+1}(x)}\,|X(A)|^p\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\bigg]\\
&\quad\le Cp^2\,\mathbb{E}\big[X(A)^{2p}\big]^{1-\frac1p}\,\mathbb{E}\bigg[\bigg(\int_0^\infty\int_{\mathbb{R}^d}\Big(\,\partial^{\rm osc}_{A,B_{\ell+1}(x)}\,X(A)\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\bigg)^p\bigg]^{\frac1p}\\
&\qquad\qquad+(Cp^2)^p\,\mathbb{E}\bigg[\int_0^\infty\int_{\mathbb{R}^d}\Big(\,\partial^{\rm osc}_{A,B_{\ell+1}(x)}\,X(A)\Big)^{2p}dx\,(\ell+1)^{-d}\pi(\ell)d\ell\bigg].
\end{aligned}$$

Combined with (3.2) and Young's inequality to absorb the factor $\mathbb{E}[X(A)^{2p}]$ into the LHS, this yields

$$\mathbb{E}\left[X(A)^{2p}\right] \le (Cp^2)^p\,\mathbb{E}\left[\left(\int_0^\infty \int_{\mathbb{R}^d} \left(\partial^{\mathrm{osc}}_{A,B_{\ell+1}(x)}\, X(A)\right)^2 dx\,(\ell+1)^{-d}\pi(\ell)d\ell\right)^p\right] + (Cp^2)^p\,\mathbb{E}\left[\int_0^\infty \int_{\mathbb{R}^d} \left(\partial^{\mathrm{osc}}_{A,B_{\ell+1}(x)}\, X(A)\right)^{2p} dx\,(\ell+1)^{-d}\pi(\ell)d\ell\right].$$

It remains to reformulate the second RHS term. By the discrete $\ell^1-\ell^p$ inequality, we have

$$\begin{aligned}
\int_{\mathbb{R}^d} \left(\partial^{\mathrm{osc}}_{A,B_{\ell+1}(x)}\, X(A)\right)^{2p} dx &\le \sum_{z\in\frac{\ell+1}{\sqrt{d}}\mathbb{Z}^d} \int_{z+\frac{\ell+1}{\sqrt{d}}Q} \left(\partial^{\mathrm{osc}}_{A,B_{\ell+1}(x)}\, X(A)\right)^{2p} dx \\
&\le \left(\frac{\ell+1}{\sqrt{d}}\right)^d \sum_{z\in\frac{\ell+1}{\sqrt{d}}\mathbb{Z}^d} \left(\partial^{\mathrm{osc}}_{A,B_{\frac{3}{2}(\ell+1)}(z)}\, X(A)\right)^{2p} \\
&\le \left(\frac{\ell+1}{\sqrt{d}}\right)^d \left(\sum_{z\in\frac{\ell+1}{\sqrt{d}}\mathbb{Z}^d} \left(\partial^{\mathrm{osc}}_{A,B_{\frac{3}{2}(\ell+1)}(z)}\, X(A)\right)^{2}\right)^p \\
&\le \left(\frac{\ell+1}{\sqrt{d}}\right)^d \left(\sum_{z\in\frac{\ell+1}{\sqrt{d}}\mathbb{Z}^d} \fint_{z+\frac{\ell+1}{\sqrt{d}}Q} \left(\partial^{\mathrm{osc}}_{A,B_{2(\ell+1)}(x)}\, X(A)\right)^{2} dx\right)^p \\
&\le \left(\frac{\sqrt{d}}{\ell+1}\right)^{d(p-1)} \left(\int_{\mathbb{R}^d} \left(\partial^{\mathrm{osc}}_{A,B_{2(\ell+1)}(x)}\, X(A)\right)^{2} dx\right)^p. \qquad (3.6)
\end{aligned}$$

Combined with the above, this yields

$$\mathbb{E}\left[X(A)^{2p}\right] \le (Cp^2)^p\,\mathbb{E}\left[\left(\int_0^\infty \int_{\mathbb{R}^d} \left(\partial^{\mathrm{osc}}_{A,B_{\ell+1}(x)}\, X(A)\right)^2 dx\,(\ell+1)^{-d}\pi(\ell)d\ell\right)^p\right] + (Cp^2)^p\,\mathbb{E}\left[\int_0^\infty \left(\int_{\mathbb{R}^d} \left(\partial^{\mathrm{osc}}_{A,B_{2(\ell+1)}(x)}\, X(A)\right)^{2} dx\right)^p (\ell+1)^{-dp}\pi(\ell)d\ell\right].$$

Since $\int_0^\infty \pi(\ell)d\ell < \infty$, the first RHS term can be absorbed into the second RHS term. Indeed, the triangle inequality and the Hölder inequality with exponents $(p, \frac{p}{p-1})$ combine to

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^\infty \int_{\mathbb{R}^d} \left(\partial^{\mathrm{osc}}_{A,B_{\ell+1}(x)}\, X(A)\right)^2 dx\,(\ell+1)^{-d}\pi(\ell)d\ell\right)^p\right] \\
&\le \left(\int_0^\infty \mathbb{E}\left[\left(\int_{\mathbb{R}^d} \left(\partial^{\mathrm{osc}}_{A,B_{\ell+1}(x)}\, X(A)\right)^2 dx\right)^p\right]^{\frac{1}{p}} (\ell+1)^{-d}\pi(\ell)d\ell\right)^p \\
&= \left(\int_0^\infty \mathbb{E}\left[\left(\int_{\mathbb{R}^d} \left(\partial^{\mathrm{osc}}_{A,B_{\ell+1}(x)}\, X(A)\right)^2 dx\right)^p (\ell+1)^{-dp}\pi(\ell)\right]^{\frac{1}{p}} \pi(\ell)^{1-\frac{1}{p}}d\ell\right)^p \\
&\le \left(\int_0^\infty \pi(\ell)d\ell\right)^{p-1} \mathbb{E}\left[\int_0^\infty \left(\int_{\mathbb{R}^d} \left(\partial^{\mathrm{osc}}_{A,B_{\ell+1}(x)}\, X(A)\right)^2 dx\right)^p (\ell+1)^{-dp}\pi(\ell)d\ell\right],
\end{aligned}$$

and the conclusion of item (ii) follows.

Step 2. Improvement of (i) for ($\tilde\partial$-WLSI).
In this step, we argue that the prefactor $(Cp^2)^p$ in item (i) can be upgraded to $(Cp)^p$ if the field A satisfies the corresponding logarithmic Sobolev inequality ($\tilde\partial$-WLSI). Starting point is the following observation (see [1, Theorem 3.4] and [4, Proposition 5.4.2]): if the random variable $X(A)$ satisfies $\mathrm{Ent}\big[X(A)^{2p}\big]<\infty$, then we have

$$\mathbb{E}\left[X(A)^{2p}\right]^{\frac1p}-\mathbb{E}\left[X(A)^2\right]=\int_1^p\frac{1}{q^2}\mathbb{E}\left[X(A)^{2q}\right]^{\frac1q-1}\mathrm{Ent}\big[X(A)^{2q}\big]\,dq. \tag{3.7}$$

It remains to estimate the entropy $\mathrm{Ent}[X(A)^{2q}]$ for all $1\le q\le p$. Applied to the $\sigma(A)$-measurable random variable $|X(A)|^q$, ($\tilde\partial$-WLSI) yields

$$\mathrm{Ent}\big[X(A)^{2q}\big]\le\mathbb{E}\left[\int_0^\infty\int_{\mathbb{R}^d}\Big(\tilde\partial_{A,B_{\ell+1}(x)}|X(A)|^q\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\right].$$

For the choice $\tilde\partial=\partial^{\mathrm{G}}$ or ∂^{fct}, the argument of Substeps 1.1–1.2, cf. (3.3), applied to the above RHS yields

$$\begin{aligned}&\mathrm{Ent}\big[X(A)^{2q}\big]\\&\quad\le Cq^2\,\mathbb{E}\left[X(A)^{2q}\right]^{1-\frac1q}\mathbb{E}\left[\bigg(\int_0^\infty\int_{\mathbb{R}^d}\Big(\tilde\partial_{A,B_{\ell+1}(x)}X(A)\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\bigg)^q\right]^{\frac1q}.\end{aligned}$$

Inserting this into (3.7), we obtain

$$\begin{aligned}\mathbb{E}\left[X(A)^{2p}\right]^{\frac1p}&\le\mathbb{E}\left[X(A)^2\right]\\&\quad+C\int_1^p\mathbb{E}\left[\bigg(\int_0^\infty\int_{\mathbb{R}^d}\Big(\tilde\partial_{A,B_{\ell+1}(x)}X(A)\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\bigg)^q\right]^{\frac1q}dq.\end{aligned}$$

We then appeal to the spectral gap ($\tilde\partial$-WSG) (which follows from ($\tilde\partial$-WLSI)) to estimate the first RHS term, and use Jensen's inequality on the second RHS to obtain

$$\mathbb{E}\left[X(A)^{2p}\right]^{\frac1p}\le Cp\,\mathbb{E}\left[\bigg(\int_0^\infty\int_{\mathbb{R}^d}\Big(\tilde\partial_{A,B_{\ell+1}(x)}X(A)\Big)^2dx\,(\ell+1)^{-d}\pi(\ell)d\ell\bigg)^p\right]^{\frac1p}.$$

This upgrades the prefactor in item (i) to $(Cp)^p$, as claimed. □

3.2. **Concentration properties.** The following results establish concentration properties implied by weighted functional inequalities, and extend the known results for standard functional inequalities. Again, these properties depend crucially on the choice of the derivative. On the one hand, spectral gaps for the Glauber and functional derivatives imply exponential tail concentration, and the corresponding logarithmic Sobolev inequalities imply stronger Gaussian tail concentration. On the other hand, for other choices of the derivative, the failure of the Leibniz rule in general only yields weaker results (except when the weight has compact support or when additional properties are assumed on the random variable, cf. Propositions 3.3(i) and 4.3(iii) below). Most of the following results are direct consequences of the p-versions of Proposition 3.1. We start with the concentration properties for the functional and Glauber derivatives.

Proposition 3.2. *Assume that the random field A satisfies ($\tilde{\partial}$-WSG) with integrable weight $\pi : \mathbb{R}_+ \to \mathbb{R}_+$ and derivative $\tilde{\partial} = \partial^{\mathrm{G}}$ or ∂^{fct}. We define the Lipschitz norm of a $\sigma(A)$-measurable random variable $X(A)$ with respect to the derivative $\tilde{\partial}$ and the weight π as*

$$\| X \|_{\tilde{\partial},\pi} := \operatorname*{sup\,ess}_A \left(\int_0^\infty \int_{\mathbb{R}^d} \left(\tilde{\partial}_{A,B_{\ell+1}(x)} X(A) \right)^2 dx \, (\ell+1)^{-d} \pi(\ell) d\ell \right)^{\frac{1}{2}}.$$

Then there exists a constant $C > 0$ depending only on d and π such for all $\sigma(A)$-measurable random variables $X(A)$ with $\| X \|_{\tilde{\partial},\pi} \leq 1$ we have exponential tail concentration in the form

$$\mathbb{E}\left[\exp\left(\frac{1}{C} |X(A) - \mathbb{E}\left[X(A)\right]| \right) \right] \leq 2,$$
$$\mathbb{P}\left[X(A) - \mathbb{E}\left[X(A)\right] \geq r \right] \leq e^{-\frac{r}{C}}, \qquad \text{for all } r \geq 0.$$

If in addition A satisfies ($\tilde{\partial}$-WLSI) with weight π, then for all $\sigma(A)$-measurable random variables $X(A)$ with $\| X \|_{\tilde{\partial},\pi} \leq 1$ we have Gaussian tail concentration in the form

$$\mathbb{E}\left[\exp\left(\frac{1}{C} \big(X(A) - \mathbb{E}\left[X(A)\right] \big)^2 \right) \right] \leq 2,$$
$$\mathbb{P}\left[X(A) - \mathbb{E}\left[X(A)\right] \geq r \right] \leq e^{-\frac{r^2 \vee r}{C}}, \qquad \text{for all } r \geq 0. \qquad \square$$

We now turn to the case of the oscillation, which yields in general weaker concentration results due to the failure of the Leibniz rule.

Proposition 3.3.

(i) Assume that the random field A satisfies (∂^{osc}-SG) with radius $R > 0$. Then for all $\sigma(A)$-measurable random variables $X(A)$ that satisfy

$$\| X \|_{\partial^{\mathrm{osc}},R} := \operatorname*{sup\,ess}_A \int_{\mathbb{R}^d} \left(\partial^{\mathrm{osc}}_{A,B_R(x)} X(A) \right)^2 dx \leq 1,$$

we have exponential tail concentration in the form

$$\mathbb{E}\left[\exp\left(\frac{1}{C} |X(A) - \mathbb{E}\left[X(A)\right]| \right) \right] \leq 2,$$
$$\mathbb{P}\left[X(A) - \mathbb{E}\left[X(A)\right] \geq r \right] \leq e^{-\frac{r}{C}}, \qquad \text{for all } r \geq 0.$$

If in addition A satisfies (∂^{osc}-LSI) with radius $R > 0$ and if the random variable $X(A)$ further satisfies

$$L := \sup_x \operatorname*{sup\,ess}_A \partial^{\mathrm{osc}}_{A,B_R(x)} X(A) < \infty,$$

we have Poisson tail concentration in the form

$$\mathbb{E}\left[\exp\left(\frac{1}{C} \psi_L\big(|X(A) - \mathbb{E}\left[X(A)\right]| \big) \right) \right] \leq 2, \qquad \psi_L(u) := \frac{u}{L} \log\left(1 + \frac{Lu}{C} \right),$$
$$\mathbb{P}\left[X(A) - \mathbb{E}\left[X(A)\right] \geq r \right] \leq e^{-\frac{1}{C}\psi_L(r)}, \qquad \text{for all } r \geq 0.$$

(ii) Assume that the random field A satisfies ($\partial^{\rm osc}$-WSG) with integrable weight $\pi:\mathbb{R}_+\to\mathbb{R}_+$. Let $X(A)$ be a $\sigma(A)$-measurable random variable, and assume that, for some $\kappa>0$, $p_0,\alpha\geq 0$, we have for all $p\geq p_0$,

$$\mathbb{E}\bigg[\int_0^\infty\bigg(\int_{\mathbb{R}^d}\Big(\partial^{\rm osc}_{A,B_{\ell+1}(x)}\,X(A)\Big)^2dx\bigg)^p(\ell+1)^{-dp}\pi(\ell)d\ell\bigg]\,\leq\,p^{\alpha p}\kappa. \tag{3.8}$$

Then there exists a constant $C>0$ depending only on d, π, p_0, and α (but not on κ) such that we have concentration in the form

$$\mathbb{E}\bigg[\psi_{p_0,\alpha}\Big(\frac{1}{C}|X(A)-\mathbb{E}\,[X(A)]\,|\Big)\bigg]\,\leq\,C\kappa,\qquad \psi_{p_0,\alpha}(u):=(1\wedge r^{2p_0})\exp(r^{\frac{2}{2+\alpha}}),$$

$$\mathbb{P}\,[\,|X(A)-\mathbb{E}\,[X(A)]\,|\geq r]\leq C\kappa\Big(\psi_{p_0,\alpha}\big(\tfrac{r}{C}\big)\Big)^{-1},\qquad \text{for all } r\geq 0. \qquad \square$$

Remark 3.4. Comments are in order.

- For spatial averages of (possibly nonlinear approximately local transformations of) the random field A, one can prove much stronger concentration results using the specific structure of averages, cf. Proposition 4.3(iii) below.
- Proposition 3.3(ii) above is used in two contexts. When the weight π is algebraic, the decay in (3.8) is typically independent of p (that is, $\alpha=0$, and κ is not to the power p so that it cannot be absorbed by rescaling of X), in which case κ is the driving quantity (see e.g. the application in Proposition 4.3(ii)). When the weight is super-algebraic, there can be an interplay between the decay of the weight and the power p, and an optimization may allow to put part of the decay to the power p at the price of losing some power of p itself — which leads to (3.8) for some $\alpha>0$ (after rescaling of X). $\square$

We start with the proof of Proposition 3.2.

Proof of Proposition 3.2. If A satisfies ($\tilde\partial$-WSG) for $\tilde\partial=\partial^{\rm G}$ or $\partial^{\rm fct}$, the assumption $\vertiii{X}_{\tilde\partial,\pi}\leq 1$ allows to apply Proposition 3.1(i) in the form

$$\mathbb{E}\Big[\big(X(A)-\mathbb{E}\,[X(A)]\,\big)^{2p}\Big]\leq(Cp^2)^p, \tag{3.9}$$

for all $p\geq 1$. Summing this estimate over p, and recalling that $n^n\leq e^n n!$, the exponential concentration result (i) follows in the form

$$\mathbb{E}\bigg[\exp\Big(\frac{1}{C}|X(A)-\mathbb{E}\,[X(A)]\,|\Big)\bigg]\leq 2,$$

and hence by Markov's inequality, for all $r\geq 0$,

$$\mathbb{P}\,\big[\big|X(A)-\mathbb{E}\,[X(A)]\,\big|\geq r\big]\leq 2e^{-\frac{r}{C}}.$$

The stronger unilateral estimate without the factor 2 is obtained by a standard application of Herbst-type techniques as in [7, Section 4] (see also [20, Section 2.5]).
If A further satisfies ($\tilde\partial$-WLSI) for $\tilde\partial=\partial^{\rm G}$ or $\partial^{\rm fct}$, Proposition 3.1(i) asserts that the RHS in (3.9) is replaced by $(Cp)^p$, which yields after summation the corresponding Gaussian concentration result (ii) in the form

$$\mathbb{E}\bigg[\exp\Big(\frac{1}{C}\big(X(A)-\mathbb{E}\,[X(A)]\,\big)^2\Big)\bigg]\leq 2,$$

and hence by Markov's inequality, for all $r \geq 0$,

$$\mathbb{P}\left[\left|X(A) - \mathbb{E}\left[X(A)\right]\right| \geq r\right] \leq 2e^{-\frac{r^2}{C}}.$$

The stronger unilateral estimate without the factor 2 is obtained by a standard application of Herbst's argument as e.g. in [21, Section 5.1]. □

We now turn to the proof of Proposition 3.3.

Proof of Proposition 3.3. We split the proof into two steps, and prove (i) and (ii) separately.

Step 1. Proof of (i).
The exponential concentration result in (i) follows from Proposition 3.1(ii) (with compactly supported weight π) as in the proof of Proposition 3.2 above. Let us now turn to the Poisson concentration result; although it could similarly be proven by first deriving suitable moment bounds, the proof is more transparent using a variation of Herbst's argument. Let A satisfy ($\partial^{\rm osc}$-LSI) and let $X(A)$ satisfy $L := \sup_x \sup \operatorname{ess}_A \partial^{\rm osc}_{A,B_R(x)} X(A) < \infty$ and $\|\!| X |\!|_{\partial^{\rm osc},R} \leq 1$. For all $t \in \mathbb{R}$, we apply ($\partial^{\rm osc}$-LSI) to the $\sigma(A)$-measurable random variable $e^{tX(A)/2}$,

$$\operatorname{Ent}[e^{tX(A)}] \leq C\,\mathbb{E}\left[\int_{\mathbb{R}^d} \Big(\,\partial^{\rm osc}_{A,B_R(x)}\, e^{tX(A)/2}\Big)^2 dx\right]. \tag{3.10}$$

By the inequality $|e^a - e^b| \leq (e^a + e^b)|a-b|$ for all $a, b \in \mathbb{R}$, the integrand turns into

$$\begin{aligned}\Big(\,\partial^{\rm osc}_{A,S}\, e^{tX(A)/2}\Big)^2 &\leq 2t^2 \sup_{A,S} e^{tX(A)} \Big(\,\partial^{\rm osc}_{A,S}\, X(A)\Big)^2 \\ &\leq 2t^2 e^{tX(A)} \exp\Big(t\,\partial^{\rm osc}_{A,S}\, X(A)\Big)\Big(\,\partial^{\rm osc}_{A,S}\, X(A)\Big)^2. \end{aligned}\tag{3.11}$$

Inserting this inequality into (3.10) and using the assumptions on $X(A)$, we obtain

$$\operatorname{Ent}[e^{tX(A)}] \leq Ct^2 e^{tL} \mathbb{E}[e^{tX(A)}].$$

Compared to the standard Herbst argument, we have to deal here with the additional exponential factor e^{tL}. We may then appeal to [20, Corollary 2.12] which indeed yields the desired Poisson concentration. We include a proof for the reader's convenience. In terms of the Laplace transform $H(t) = \mathbb{E}[e^{tX(A)}]$, the above takes the form

$$tH'(t) - H(t)\log H(t) \leq Ct^2 e^{tL} H(t),$$

or equivalently,

$$\frac{d}{dt}\Big(\frac{1}{t}\log H(t)\Big) \leq Ce^{tL},$$

and hence by integration

$$H(t) \leq \exp\Big(\frac{Ct}{L}(e^{tL}-1) + t\frac{H'(0)}{H(0)}\Big) = e^{\frac{Ct}{L}(e^{tL}-1) + t\mathbb{E}[X(A)]}.$$

The Markov inequality then implies for all $r, t \geq 0$,

$$\begin{aligned}\mathbb{P}\left[X(A) \geq \mathbb{E}\left[X(A)\right] + r\right] = \mathbb{P}\left[e^{tX(A)} \geq e^{t\mathbb{E}[X(A)]+tr}\right] &\leq e^{-t\mathbb{E}[X(A)]-tr}\mathbb{E}[e^{tX(A)}] \\ &\leq e^{\frac{Ct}{L}(e^{tL}-1)-tr}. \end{aligned}\tag{3.12}$$

Let $r \geq 0$ be momentarily fixed, and denote by $t_* \geq 0$ the value of $t \geq 0$ that minimizes $f_r(t) := \frac{Ct}{L}(e^{tL}-1) - tr$, that is the (unique) solution $t_* \geq 0$ of the equation

$$Ce^{t_*L} = (Lr+C)/(1+t_*L) \tag{3.13}$$

(note that f_r is strictly convex, $f_r(0)=0$, and $f_r'(0) \leq 0$). We now give two estimates on $f_r(t_*)$ depending on the value of r. Assume first that $r \geq \frac{2eC}{L}$. We may then compute

$$f_r(t_*) := \frac{Ct_*}{L}(e^{t_*L}-1) - t_*r \overset{(3.13)}{=} -\frac{t_*^2(Lr+C)}{1+t_*L}.$$

Using the bound $2t_*L \geq t_*L + \log(1+t_*L) \overset{(3.13)}{=} \log(1+Lr/C)$, and the fact that $t \mapsto -\frac{t^2(Lr+C)}{1+tL}$ is decreasing on $\mathbb{R}_+$, we obtain

$$f_r(t_*) \leq -\frac{Lr+C}{2L^2}\frac{\log(1+Lr/C)^2}{2+\log(1+Lr/C)}.$$

Hence, for $r \geq \frac{2eC}{L}$, we obtain using in addition $\log(1+Lr/C) \geq \log(1+2e) > 9/5$,

$$f_r(t_*) \leq -\frac{r}{5L}\log\Big(1+\frac{Lr}{C}\Big). \tag{3.14}$$

We now turn to the case $0 \leq r \leq \frac{2eC}{L}$. Comparing the minimal value $f_r(t_*)$ to the choice $t = \frac{r}{2eC}$, and using the bound $e^a - 1 \leq ea$ for $a \in [0,1]$, we obtain for all $r \leq \frac{2eC}{L}$,

$$f_r(t_*) \leq f_r\Big(\frac{r}{2eC}\Big) = \frac{r}{2eL}\big(e^{\frac{rL}{2eC}}-1\big) - \frac{r^2}{2eC} \leq -\frac{r^2}{4eC},$$

which yields, using that $\log(1+a) \leq a$ for all $a \geq 0$,

$$f_r(t_*) \leq -\frac{r}{4eL}\log\Big(1+\frac{Lr}{C}\Big) \leq -\frac{r}{11L}\log\Big(1+\frac{Lr}{C}\Big).$$

Combining this with (3.12) and (3.14), we conclude

$$\mathbb{P}\left[X(A) \geq \mathbb{E}\left[X(A)\right] + r\right] \leq e^{-\frac{r}{11L}\log(1+\frac{Lr}{C})},$$

and the corresponding integrability result follows by integration.

Step 2. Proof of (ii).
Let A satisfy ($\partial^{\rm osc}$-WSG) with weight π, and let the random variable $X(A)$ satisfy (3.8) for some $\kappa > 0$, $p_0, \alpha \geq 0$. Proposition 3.1(ii) then yields for all $p \geq p_0$,

$$\mathbb{E}\Big[\big(X(A) - \mathbb{E}\left[X(A)\right]\big)^{2p}\Big] \leq C^p p^{(2+\alpha)p}\kappa,$$

or alternatively, for all $p \geq (2+\alpha)p_0$,

$$\mathbb{E}\Big[\Big(\big|X(A) - \mathbb{E}\left[X(A)\right]\big|^{\frac{2}{2+\alpha}}\Big)^p\Big] \leq C^p p!\,\kappa.$$

Summing this estimate over p, we obtain

$$\mathbb{E}\left[\tilde\psi_{p_0,\alpha}\Big(\frac{1}{C}|X(A) - \mathbb{E}\left[X(A)\right]|^{\frac{2}{2+\alpha}}\Big)\right] \leq \kappa,$$

where we have set $\tilde\psi_{p_0,\alpha}(u) := \sum_{n=0}^\infty \frac{u^{n+(2+\alpha)p_0}}{(n+(2+\alpha)p_0)!}$. Noting that $\tilde\psi_{p_0,\alpha}(u) \simeq_{p_0,\alpha} (1\wedge u)^{(2+\alpha)p_0}e^u$ holds for all $u \geq 0$, the conclusion follows. □

4. Application to spatial averages of the random field

Although the primary aim of this contribution is to address concentration properties for general *nonlinear* functions of correlated random fields, we illustrate the use of weighted functional inequalities on the simplest functions possible, that is, (linear) spatial averages of (a possibly nonlinear yet approximately local transformation of) the random field itself.

Given a jointly measurable stationary random field A, we typically consider a $\sigma(A)$-measurable random variable $f(A)$ that is approximately 1-local with respect to the field A, in the following sense: for all $r>0$ we assume

$$\sup_A \operatorname{ess} \left| f(A) - \mathbb{E}\left[f(A) \,\|\, A|_{B_r} \right] \right| \le C e^{-\frac{r}{C}}. \tag{4.1}$$

More precisely, given $\tilde\partial = \partial^{\rm G}$, $\partial^{\rm fct}$, or $\partial^{\rm osc}$, we will use the following finer notion of approximate 1-locality: for all $x\in\mathbb{R}^d$ and $\ell\ge 0$,

$$\sup_A \operatorname{ess}\, \tilde\partial_{A,B_{\ell+1}(x)} f(A) \le C e^{-\frac{1}{C}(|x|-\ell)_+}. \tag{4.2}$$

(An important particular case is when the random variable $f(A)$ is exactly 1-local, that is, when $f(A)$ is $\sigma(A|_{B_1})$-measurable.) We then set $F(x):=f(A(\cdot+x))$ for all $x\in\mathbb{R}^d$, and for all $L\ge 0$ we consider the random variable

$$X_L := X_L(A) := L^{-d}\int_{\mathbb{R}^d} e^{-\frac{1}{L}|y|}(F(y)-\mathbb{E}[F])\,dy,$$

that is, the spatial average of (the nonlinear approximately local transformation F of) the random field A at the scale L. Note that the results below hold in the same form if X_L is replaced by $\fint_{Q_L}(F-\mathbb{E}[F])$.

As emphasized in the introduction, this example turns out to be relevant for quantitative stochastic homogenization, and more precisely to quantify the quenched large-scale regularity theory for random elliptic systems in divergence form (that is, operators of the form $-\nabla\cdot A\nabla$ with A a matrix-valued random coefficient field as considered throughout this article). In [14] the second author, Neukamm, and Otto indeed reduce the validity of this large-scale regularity to concentration properties of spatial averages X_L of the square of an approximately local version of the modified extended corrector (cf. [14, Proposition 3]), and then make direct use of weighted functional inequalities in the form of Proposition 4.3 below. More precisely, large-scale regularity is characterized in [14] by the so-called minimal radius r_*, an almost surely finite stationary random field whose stochastic integrability essentially coincides with the scaling in L of the probability $\mathbb{P}[X_L\ge\delta]$ for some fixed (small) $\delta>0$, in the sense that a property of the form $\mathbb{P}[X_L\ge\delta]\le g_\delta(L)^{-1}$ for all $L\ge 1$ essentially implies $\mathbb{E}[g_\delta(\frac{1}{C_\delta}r_*)]<\infty$.

We believe that these concentration results can also be used in the approach to large-scale regularity by Armstrong & Smart [3] and Armstrong & Mourrat [2] (which are rather formulated in terms of α-mixing assumptions). As shown below, concentration properties implied by weighted functional inequalities are in general stronger than those implied by the corresponding α-mixing.

4.1. **Scaling of spatial averages.**

We start with the scaling of the variance of the spatial average X_L. A similar result holds in stochastic homogenization, where X_L is replaced by the spatial average of the gradient of the extended corrector (cf. [13]).

Proposition 4.1. *If A satisfies ($\tilde\partial$-WSG) with integrable weight π and derivative $\tilde\partial=\partial^{\rm G}$, $\partial^{\rm fct}$, or $\partial^{\rm osc}$, and if the random variable $f(A)$ satisfies (4.2), then we have for all $L>0$,*

$$\mathrm{Var}\,[X_L]\lesssim \pi_*(L)^{-1},$$

where we define

$$\pi_*(\ell):=\Big(\fint_{B_\ell}\int_{|x|}^\infty \pi(s)dsdx\Big)^{-1}.$$ □

Remark 4.2. If $\pi(\ell)\simeq(\ell+1)^{-1-\beta}$ for some $\beta>0$, then

$$\pi_*(\ell)\simeq\begin{cases}(\ell+1)^\beta, & \text{if } \beta<d;\\ (\ell+1)^d\log^{-1}(2+\ell), & \text{if } \beta=d;\\ (\ell+1)^d, & \text{it } \beta>d.\end{cases}$$

In particular if correlations are integrable (corresponding to the case $\beta>d$), we recover the central limit theorem scaling: $\mathrm{Var}\,[X_L]\lesssim\pi_*(L)^{-1}\simeq L^{-d}$ for all $L\ge1$. □

Proof of Proposition 4.1. Let $L>0$. Given $\tilde\partial=\partial^{\rm G}$, $\partial^{\rm fct}$, and $\partial^{\rm osc}$, assumption (4.2) yields

$$\begin{aligned}|\tilde\partial_{A,B_{\ell+1}(x)}X_L|&\lesssim L^{-d}\int_{\mathbb{R}^d}e^{-\frac1L|y|}e^{-\frac1C(|x-y|-\ell)_+}dy\lesssim L^{-d}\int_{\mathbb{R}^d}e^{-\frac1L|y|}e^{-\frac{1}{C(\ell+1)}|x-y|}dy\\&\lesssim L^{-d}(L\wedge(\ell+1))^de^{-\frac{1}{C(L+\ell+1)}|x|},\end{aligned}$$

so that the weighted spectral gap yields

$$\begin{aligned}\mathrm{Var}\,[X_L]&\lesssim\int_0^\infty\int_{\mathbb{R}^d}L^{-2d}(L\wedge(\ell+1))^{2d}e^{-\frac{1}{C(L+\ell+1)}|x|}dx\,(\ell+1)^{-d}\pi(\ell)d\ell\\&\lesssim\int_0^\infty L^{-2d}(L\wedge(\ell+1))^{2d}(L+\ell)^d(\ell+1)^{-d}\pi(\ell)d\ell\\&\lesssim L^{-d}\int_0^L(\ell+1)^d\pi(\ell)d\ell+\int_L^\infty\pi(\ell)d\ell.\end{aligned}$$

An integration by parts yields $\pi_*(L)^{-1}\simeq L^{-d}\int_0^L\pi(\ell)\ell^d d\ell+\int_L^\infty\pi(\ell)d\ell$, and the conclusion follows. □

4.2. **Concentration properties.** In view of their applications to large-scale regularity of random elliptic systems in [14], we study the concentration properties of the spatial average X_L. The following result shows that the scaling crucially depends on three properties: the type of weighted functional inequality, the type of derivative, and the decay of the weight.

Proposition 4.3. *Assume that the random variable $f(A)$ satisfies (4.2).*

(i) Let A satisfy ($\tilde\partial$-WSG) with integrable weight π and derivative $\tilde\partial=\partial^{\rm G}$ or $\partial^{\rm fct}$, and let π_* be defined as in Proposition 4.1. Then for all $\delta,L>0$ we have

$$\mathbb{P}\,[X_L\ge\delta]\le\exp\Big(-\frac{\delta}{C}\pi_*(L)^{\frac12}\Big).\tag{4.3}$$

If in addition A satisfies ($\tilde\partial$-WLSI) with weight π, then for all $\delta,L>0$ we have

$$\mathbb{P}\,[X_L\ge\delta]\le\exp\Big(-\frac{\delta^2}{C}\pi_*(L)\Big).\tag{4.4}$$

(ii) Let A satisfy ($\partial^{\rm osc}$-WSG) with weight $\pi(\ell) \lesssim (\ell+1)^{-\beta-1}$ for some $\beta > 0$. Then for all $\delta, L > 0$ we have

$$\mathbb{P}\left[X_L \geq \delta\right] \leq C e^{-\frac{1}{C}\delta}\big(1 + \delta^{-\frac{2\beta}{d}} |\log \delta|\big)\, L^{-\beta}. \tag{4.5}$$

(iii) Let A satisfy ($\partial^{\rm osc}$-WSG) with weight $\pi(\ell) \lesssim \exp(-\frac{1}{C}\ell^\beta)$ for some $\beta > 0$. Then for all $\delta > 0$ and all $L \geq 1$ we have

$$\mathbb{P}\left[X_L \geq \delta\right] \leq \exp\Big(-\frac{\delta \wedge \delta^2}{C} L^{\beta \wedge \frac{d}{2}}\Big). \tag{4.6}$$

If in addition A satisfies ($\partial^{\rm osc}$-WLSI) with weight $\pi(\ell) \lesssim \exp(-\frac{1}{C}\ell^\beta)$ for some $\beta > 0$, then for all $\delta > 0$ and all $L \geq 1$ we have

$$\mathbb{P}\left[X_L \geq \delta\right] \leq \exp\Big(-\frac{\delta \wedge \delta^2}{C} L^{\beta \wedge d}\Big). \tag{4.7}$$

□

Remark 4.4. If we further assume that the random variable $f(A)$ is a.s. bounded by a deterministic constant $C_0 \geq 1$, then there holds $\mathbb{P}\left[|X_L| > C_0\right] = 0$, and hence in (4.6) and (4.7) we may replace $\delta \wedge \delta^2$ by $\frac{1}{C_0}\delta^2$. □

In the case of a super-algebraic weight, it is instructive to compare these (nonlinear) concentration results to the corresponding (linear) concentration result implied by the α-mixing properties of the field A (see also [2, Appendix A]). Note that the same result holds under the corresponding weighted covariance inequality (which is natural in view of Proposition 2.5(iii)). (As we are basically interested in the scaling in L, we do not try to optimize the $\log\delta$-dependence below.)

Proposition 4.5. *Given $\beta > 0$, assume that the random field A either is α-mixing with $\tilde\alpha(\ell, D; A) \lesssim (1+D)^C \exp(-\frac{1}{C}\ell^\beta)$ for all $D, \ell \geq 0$, or satisfies ($\tilde\partial$-WCI) with weight $\pi(\ell) \lesssim \exp(-\frac{1}{C}\ell^\beta)$ and derivative $\tilde\partial = \partial^{\rm G}$ or $\partial^{\rm osc}$. Further assume that the random variable $f(A)$ is a.s. bounded by a deterministic constant, that is, $\sup\operatorname{ess}_A |f(A)| \lesssim 1$, and that it satisfies* (4.1)*. Then for all $\delta > 0$ and all $L \geq 1$ we have*

$$\mathbb{P}\left[X_L > \delta\right] \leq C \exp\left(-\frac{\delta^2 (|\log\delta| + 1)^{-\frac{d\beta}{d+\beta}}}{C}\, L^{\frac{d\beta}{d+\beta}}\right). \qquad \square$$

Remark 4.6. Let us briefly compare the concentration results of Propositions 4.3(iii) and 4.5. Assume that the random field A satisfies a weighted functional inequality with super-algebraic weight $\pi(\ell) \lesssim \exp(-\frac{1}{C}\ell^\beta)$ and derivative $\partial^{\rm osc}$, and that A is α-mixing with $\tilde\alpha(\ell, D; A) \lesssim (1+D)^d \exp(-\frac{1}{C}\ell^\beta)$ (these assumptions are indeed compatible in view of Proposition 2.5(iii)). Then the decay in L of the probability $\mathbb{P}\left[X_L \geq \delta\right]$ obtained from the α-mixing is better than the one obtained from ($\partial^{\rm osc}$-WSG) only for $\beta > d$, and is always worse than the one obtained from ($\partial^{\rm osc}$-WLSI). Similarly, in the case of an algebraic weight $\pi(\ell) \lesssim (\ell+1)^{-\beta-1}$, the functional inequality ($\partial^{\rm osc}$-WSG) yields the optimal decay $L^{-\beta}$ (cf. Proposition 4.3(ii)), while one can check that the corresponding α-mixing only leads to this decay up to a small (sub-algebraic) loss. □

We start with the proof of Proposition 4.3.

Proof of Proposition 4.3. We split the proof into three steps. We start with the proofs of (4.3), (4.4), and (4.5), which directly follow from Propositions 3.2 and 3.3(ii). The proof of estimates (4.6) and (4.7) is more subtle and is based on a fine tuning of Herbst's argument using specific features of the random variable Z_L.

Step 1. Proof of (4.3), (4.4), and (4.5).
For $\tilde\partial = \partial^{\rm G}$ or $\partial^{\rm fct}$, let the Lipschitz norm $\|\!|\!| \cdot |\!|\!|_{\tilde\partial,\pi}$ be defined as in the statement of Proposition 3.2. The same computation as in the proof of Proposition 4.1 ensures that the random variable $Z_L := \pi_*(L)^{1/2} X_L = \pi_*(L)^{1/2} \fint_{Q_L} (F - \mathbb{E}\,[F])$ satisfies

$$|\!|\!| Z_L |\!|\!|_{\tilde\partial,\pi} \lesssim 1.$$

Hence, estimates (4.3) and (4.4) follow from Proposition 3.2. We now turn to the proof of (4.5). If A satisfies ($\partial^{\rm osc}$-WSG) with weight $\pi(\ell) \lesssim (\ell+1)^{-\beta-1}$, $\beta > 0$, we compute for all $p \ge p_0 > \frac{\beta}{d}$, using assumption (4.2),

$$\begin{aligned}\mathbb{E}\bigg[\int_0^\infty \bigg(\int_{\mathbb{R}^d} \Big(\partial^{\rm osc}_{A,B_\ell(x)} X_L\Big)^2 dx\bigg)^p (\ell+1)^{-dp-\beta-1} d\ell\bigg]&\\ \lesssim L^{-2dp} \int_1^\infty (L+\ell)^{dp} (L\wedge\ell)^{2dp} \ell^{-dp-\beta-1} d\ell \;\lesssim\; \big(1+(dp_0-\beta)^{-1}\big)\, L^{-\beta}.&\end{aligned}$$

Then applying Proposition 3.3(ii) and optimizing the choice of $p_0 > \frac{\beta}{d}$, the result (4.5) follows.

Step 2. Proof of (4.6).
Let $L \ge 1$, and define $Z_L := L^{d/2} X_L$. As in the proof of Proposition 4.1, assumption (4.2) yields

$$\partial^{\rm osc}_{A,B_\ell(x)} Z_L \lesssim L^{-\frac{d}{2}} (L \wedge (\ell+1))^d e^{-\frac{1}{C(L+\ell+1)}|x|}. \tag{4.8}$$

We make use of a variant of Herbst's argument as in [7, Section 4] (see also [20, Section 2.5]). For all $t \ge 0$ we apply ($\partial^{\rm osc}$-WSG) to the random variable $\exp(\frac{1}{2} t Z_L)$: using the inequality $|e^a - e^b| \le (e^a + e^b)|a-b|$ for all $a, b \in \mathbb{R}$, we obtain

$$\begin{aligned}\mathrm{Var}[e^{\frac{1}{2}tZ_L}] &\le \int_0^\infty \int_{\mathbb{R}^d} \mathbb{E}\bigg[\Big(\partial^{\rm osc}_{A,B_\ell(x)} e^{\frac{1}{2}tZ_L}\Big)^2\bigg] dx\, (\ell+1)^{-d} \pi(\ell) d\ell\\ &\lesssim t^2\, \mathbb{E}\big[e^{tZ_L}\big] \operatorname*{sup\,ess}_A \int_0^\infty \int_{\mathbb{R}^d} e^{t\partial^{\rm osc}_{A,B_\ell(x)} Z_L} \Big(\partial^{\rm osc}_{A,B_\ell(x)} Z_L\Big)^2 dx\, e^{-\frac{1}{C}\ell^\beta} d\ell,\end{aligned}$$

and hence, in terms of the Laplace transform $H_L(t) := \mathbb{E}\big[e^{tZ_L}\big]$,

$$H_L(t) - H_L(t/2)^2 \le\; t^2 H_L(t) \operatorname*{sup\,ess}_A \int_0^\infty \int_{\mathbb{R}^d} e^{t\partial^{\rm osc}_{A,B_\ell(x)} Z_L} \Big(\partial^{\rm osc}_{A,B_\ell(x)} Z_L\Big)^2 dx\, e^{-\frac{1}{C}\ell^\beta} d\ell.$$

Using the property (4.8) of the random variable Z_L, we find

$$\begin{aligned} &H_L(t) - H_L(t/2)^2 \\ &\lesssim\ t^2 H_L(t) \int_1^\infty \Big(\frac{L\wedge\ell}{\sqrt{L}}\Big)^{2d} \exp\bigg(Ct\Big(\frac{L\wedge\ell}{\sqrt{L}}\Big)^d - \frac{\ell^\beta}{C}\bigg) \int_{\mathbb{R}^d} e^{-\frac{1}{C(L+\ell+1)}|x|} dx\, d\ell \\ &\lesssim\ t^2 H_L(t) \int_1^\infty (L+\ell)^d \Big(\frac{L\wedge\ell}{\sqrt{L}}\Big)^{2d} \exp\bigg(Ct\Big(\frac{L\wedge\ell}{\sqrt{L}}\Big)^d - \frac{\ell^\beta}{C}\bigg) d\ell \\ &\lesssim\ t^2 H_L(t) \bigg(\int_0^L \exp\bigg(Ct\Big(\frac{\ell}{\sqrt{L}}\Big)^d - \frac{\ell^\beta}{C}\bigg) d\ell + L^d e^{CtL^{\frac{d}{2}}} \int_L^\infty e^{-\frac{1}{C}\ell^\beta} d\ell \bigg). \end{aligned} \tag{4.9}$$

Without loss of generality we may assume that $\beta \le \frac{d}{2}$ (the statement (4.6) is indeed not improved for $\beta > \frac{d}{2}$). We then restrict to

$$0 \le t \le T := \frac{1}{K} L^{\beta - \frac{d}{2}}, \tag{4.10}$$

for some $K \gg 1$ to be chosen later (with in particular $K \ge 2C^2$). As a consequence of $\beta \le \frac{d}{2}$, this choice yields $T \le K^{-1}$. On the one hand, for all $0 \le \ell \le L$ and all $0 \le t \le T$, the choice of T with $K \ge 2C^2$ yields

$$Ct\Big(\frac{\ell}{\sqrt{L}}\Big)^d - \frac{\ell^\beta}{C} = -\frac{L^\beta}{C}\bigg(\Big(\frac{\ell}{L}\Big)^\beta - \frac{C^2 t}{L^{\beta-\frac{d}{2}}}\Big(\frac{\ell}{L}\Big)^d\bigg) \le -\frac{L^\beta}{C}\bigg(\Big(\frac{\ell}{L}\Big)^\beta - \frac{1}{2}\Big(\frac{\ell}{L}\Big)^d\bigg) \le -\frac{\ell^\beta}{2C},$$

and hence

$$\int_0^L \exp\bigg(Ct\Big(\frac{\ell}{\sqrt{L}}\Big)^d - \frac{\ell^\beta}{C}\bigg) d\ell \lesssim \int_0^\infty e^{-\frac{\ell^\beta}{2C}} d\ell \lesssim 1.$$

On the other hand, for all $0 \le t \le T$, the choice of T with $K \ge 2C^2$ yields

$$L^d e^{CtL^{\frac{d}{2}}} \int_L^\infty e^{-\frac{1}{C}\ell^\beta} d\ell \lesssim \exp\Big(CtL^{\frac{d}{2}} - \frac{L^\beta}{2C}\Big) \le \exp\Big(\frac{CL^\beta}{K} - \frac{L^\beta}{2C}\Big) \le 1.$$

Injecting these estimates into (4.9), we obtain for all $0 \le t \le T$,

$$H_L(t) - H_L(t/2)^2 \le Ct^2 H_L(t),$$

and hence

$$H_L(t) \le \frac{H_L(t/2)^2}{1 - Ct^2}.$$

Applying the same inequality for $t/2$, iterating, and noting that $H_L(2^{-n}t)^{2^n} \to e^{t\mathbb{E}[Z_L]} = 1$ as $n \uparrow \infty$, we obtain for all $0 \le t \le T$,

$$H_L(t) \le \prod_{n=0}^\infty \big(1 - C(2^{-n}t)^2\big)^{-2^n}.$$

For K large enough such that $CT^2 \le CK^{-2} \le \frac{1}{2}$, the inequality $\log(1-x) \ge -2x$ for all $0 \le x \le \frac{1}{2}$ then yields for all $0 \le t \le T$,

$$\log H_L(t) \le -\sum_{n=0}^\infty 2^n \log\big(1 - C(2^{-n}t)^2\big) \le 2Ct^2 \sum_{n=0}^\infty 2^{-n} \lesssim t^2,$$

and thus $H_L(T) \le e^{CT^2}$. Using Markov's inequality and the choice (4.10) of T, we deduce for all $r \ge 0$,

$$\mathbb{P}\left[Z_L > r\right] \le e^{-Tr+CT^2} = \exp\Big(- \frac{L^{\beta-\frac{d}{2}} r}{K} + \frac{C}{K^2} L^{2\beta-d} \Big).$$

With the choice $r = \delta L^{\frac{d}{2}}$ for $\delta > 0$, this turns into

$$\mathbb{P}\left[X_L > \delta\right] \le \exp\Big(- \frac{\delta}{K} L^{\beta} + \frac{C}{K^2} L^{2\beta-d} \Big) \le \exp\bigg(- \frac{1}{K} \Big(\delta - \frac{C}{K} \Big) L^{\beta} \bigg).$$

Choosing $K \simeq 1 \vee \delta^{-1}$ large enough, the desired estimate (4.6) follows.

Step 2. Proof of (4.7).
Let $L \ge 1$, and define $Z_L := L^{d/2} X_L$. We make use of Herbst's classical argument as presented e.g. in [21, Section 5.1]. For all $t \ge 0$ we apply ($\partial^{\rm osc}$-WLSI) to the random variable $\exp(\frac{1}{2} t Z_L)$,

$$\mathrm{Ent}\big[e^{tZ_L}\big] \le \int_0^\infty \int \mathbb{E}\bigg[\Big(\partial^{\rm osc}_{A, B_\ell(x)} e^{\frac{1}{2} t Z_L} \Big)^2 \bigg] dx \, (\ell+1)^{-d} \pi(\ell) d\ell.$$

Estimating the RHS as in (4.9), we obtain in terms of $H_L(t) := \mathbb{E}[e^{tZ_L}]$,

$$\frac{d}{dt} \Big(\frac{1}{t} \log H_L(t) \Big) \lesssim \int_0^L \exp\bigg(Ct \Big(\frac{\ell}{\sqrt{L}} \Big)^d - \frac{\ell^\beta}{C} \bigg) d\ell + L^d e^{CtL^{\frac{d}{2}}} \int_L^\infty e^{-\frac{1}{C}\ell^\beta} d\ell.$$

Without loss of generality we may assume that $\beta \le d$ (the statement (4.7) is indeed not improved for $\beta > d$). We then restrict to

$$0 \le t \le T := \frac{1}{K} L^{\beta - \frac{d}{2}}, \tag{4.11}$$

for some $K \gg 1$ to be chosen later (with in particular $K \ge 2C$). Arguing as in Step 1, we obtain for all $0 \le t \le T$,

$$\frac{d}{dt} \Big(\frac{1}{t} \log H_L(t) \Big) \lesssim 1,$$

which yields by integration with respect to t on $[0, T]$,

$$\frac{1}{T} \log H_L(T) = \frac{1}{T} \log H_L(T) - \mathbb{E}\left[Z_L\right] \lesssim T,$$

that is, $H_L(T) \le e^{CT^2}$. The desired estimate (4.7) then follows as in Step 1, using Markov's inequality and choosing K large enough. □

We now turn to the proof of Proposition 4.5.

Proof of Proposition 4.5. Without loss of generality we assume that $\sup \mathrm{ess}_A \, |f(A)| \le 1$, which implies $\mathbb{P}\left[|X_L| > 1\right] = 0$. It is then sufficient to establish the result for $0 < \delta \le 1$. We split the proof into two steps. In the first step we prove the result in the case when the random variable $f(A)$ is exactly 1-local. We then extend the result in Step 2 when $f(A)$ is only approximately local in the sense (4.1). Since (WCI) implies α-mixing by Proposition 2.5(iii), it is enough to prove the result under the sole assumption of α-mixing.

Step 1. Exactly 1-local random variable $f(A)$.
In this step we assume in addition that $f(A)$ is $\sigma(A|_{B_1})$-measurable, and we prove that for all $\delta, L > 0$,

$$\mathbb{P}\left[X_L > \delta\right] \leq C \exp\Big(- \frac{\delta^2}{C} L^{\frac{d\beta}{d+\beta}} \Big). \tag{4.12}$$

Let $p \geq 1$ be an integer and let $R > 0$. Setting

$$E_{R,p} := \{(x_1, \ldots, x_{2p}) \in (\mathbb{R}^d)^{2p} : |x_1 - x_j| > R, \, \forall j \neq 1\},$$

and noting that for all x the random variable $F(x)$ is $\sigma(A|_{B(x)})$-measurable, α-mixing leads to

$$\begin{aligned} L^{-2dp} \Big| \int \ldots \int_{E_{R,p}} e^{-\frac{1}{L}\sum_{i=1}^{2p}|x_i|} \, \mathbb{E}\left[(F(x_1) - \mathbb{E}\left[F\right]) \ldots (F(x_{2p}) - \mathbb{E}\left[F\right])\right] dx_1 \ldots dx_{2p} \Big| \\ \leq C^p L^{-2dp} \int \ldots \int_{E_{R,p}} e^{-\frac{1}{L}\sum_{i=1}^{2p}|x_i|} \, \tilde{\alpha}\Big(R - 2, 2 + 2\sum_{i=1}^{2p} |x_i|; A \Big) dx_1 \ldots dx_{2p} \\ \leq C^p L^{-2dp} e^{-\frac{1}{C}R^\beta} \int_{\mathbb{R}^{2dp}} e^{-\frac{1}{L}|x|} (1 + |x|)^C dx \leq C^p L^C e^{-\frac{1}{C}R^\beta}. \end{aligned} \tag{4.13}$$

Using this estimate, we compute

$$\begin{aligned} \mathbb{E}[X_L^{2p}] &= L^{-2dp} \int \ldots \int_{\mathbb{R}^d} e^{-\frac{1}{L}\sum_{i=1}^{2p}|x_i|} \, \mathbb{E}\left[(F(x_1) - \mathbb{E}\left[F\right]) \ldots (F(x_{2p}) - \mathbb{E}\left[F\right])\right] dx_1 \ldots dx_{2p} \\ &\leq C^p L^C e^{-\frac{1}{C}R^\beta} + C^p L^{-2dp} \int \ldots \int_{\mathbb{R}^d} e^{-\frac{1}{L}\sum_{i=1}^{2p}|x_i|} \, \mathbb{1}_{\forall i, \, \exists j \neq i : \, |x_i - x_j| \leq R} \, dx_1 \ldots dx_{2p}. \end{aligned}$$

We consider the partitions $P := \{P_1, \ldots, P_{N_P}\}$ of the index set $[2p] := \{1, \ldots, 2p\}$ into nonempty subsets of cardinality ≥ 2 (that is, $\cup_j P_j = [2p]$, $\sharp P_j \geq 2$ for all j, and $P_j \cap P_l = \varnothing$ for all $j \neq l$), and we use the notation $P \vdash_2 [2p]$ for such partitions. The above then takes the form

$$\mathbb{E}[X_L^{2p}] \leq C^p L^C e^{-\frac{1}{C}R^\beta} + C^p L^{-2dp} \sum_{P \vdash_2 [2p]} L^{dN_P} R^{d(2p - N_P)}.$$

Since for all $1 \leq k \leq p$ the number of partitions $P \vdash_2 [2p]$ with $N_P = k$ is bounded by the Stirling number of the second kind $\left\{ {2p \atop k} \right\} \leq \frac{1}{2}\binom{2p}{k} k^{2p-k} \leq C^p p^{2p} k^{2(p-k)} (2p-k)^{-(2p-k)}$, we deduce

$$\mathbb{E}[X_L^{2p}] \leq C^p L^C e^{-\frac{1}{C}R^\beta} + C^p \Big(\frac{R}{L}\Big)^{dp} \sum_{k=1}^{p} \frac{p^{2p} k^{2(p-k)}}{(2p-k)^{2p-k}} \Big(\frac{R}{L}\Big)^{d(p-k)},$$

and hence by Markov's inequality, for all $\delta > 0$,

$$\mathbb{P}\left[X_L > \delta\right] \leq \delta^{-2p} C^p L^C e^{-\frac{1}{C}R^\beta} + \delta^{-2p} C^p \Big(\frac{R}{L}\Big)^{dp} \sum_{k=1}^{p} \frac{p^{2p} k^{2(p-k)}}{(2p-k)^{2p-k}} \Big(\frac{R}{L}\Big)^{d(p-k)}. \tag{4.14}$$

Recall that we may restrict to $0 < \delta \leq 1$. Choosing $R = L^\alpha$, $p = \delta^2 C_0^{-1} L^{\alpha\beta}$, and $\alpha = \frac{d}{d+\beta}$, for some $C_0 \simeq 1$ large enough, the estimate (4.14) above leads to

$$\mathbb{P}\left[X_L > \delta\right] \leq C e^{-\frac{1}{C}L^{\alpha\beta}} + \delta^{-2p} C^p L^{-\alpha\beta p} \sum_{k=1}^{p} \frac{p^{p+k} k^{2(p-k)}}{(2p-k)^{2p-k}}.$$

Noting that the summand is increasing in k, and using the choice of p with C_0 large enough, we deduce

$$\mathbb{P}\left[X_L > \delta\right] \le Ce^{-\frac{1}{C}L^{\alpha\beta}} + \delta^{-2p}C^pL^{-\alpha\beta p}p^p \le Ce^{-\frac{1}{C}\delta^2L^{\alpha\beta}}, \tag{4.15}$$

from which the desired result (4.12) follows.

Step 2. Approximately 1-local random variable $f(A)$.
For all $r > 1$, we define the (r-local) random variable $f_r(A) := \mathbb{E}\left[f(A) \,\|\, A|_{B_r}\right]$, and we set $F_r(x) := f_r(A(\cdot + x))$ and $X_{r,L} := L^{-d}\int_{\mathbb{R}^d} e^{-\frac{1}{L}|y|}(F_r(y) - \mathbb{E}\left[F_r\right])dy$. The approximate locality assumption (4.1) implies a.s. for all $r, L > 0$,

$$|X_{r,L} - X_L| \le Ce^{-\frac{r}{C}}. \tag{4.16}$$

Setting $\tilde{F}_r(x) := F(rx)$ and $A_r(x) := A(rx)$, we note that for all $x \in \mathbb{R}^d$ the random variable $\tilde{F}_r(x)$ is $\sigma(A|_{B_r(rx)})$-measurable, that is, $\sigma(A_r|_{B(x)})$-measurable. For all $r \ge 1$, the α-mixing assumption on A implies that the contracted random field A_r satisfies α-mixing with coefficient

$$\tilde{\alpha}_r(\ell, D; A_r) := \Big((1+rD)^C \exp(-\frac{1}{C}(r\ell)^\beta)\Big) \wedge 1 \;\le\; C(1+rD)^C \exp(-\frac{1}{C}(r\ell)^\beta),$$

so that the α-mixing coefficient for $r \ge 1$ can basically be estimated by the one for $r = 1$. We may therefore apply Step 1 in the following form for all $\delta, L > 0$ and all $r \ge 1$,

$$\mathbb{P}\left[X_{r,L} > \delta\right] = \mathbb{P}\left[\fint_{Q_{L/r}} (\tilde{F}_r - \mathbb{E}[\tilde{F}_r]) > \delta\right] \le C\exp\Big(-\frac{\delta^2}{C}\Big(\frac{L}{r}\Big)^{\frac{d\beta}{d+\beta}}\Big),$$

where the constant $C \ge 1$ is independent of r. Combining this with (4.16) and choosing $r := C|\log(\frac{\delta}{eC})| \ge 1$, we obtain for all $0 < \delta \le 1$ and $L > 0$,

$$\mathbb{P}\left[X_L > \delta\right] \le \mathbb{P}\Big[X_{r,L} > \delta - Ce^{-\frac{r}{C}}\Big] \le \mathbb{P}\left[X_{r,L} > \tfrac{\delta}{2}\right] \le C\exp\Big(-\frac{\delta^2}{C}\Big(\frac{L}{|\log(\frac{\delta}{eC})|}\Big)^{\frac{d\beta}{d+\beta}}\Big),$$

and the conclusion follows. □

Acknowledgements

The authors acknowledge financial support from the European Research Council under the European Community's Seventh Framework Programme (FP7/2014-2019 Grant Agreement QUANTHOM 335410).

References

[1] S. Aida and D. Stroock. Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. *Math. Res. Lett.*, 1(1):75–86, 1994.
[2] S. N. Armstrong and J.-C. Mourrat. Lipschitz regularity for elliptic equations with random coefficients. *Arch. Ration. Mech. Anal.*, 219(1):255–348, 2016.
[3] S. N. Armstrong and C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér. (4)*, 49(2):423–481, 2016.
[4] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.
[5] E. N. Barron, P. Cardaliaguet, and R. Jensen. Radon-Nikodym theorem in L^∞. *Appl. Math. Optim.*, 42(2):103–126, 2000.

[6] E. N. Barron, P. Cardaliaguet, and R. Jensen. Conditional essential suprema with applications. *Appl. Math. Optim.*, 48(3):229–253, 2003.
[7] S. Bobkov and M. Ledoux. Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, 107(3):383–400, 1997.
[8] P. Doukhan. *Mixing. Properties and examples*, volume 85 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1994.
[9] M. Duerinckx and A. Gloria. Weighted functional inequalities: Constructive approach. Preprint, 2017.
[10] M. Duerinckx and A. Gloria. Weighted second-order Poincaré inequalities: Application to the jamming limit. Preprint, 2017.
[11] J. Fischer and F. Otto. A higher-order large-scale regularity theory for random elliptic operators. *Comm. Partial Differential Equations*, 41(7):1108–1148, 2016.
[12] A. Gloria, S. Neukamm, and F. Otto. Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.*, 199(2):455–515, 2015.
[13] A. Gloria, S. Neukamm, and F. Otto. Quantitative stochastic homogenization for correlated fields. Preprint, arXiv:1409.2678v3, 2017.
[14] A. Gloria, S. Neukamm, and F. Otto. A regularity theory for random elliptic operators. Preprint, arXiv:1409.2678v3, 2017.
[15] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3):779–856, 2011.
[16] A. Gloria and F. Otto. An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, 22(1):1–28, 2012.
[17] A. Gloria and F. Otto. Quantitative estimates on the periodic approximation of the corrector in stochastic homogenization. In *CEMRACS 2013—modelling and simulation of complex systems: stochastic and deterministic approaches*, volume 48 of *ESAIM Proc. Surveys*, pages 80–97. EDP Sci., Les Ulis, 2015.
[18] A. Gloria and F. Otto. Quantitative results on the corrector equation in stochastic homogenization. *J. Eur. Math. Soc.*, 19:3489–3548, 2017.
[19] I. A. Ibragimov and Y. A. Rozanov. *Gaussian random processes*, volume 9 of *Applications of Mathematics*. Springer-Verlag, New York-Berlin, 1978. Translated from the Russian by A. B. Aries.
[20] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. Notes (Berlin, 1997).
[21] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.
[22] A. Naddaf and T. Spencer. Estimates on the variance of some homogenization problems. Preprint, 1998.
[23] M. Rosenblatt. A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U. S. A.*, 42:43–47, 1956.
[24] S. Torquato. *Random heterogeneous materials. Microstructure and macroscopic properties*, volume 16 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 2002.

(Mitia Duerinckx) Université Libre de Bruxelles (ULB), Brussels, Belgium
E-mail address: mduerinc@ulb.ac.be

(Antoine Gloria) Sorbonne Université, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France
Université Libre de Bruxelles (ULB), Brussels, Belgium
E-mail address: antoine.gloria@upmc.fr