Habilitation à diriger des recherches

Université de Lyon – CNRS
Ecole normale supérieure de Lyon
Unité de mathématiques pures et appliquées

EDP aléatoires : questions de régularité

Jean-Christophe Mourrat

Document de synthèse présenté le 6 février 2017

JURY

Alice Guionnet	E.N.S. Lyon	Examinatrice
Martin Hairer	University of Warwick	Rapporteur
Grégory Miermont	E.N.S. Lyon	Examinateur
Christophe Sabot	Université Lyon 1	Examinateur
Alain-Sol Sznitman	E.T.H. Zürich	Rapporteur
Lorenzo Zambotti	Université Paris 6	Examinateur
Ofer Zeitouni	Weizmann Institute	Rapporteur

Remerciements

En premier lieu, je suis très reconnaissant à Martin Hairer, Alain-Sol Sznitman et Ofer Zeitouni d'avoir accepté d'être les rapporteurs de mon habilitation. Merci d'avoir consacré une partie de votre temps précieux à ce travail. Je suis également très reconnaissant à Alice Guionnet, Grégory Miermont, Christophe Sabot et Lorenzo Zambotti d'avoir bien voulu faire partie du jury. Chacun d'entre vous est une source d'inspiration constante pour moi.

J'aimerais dire un immense merci à mes collaborateurs. Ce mémoire donne la plus grande place aux travaux récents réalisés en collaboration avec Scott Armstrong et Tuomo Kuusi d'une part, et Hendrik Weber d'autre part. Difficile de résumer en quelques mots toutes nos aventures ! Je remercie aussi chaleureusement mes collaborateurs lausannois Tom Mountford et Daniel Valesin. A vous tous, je voudrais dire combien j'ai aimé faire ce bout de chemin à vos côtés. Je pense également à mes autres collaborateurs bien sûr, qui m'ont chacun tant apporté.

Et bien sûr, amis et collègues de l'UMPA, merci à vous ! J'apprécie énormément l'ambiance du labo, à la fois mathématiquement stimulante et humainement si bienveillante. Je n'ose pas me lancer dans une liste des gens formidables de l'UMPA et de l'ICJ, ils sont si nombreux.

Merci aux amis non mathématiciens, et qui ne liront jamais ces lignes ! Merci aussi à mes parents pour leur soutien constant. Enfin, merci à Rosanne pour ces années passées ensemble.

Contents

CHAPTER I

Overview

The goal of this document is to survey my research from the completion of my Ph.D. thesis until March 2016. For convenience, the list of the corresponding research articles (including [Mou10, Mou11b, Mou11a] from my Ph.D. thesis) is displayed in a separate section of the bibliography, at the end of this document.

The present chapter aims at giving a bird's-eye view of my research activity. Besides presenting the results I obtained, I aim to put them in context and underline the links with other developments or questions. Each section presents a different aspect and can be read independently.

Chapters II and III highlight certain more specific aspects of my research, in relation with Sections 1 and 2 below. Although the presentation is still informal, the mathematics become more precise there.

1. Homogenization

1.1. Qualitative homogenization. Elliptic and parabolic partial differential equations of second order appear in a large variety of contexts. As an example, consider the diffusion of temperature in a medium. Denote by $u(t,x)$ the thermal energy density at time t and position x. This quantity is proportional to the temperature. In the absence of external sources of energy, we expect the conservation law

$$\partial_t u + \nabla \cdot \mathbf{j} = 0,$$

where $\mathbf{j}$ is the flux of thermal energy. Fourier's law predicts that this flux is proportional to the gradient of thermal energy: $\mathbf{j} = -\mathbf{a}\nabla u$, which leads to the parabolic equation

$$\partial_t u = \nabla \cdot (\mathbf{a}\nabla u). \tag{1.1}$$

Similarly, the steady state thermal energy profile can be recovered by solving the elliptic equation

$$\nabla \cdot (\mathbf{a}\nabla u) = 0, \tag{1.2}$$

with suitable boundary conditions.

The generality of this derivation makes it clear that the equations (1.1) and (1.2) are relevant in a large variety of contexts. Indeed, it only relies on the assumption of Fourier's law, which can be expected to hold in great generality, at least in the regime of small gradients. Examples include electromagnetism (Ohm's law), the diffusion of the concentration of a substance (Fick's law), or the linearization of Lagrange's equation for minimal surfaces. The diffusion of the concentration of molecules can be understood microscopically as the result of their Brownian movement. In reference to the context of electromagnetism, the matrix $\mathbf{a}$ is often called the conductivity.

We are interested in the situation where the medium is heterogeneous. Mathematically, this translates into the fact that the conductivity matrix $\mathbf{a}$ depends on the position within the medium. We assume however that the statistics of the heterogeneities are homogeneous in space. More precisely, we assume that the law

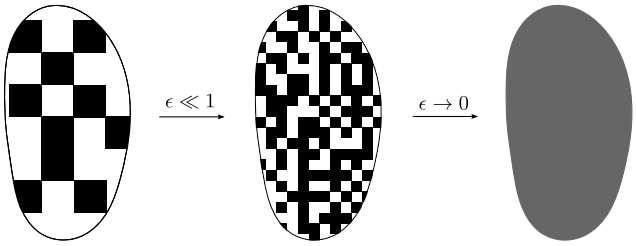

FIGURE 1.1. The first two drawings display realizations of the "random checkerboard model", where the conductivity is $\alpha\,\mathrm{Id}$ in the white region, and $\beta\,\mathrm{Id}$ in the black region. The squares are independently black or white with probability 1/2. In the limit $\varepsilon \to 0$, the material becomes equivalent to a homogeneous material of conductivity $\sqrt{\alpha\beta}\,\mathrm{Id}$.

of the mapping $\mathbb{R}^d \ni x \mapsto \mathbf{a}(x)$ is invariant under translations by any vector of $\mathbb{Z}^d$. We also ask that this law be ergodic under the action of these translations. Finally, we assume that $\mathbf{a}(x)$ is symmetric and uniformly elliptic; more precisely, that there exists a constant $\Lambda < \infty$ such that for every x,

$$\mathrm{Id} \le \mathbf{a}(x) \le \Lambda\,\mathrm{Id}. \tag{1.3}$$

We denote by $\mathbb{P}$ the law of the coefficient field $\mathbf{a}$, with associated expectation $\mathbb{E}$. We want to consider the situation where the correlation length of the coefficient field is much shorter than the typical length scale at which we want to describe the solution of an equation involving $\mathbf{a}$. In this case, it is reasonable to expect that the local fluctuations of the conductivity will be "averaged out" in the limit of large distances, resulting in an equivalent equation with constant, *homogenized* coefficients. One way to formulate this assumption of separation of scales is to introduce a small parameter $\varepsilon > 0$, and solve the family of problems

$$\begin{cases} -\nabla \cdot \left(\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right) \nabla u_\varepsilon\right) = 0 & \text{in } U, \\ u_\varepsilon = f & \text{on } \partial U, \end{cases} \tag{1.4}$$

where U is a given domain with Lipschitz boundary, and f is a fixed function with sufficient regularity. The statement of homogenization is that there exists a matrix $\bar{\mathbf{a}}$ which is constant in space, deterministic, and not depending on f or U, such that u_ε converges to $\overline{u}$ solution to

$$\begin{cases} -\nabla \cdot \bar{\mathbf{a}} \nabla \overline{u} = 0 & \text{in } U, \\ \overline{u} = f & \text{on } \partial U. \end{cases}$$

The matrix $\bar{\mathbf{a}}$ is usually called the homogenized or effective matrix. We may say that the solution operator associated with $-\nabla \cdot \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right) \nabla$ converges to that of $-\nabla \cdot \bar{\mathbf{a}} \nabla$. The result is not specific to the choice of Dirichlet problems.

The fact that we can asymptotically replace the rapidly oscillating coefficient field $\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)$ by a constant one $\bar{\mathbf{a}}$ is illustrated on Figure 1.1. This figure is so suggestive that it may lead us to believe that $\bar{\mathbf{a}}$ is simply the average of $\mathbf{a}$. This is however not the case. One way to realize this is to imagine a medium with long and thin vertical rods of very low conductivity, arranged randomly or periodically in the plane. The presence of these rods will have essentially no effect on a flux going in the vertical direction. On the other hand, it will have a dramatic slowdown effect on a flux going in the horizontal direction. This example shows that the computation of $\bar{\mathbf{a}}$ must incorporate geometric information about the random coefficient field; knowing

its one-point statistics is not sufficient. In fact, Figure 1.1 displays one of the very rare examples where the homogenized matrix is given by a simple formula.

The statement of homogenization was first proved under the assumption that $x \to \mathbf{a}(x)$ is periodic; we refer to [34] for a comprehensive study of homogenization under this assumption. Under the assumptions we stated on the coefficient field $\mathbf{a}(x)$, the result was obtained by Kozlov [147], Yurinskiĭ [200], and Papanicolaou and Varadhan [173].

This result of homogenization can be recast in terms of diffusions in random environment. Indeed, denote by $(X_t)_{t\geq 0}$ the diffusion associated with the infinitesimal generator $-\nabla\cdot\mathbf{a}\nabla$, with $\mathbf{E}_x^{\mathbf{a}}$ its expectation started from x, and let $f \in C_c^\infty(\mathbb{R}^d,\mathbb{R})$. The function

$$u_\varepsilon(t,x) := \mathbf{E}^{\mathbf{a}}_{\varepsilon^{-1}x}\left[f\left(\varepsilon X_{\varepsilon^{-2}t}\right)\right]$$

satisfies the equation

$$\begin{cases} \partial_t u_\varepsilon = \nabla\cdot\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u_\varepsilon & \text{in } \mathbb{R}_+\times\mathbb{R}^d,\\ u_\varepsilon(t=0,\cdot) = f. \end{cases} \tag{1.5}$$

As was shown in [173], one can deduce from the statement of elliptic homogenization stated above that this new function u_ε converges to $\overline{u}$ solution to

$$\begin{cases} \partial_t \overline{u} = \nabla\cdot\bar{\mathbf{a}}\nabla\overline{u} & \text{in } \mathbb{R}_+\times\mathbb{R}^d,\\ \overline{u}(t=0,\cdot) = f. \end{cases}$$

This can be rephrased in probabilistic terms as

$$\mathbf{E}^{\mathbf{a}}_{\varepsilon^{-1}x}\left[f\left(\varepsilon X_{\varepsilon^{-2}t}\right)\right] \xrightarrow[\varepsilon\to 0]{} \mathbf{E}_x\left[f\left(B_t\right)\right], \tag{1.6}$$

where under $\mathbf{E}_x$, the process B is a Brownian motion starting at x and with covariance matrix $2\bar{\mathbf{a}}$. One can then deduce the convergence in law of $\varepsilon X_{\varepsilon^{-2}\cdot}$ to B. The precise statement obtained in this way involves an averaging on the initial starting point of the diffusion, see [173].

A more direct and probabilistic approach to prove the convergence of the rescaled diffusion to Brownian motion was developped by Osada [170]. Putting more emphasis on the process of the environment viewed by the particle and using heat kernel upper bounds, Osada shows that $\varepsilon X_{\varepsilon^{-2}\cdot}$ converges in law under $\mathbf{P}_0^{\mathbf{a}}$ to B, for $\mathbb{P}$-almost every realization of the coefficient field $\mathbf{a}$. Such a statement is usually called a *quenched* central limit theorem. The weaker statement of convergence in law under $\mathbb{P}\mathbf{P}_0^{\mathbf{a}}$ is called an *annealed* central limit theorem.

The approach based on the environment viewed by the particle was extended to general reversible Markov processes by Kipnis and Varadhan [144] (see also [73, 74]). This general viewpoint covers at once a random walk evolving on a percolation cluster and a tagged particle in the symmetric exclusion process, among other examples. However, the result takes the form of an annealed central limit theorem.

There are several ways to prove the statement of homogenization. We will fisrt focus on a PDE-oriented approach, in the elliptic setting, with u_ε solution to (1.4), and then outline briefly a more probabilistic approach. A powerful heuristic consists in postulating that the function u_ε should be well-approximated by a function of the form

$$u_\varepsilon(x) \simeq \overline{u}(x) + \varepsilon v_1\left(x,\tfrac{x}{\varepsilon}\right) + \varepsilon^2 v_2\left(x,\tfrac{x}{\varepsilon}\right) + \cdots, \tag{1.7}$$

where the functions $v_1(x,y)$, $v_2(x,y)$, etc. are "reasonable". This ansatz is called a two-scale expansion. Under the assumption that the coefficient field is periodic, we would look for v_1 such that for each fixed x, the function $y \mapsto u_1(x,y)$ is periodic

(and similarly for v_2, etc.). It is indeed natural to expect a form of local limit for $u_\varepsilon - \overline{u}$, in the sense that for each fixed x, the function

$$y \mapsto \varepsilon^{-\alpha}(u_\varepsilon(x+\varepsilon y) - \overline{u}(x))$$

should have a well-defined limit which would capture the microscopic oscillations of u_ε, for a suitable choice of exponent $\alpha > 0$. Formally replacing u_ε by the development $\overline{u}(x) + \varepsilon^\alpha v_1\left(x, \frac{x}{\varepsilon}\right)$ in the equation $\nabla \cdot \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u_\varepsilon = 0$ then suggests the choice of exponent $\alpha = 1$. The form of the equation also makes it intuitive that the correction $v_1(x,y)$ should depend on x only through the quantity $\nabla\overline{u}(x)$. Moreover, this dependence should be linear. (All this can be checked by doing the explicit replacement of u_ε by the two-scale expansion ansatz.) In other words, on a scale intermediate between the microscopic scale ε and the unit scale around point x, the solution u_ε should be very close to the solution of a problem homogenizing to an affine function with slope $\nabla\overline{u}(x)$.

We therefore postpone for a moment the analysis of this two-scale expansion, and study this "affine" solution more precisely. Given $p \in \mathbb{R}^d$, we wish to find an $\mathbf{a}$-harmonic function which is as close as possible to the affine function $x \mapsto p \cdot x$. In other words, we wish to find a function $\phi(\cdot, p) : \mathbb{R}^d \to \mathbb{R}$ with slow growth at infinity and such that $x \mapsto p \cdot x + \phi(x,p)$ is $\mathbf{a}$-harmonic. This last condition can be rewritten as

$$-\nabla \cdot \mathbf{a}(p + \nabla\phi) = 0. \tag{1.8}$$

Here and below, we will often abuse notation and simply write ϕ instead of $\phi(\cdot, p)$. In the periodic setting, this equation can be solved with the further requirement that $x \mapsto \phi(x)$ be periodic and of zero mean. In the random setting we chose to work with, we can only ask for $\nabla\phi$ to be a $\mathbb{Z}^d$-stationary field with mean zero, in the sense that

$$\mathbb{E}\left[\int_{[0,1]^d} \nabla\phi\right] = 0. \tag{1.9}$$

The function ϕ is then only well-defined up to a constant, which we may fix by requiring $\phi(0) = 0$. (We will discuss later the possibility of constructing a stationary version of ϕ, in which case this condition must be lifted.) General arguments then ensure that ϕ is sublinear, in the sense that

$$r^{-1}\left(\fint_{B_r} |\phi|^2\right)^{\frac{1}{2}} \xrightarrow[r\to\infty]{\text{a.s.}} 0, \tag{1.10}$$

where B_r is the Euclidean ball of radius r centered at the origin, and $\fint_{B_r}$ is the normalized integral $|B_r|^{-1}\int_{B_r}$. The function ϕ is called the corrector (in the direction of p). Notice that the mapping $p \mapsto \phi(\cdot, p)$ is linear.

Equipped with these correctors, we can now come back to the two-scale expansion (1.7) and postulate

$$u_\varepsilon(x) \simeq \overline{u}(x) + \varepsilon\phi\left(\tfrac{x}{\varepsilon}, \nabla\overline{u}(x)\right) = \overline{u}(x) + \varepsilon\sum_{i=1}^d \partial_{x_i}\overline{u}(x)\,\phi^{(i)}\left(\tfrac{x}{\varepsilon}\right), \tag{1.11}$$

where $(\phi^{(1)}, \dots, \phi^{(d)})$ denote the correctors in the directions of the canonical basis of $\mathbb{R}^d$. This approximation is indeed plausible since it suggests

$$\nabla u(x) \simeq \nabla\overline{u}(x) + \nabla\phi\left(\tfrac{x}{\varepsilon}, \nabla\overline{u}(x)\right),$$

in agreement with the idea that the gradient of the solution is close, on a mesoscopic scale, to $p + \nabla\phi(\cdot, p)$, with a slowly varying $p = \nabla\overline{u}(x)$. The approach to homogenization exposed for instance in [173] consists in defining the difference

$$z(x) := u_\varepsilon(x) - \overline{u}(x) - \varepsilon\phi\left(\tfrac{x}{\varepsilon}, \nabla\overline{u}(x)\right), \tag{1.12}$$

and leveraging on the cancellations appearing in the computation of $-\nabla\cdot\mathbf{a}\nabla z$ to infer that z converges to 0 in $H^1(U)$. The computation reveals that

$$\bar{\mathbf{a}}p = \mathbb{E}\left[\int_{[0,1]^d}\mathbf{a}(p+\nabla\phi(\cdot,p))\right]. \tag{1.13}$$

This relation is very natural, since the right side is the average flux of the $\mathbf{a}$-adapted "affine" function with slope p, and the left side is the flux of the affine function with slope p for the homogenized limit. We also have the energy identity

$$\frac{1}{2}p\cdot\bar{\mathbf{a}}p = \mathbb{E}\left[\int_{[0,1]^d}\frac{1}{2}(p+\nabla\phi(\cdot,p))\cdot\mathbf{a}(p+\nabla\phi(\cdot,p))\right], \tag{1.14}$$

to which we will return in Subsection 1.3. From the convergence of z to 0 in $H^1(U)$, we can deduce that

$$\|u_\varepsilon-\overline{u}\|_{L^2(U)}\xrightarrow[\varepsilon\to0]{}0,$$

as well as the weak convergences in $L^2(U)$:

$$\nabla u_\varepsilon\rightharpoonup\nabla\overline{u}\qquad\text{and}\qquad\mathbf{a}\nabla u_\varepsilon\rightharpoonup\bar{\mathbf{a}}\nabla\overline{u}. \tag{1.15}$$

This weak convergence cannot be improved to a strong convergence, since the convergence to 0 of z in $H^1(U)$ can be rephrased as

$$\left\|\nabla u_\varepsilon-\nabla\overline{u}-\sum_{i=1}^d\partial_{x_i}\overline{u}(x)\,\nabla\phi^{(i)}\left(\tfrac{x}{\varepsilon}\right)\right\|_{L^2(U)}\xrightarrow[\varepsilon\to0]{}0. \tag{1.16}$$

The probabilistic view on homogenization is slightly different, although the correctors also play a central role. In this view, the corrector provides us with harmonic coordinates enabling to turn the diffusion into a martingale. More precisely, the condition that $x\mapsto p\cdot x+\phi(x,p)$ be $\mathbf{a}$-harmonic can be rephrased as the fact that $t\mapsto p\cdot X_t+\phi(X_t,p)$ is a martingale. We thus deduce a decomposition of $p\cdot X_t$ as

$$p\cdot X_t=(p\cdot X_t+\phi(X_t,p))-\phi(X_t,p),$$

where the term between parentheses is a martingale, and the remainder $\phi(X_t,p)$ is hopefully of lower order. One can then show that the martingale rescales to a Brownian motion. This consists in checking that the associated quadratic variation grows asymptotically linearly. This can be obtained as a consequence of the ergodicity of the process of the environment seen by the particle. In this way, one obtains a quenched central limit theorem for this martingale part. There remains to show the asymptotic smallness of the remainder $\phi(X_t,p)$. At this stage, we only have the relatively weak information (1.10) on the sublinearity of the corrector. This limits us a priori to an annealed control of this remainder, resulting in an annealed central limit theorem for X_t itself.

The general problem driving my research in the area is to make the statement of homogenization *quantitative*, in any of the essentially equivalent forms outlined above. This requires to strengthen the mixing assumption on the coefficient field. For simplicity, we will assume that the coefficient field has a finite range of dependence (and sometimes make the stronger assumption that it can be written as a local function of a field of i.i.d. random variables).

Until recently, little was known about this problem. One notable exception is the result of Yurinskiĭ [201] giving a rate of convergence of u_ε to $\overline{u}$ in the elliptic setting, in dimension $d\geq3$. The rate of convergence is ε^α, but the exponent $\alpha>0$ is not explicit.

A related question concerns efficient approximations of the homogenized conductivity $\bar{\mathbf{a}}$. The formulas (1.13) or (1.14) are difficult to use for practical purposes, since evaluating the right sides of these identities requires to solve the corrector

equation on the full space, and then compute an expectation involving this corrector. One can replace the expectation on the right side of (1.13) by a space integral

$$\fint_{B_r} \mathbf{a}(p + \nabla\phi(\cdot, p)),$$

which is asymptotically justified in the limit of large r, by the ergodic theorem. We can also replace the corrector itself by a finite-volume approximation (e.g. we approximate the function $x \mapsto p \cdot x + \phi(x)$ by the solution to the Dirichlet problem on a very large box, with the affine boundary condition $x \mapsto p \cdot x$.) Bourgeat and Piatnitski [45] showed that under suitable mixing assumptions, such approximate schemes yield an error of ε^α, but again the exponent $\alpha > 0$ is not explicit.

From the probabilistic perspective, the question is phrased in terms of quantitative central limit theorems for the diffusion in random environment. Over the past ten years, an intense research has focused on the related but slightly different question of obtaining quenched central limit theorems under the most general possible conditions. This research has focused on the discrete counterpart of the diffusion process, which is often called the random walk among random conductances. As was explained above, the difficulty consists in showing that $t^{-\frac{1}{2}}\phi(X_t, p)$ converges to 0 in $\mathbf{P}_0^{\mathbf{a}}$ probability, for almost every realization of the coefficient field $\mathbf{a}$. One possibility would be to try to improve on the sublinear growth estimate (1.10). The route chosen instead consists in showing that the law of X_t is sufficiently spread out that the bound (1.10) is sufficient to conclude. This relies on heat kernel upper estimates. Proofs of these heat kernel bounds for the random walk on supercritical percolation clusters were obtained by Mathieu and Remy [158] and Barlow [21]. The proof of quenched central limit theorem for this walk was obtained by Sidoravicius and Sznitman [185] in dimension $d \geq 4$, and then in every dimension by Mathieu and Piatnitski [157] and Berger and Biskup [35]. The arguments were then extended in a series of works [41, 156, 13] to cover arbitrary i.i.d. conductance fields, provided that the set of non-zero conductances defines a supercritical percolation cluster.

1.2. Quantitative homogenization – the nonlinear approach. The principal difficulty in quantifying stochastic homogenization is that solutions are nonlinear, nonlocal functions of the coefficient field. In other words, there is no simple method to transfer the strong mixing properties we assume on the coefficient field to the solutions themselves.

One approach to overcome this difficulty is to use concentration inequalities that hold for possibly nonlinear functionals. The prototypical example is the Efron-Stein inequality. This inequality states that if $X = (X_1, \ldots, X_n)$ are independent random variables, and if $(X_1', \ldots, X_n')$ is an independent copy of X, then for every function F,

$$\mathbb{V}\mathrm{ar}\,(F(X)) \leq \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}\Big[\big(F(X_1, \ldots, X_{i-1}, X_i', X_{i+1}, \ldots, X_n) - F(X)\big)^2\Big]. \tag{1.17}$$

One of the early applications of this inequality in the context of quantitative homogenization is due to Kesten [143], who obtained his celebrated variance upper bound for first passage percolation using this inequality.

In my Ph.D. thesis, I got interested in showing quantitative estimates on the rate of convergence to equilibrium of the process of the environment seen by the particle. This question was motivated by the problem of identifying the scaling limit of a toy model for aging, see Section 5. I used the Efron-Stein inequality to implement a Nash-type strategy [149, 37] and obtain a sub-optimal, but explicit polynomial rate of convergence for this process [Mou11b]. From a probabilistic point of view and in the discrete setting, applying the Efron-Stein inequality asks for comparing expectations of functionals for two environments differing only on

a single edge. The main reason for the suboptimality of the estimate I obtained is that from this point of view, there is not much to do to control this difference except to use a crude bound involving the probability that the walk touches the edge where the two environments differ.

Simultaneously and independently, Gloria and Otto [115, 116] considered the problem of estimating the sharpness of finite-volume approximations of the right side of (1.14) for the computation of the homogenized matrix. In a more PDE-based approach, they also relied on the Efron-Stein inequality to control the variance of quantities involving the corrector, or a finite-volume approximation thereof. The interesting aspect of the PDE perspective is that one can write an equation for the quantities appearing on the right side of (1.17). Indeed, the term inside the square on the right side of (1.17) is essentially a derivative with respect to the value of the conductance on the edge under consideration, and we can write an equation for derivatives of quantities of interest. We illustrate this idea by doing a formal computation for the derivative of the corrector with respect to the value of a conductance e. Our notation is relatively ambiguous at this point, because it is easier to think of this operation in the discrete setting, although we will keep using continuous-space notation. We differentiate the equation (1.8) for the corrector to get

$$-\nabla \cdot \mathbf{a} \nabla (\partial_e \phi) = \nabla \cdot \mathbf{1}_e (p + \nabla \phi).$$

In the expression above, we write $\partial_e \phi$ to denote the derivative of ϕ with respect to the value of the conductance on edge e. Using the Green representation formula, we obtain

$$\begin{aligned} \partial_e \phi(x) &= \sum_y G(x,y) \left(\nabla \cdot \mathbf{1}_e (p + \nabla \phi) \right)(y) \\ &= -\nabla G(x,e)(p + \nabla \phi(e)), \end{aligned} \tag{1.18}$$

where $G(x,y)$ is the Green function associated with $-\nabla \cdot \mathbf{a} \nabla$, and we used a summation by parts in the last step. The heart of the work is then to prove sufficiently good estimates on the Green function itself. The technique is inspired by previous work of Naddaf and Spencer [166], who implemented a comparable strategy under the additional assumption that the conductivity matrix is uniformly sufficiently close to the identity (the regime of "small ellipticity contrast"). This work was in turn influenced by their earlier, fundamental contribution to the understanding of the so-called $\nabla \phi$ interface model [165], to which we will return.

By this method, Gloria and Otto could obtain optimal estimates on the quantities they chose to focus on. This was therefore a major milestone in our understanding of quantitative homogenization. In an updated version of [Mou11b], I could show that the questions I considered of convergence to equilibrium of the environment seen by the particle, and their questions of numerical approximation of the homogenized coefficients, are very closely related. In fact, they are equivalent in small dimension, and otherwise the "parabolic" statement concerning the environment seen by the particle is stronger [Mou11b, Section 9]. The realization of this correspondence led to the development of new methods of numerical approximation of the homogenized coefficients [GM12], which are based on Richardson extrapolations. We could identify precisely the order up to which Richardson extrapolations really improve the rate of convergence. Partially inspired by these observations, Gloria, Neukamm and Otto took up the study of the rate of convergence of the process of the environment viewed by the particle. In [113], they obtained optimal convergence rates for this process, in the sense of polynomial moments. As can be glimpsed from the formula (1.18), a crucial ingredient of the proof rests in proving sufficiently strong estimates

on the gradient of the Green function. The estimates proved there are phrased in terms of spatial averages of high polynomial moments of the gradient.

Estimates on gradients of the Green function were subsequently refined by Marahrens and Otto [153], who showed the following. Recall that we denote by $G(x,y)$ the elliptic Green function. For every $p<\infty$, there exists a constant $C<\infty$ such that for every $x,y\in\mathbb{Z}^d$,

$$\mathbb{E}[|\nabla G(x,y)|^p]^{\frac{1}{p}}\le\frac{C}{|y-x|^{d-1}}, \tag{1.19}$$

$$\mathbb{E}[|\nabla\nabla G(x,y)|^p]^{\frac{1}{p}}\le\frac{C}{|y-x|^{d}}, \tag{1.20}$$

where in the second line, the quantity $\nabla\nabla G(x,y)$ denotes the mixed second derivative of G. In other words, one ∇ acts on the first variable, and the other on the second variable. The estimates match the behavior of the constant-coefficient case, and extend the results of Delmotte and Deuschel [75], who showed (1.19) for $p=2$ and (1.20) for $p=1$ (these weaker estimates were proved without any mixing assumption on the coefficients; see also [31] for a different perspective on this problem). In the work [153], the Efron-Stein inequality is replaced with a stronger log-Sobolev inequality. Strictly speaking, the estimates (1.19) and (1.20) are only true in the discrete setting, due to the possible local irregularity of the coefficient field (which should be irrelevant to the large-scale question of homogenization). The result was extended to continuous equations in [111], with local averages of the gradients replacing these pointwise bounds.

With such estimates, one can prove optimal bounds on essentially every quantity of interest, with a probabilistic control in the sense of polynomial moments. In particular, one can show that in dimension $d\ge 3$, the corrector can be constructed as a stationary field, and has bounded polynomial moments. In dimension 2, the corrector grows at most logarithmically, in the sense of polynomial moments. For the elliptic problem

$$-\nabla\cdot\mathbf{a}\left(\tfrac{\cdot}{\varepsilon}\right)\nabla u_\varepsilon=f\qquad\text{in }\mathbb{R}^d, \tag{1.21}$$

where $f\in C_c^\infty(\mathbb{R}^d,\mathbb{R})$, we have the estimate

$$\mathbb{E}\left[\|u_\varepsilon-\overline{u}\|_{L^q(\mathbb{R}^d)}^p\right]^{\frac{1}{p}}\le C\varepsilon, \tag{1.22}$$

in dimension $d\ge 3$, where the function $\overline{u}$ is the solution of the corresponding homogenized problem. In dimension $d=2$, assuming further that f is of zero mean, the estimate (1.22) holds up to the multiplication by $\log^{\frac{1}{2}}\varepsilon^{-1}$.

The important difference between the problem (1.21) and that displayed in (1.4) is that (1.21) is posed in the full space, while (1.4) is posed in a domain. Solutions of problems such as (1.4) display a degraded rate of convergence within a layer close to the boundary of the domain, and the estimate corresponding to (1.22) breaks down. In the random setting, these boundary layers are not well-understood.

I will now discussion those of my contributions that relate to this work, separating them into three groups.

First, in parallel to these developments, I explored the possibility to use the results and methods of the early papers [115, 116] of Gloria and Otto to pursue a probabilistic approach to stochastic homogenization. The approach rests on the development and application of quantitative martingale convergence theorem. I initially focused on deriving annealed quantitative central limit theorems, partly because it is simpler technically, and partly because it gives access to the "systematic error" $\mathbb{E}[u_\varepsilon(x)]-\overline{u}(x)$, while the "statistical error" measured by the variance $\mathbb{V}\mathrm{ar}(u_\varepsilon(x))$ is more easily accessible via an application of the Efron-Stein inequality.

After a first partial result [Mou12c], (the cause of the limitation is partly explained in [Mou13]), I showed in [Mou14] that for the solution $u_\varepsilon(t,x)$ to the (discrete version of the) parabolic problem (1.5), we have, for each fixed $t>0$, $x\in\mathbb{R}^d$ and $\delta>0$, the existence of a constant $C<\infty$ such that for every $\varepsilon\in(0,1]$,

$$|\mathbb{E}[u_\varepsilon(t,x)]-\overline{u}(t,x)|\le C\varepsilon^{1-\delta},$$

for arbitrary $\delta>0$, in dimension $d\ge 3$. In agreement with the discussion surrounding (1.5)-(1.6), this estimate is obtained through a quantitative annealed central limit theorem for the random walk in random conductances. The proof of a similar result, based on the method of two-scale expansion, then appeared in [112]. Together with Yu Gu, we pursued the probabilistic approach further in [GM14] and showed that in dimension $d\ge 3$, for each fixed $t>0$ and $x\in\mathbb{R}^d$,

$$u_\varepsilon(t,x)=\overline{u}(t,x)+\varepsilon\sum_{i=1}^d \partial_{x_i}\overline{u}(t,x)\,\phi^{(i)}\left(\tfrac{x}{\varepsilon}\right)+o(\varepsilon), \tag{1.23}$$

where $o(\varepsilon)$ is a quantity satisfying $\varepsilon^{-1}\mathbb{E}[|o(\varepsilon)|]\to 0$ as $\varepsilon\to 0$. Despite the heuristic of the two-scale expansion, this actually came as a surprise. The two-scale expansion, as quantified in [112], shows that the left side of (1.16) is bounded by $C\varepsilon$, from which we can only infer that

$$\|u_\varepsilon-\overline{u}\|_{L^2(U)}\le C\varepsilon.$$

A signature of the application of the probabilistic method is that the estimate (1.23) holds pointwise, or in other words, for any fixed pair (t,x). Exploring higher-order two-scale expansions, Yu Gu [119] later uncovered a cancellation which enabled him to prove in the elliptic setting that $\|u_\varepsilon-\overline{u}-\varepsilon\sum_{i=1}^d\partial_{x_i}\overline{u}\,\phi^{(i)}\left(\frac{\cdot}{\varepsilon}\right)\|_{L^2}=o(\varepsilon)$.

Second, the following question caught my interest. As was already said, it is computationally expensive to compute the homogenized coefficients of a given random field. However, Anantharaman and Le Bris [9, 10, 11] observed that if the medium is a small random perturbation of a homogeneous medium, then there are much more efficient methods to compute the homogenized matrix. This observation follows a long tradition of perturbative formulas for effective parameters of homogeneous materials perturbed by a small periodic or random structure, which date back to the 19th century and go under the names of Clausius-Mossotti, Lorentz-Lorenz, Maxwell, or Rayleigh formulas. The precise setting usually referred to by physicists is the following. From a homogeneous material of conductivity $\alpha\,\mathrm{Id}$, add unit-ball inclusions of conductivity $\beta\,\mathrm{Id}$, with density $p\ll 1$, in a stationary manner. Then the effective conductivity of the resulting material is

$$\left(\alpha+\frac{\alpha d(\beta-\alpha)}{\beta+\alpha(d-1)}p+o(p)\right)\mathrm{Id}. \tag{1.24}$$

This formula was proved by Rayleigh [192] assuming that the balls are arranged along a periodic lattice.

The formula (1.24) rests on the assumption that the inclusions are spherical. One aspect of the contribution of Anantharaman and Le Bris is to observe that for more general inclusions, although no closed analytic formula is expected to exist, we can think of efficient numerical strategies allowing to compute the homogenized matrix. The idea is to write a finite-volume periodic approximation of the homogenized matrix, and formally interchange the limits of large volume and small density of inclusions. By doing so, we obtain the first-order correction by computing a corrector-type quantity in a periodic box containing exactly one inclusion. This computation can be performed efficiently, and should indicate the first-order perturbation to the homogenized coefficients.

The Clausius-Mossotti formula was first proved by Almog [7] in dimension $d = 3$, by methods specific to this dimension. In [Mou15a], I justified and extended the picture proposed by Anantharaman and Le Bris, in every dimension and in a discrete setting. The adaptation of the arguments to the continuous setting yields a proof of the Clausius-Mossotti formula in every dimension as a particular case. The result was later generalized to cover general ergodic random fields and expansions of arbitrarily high order by Duerinckx and Gloria [83].

My interest in this problem was also motivated by the question of the behavior of the effective diffusivity of near-critical percolation. More precisely, denote by $\bar{\mathbf{a}}(p) = \bar{a}(p)\,\mathrm{Id}$ the homogenized matrix associated with the two-dimensional square bond percolation (say) of parameter p. As is well-known, the critical percolation p_c is equal to $1/2$ in this case. Does there exist an exponent σ such that $\bar{a}(p) \simeq (p-p_c)^\sigma$ as $p \to p_c^+$? Can we compute σ? Denote by $\bar{a}(n, p_c)$ a finite-volume approximation of $\bar{a}(p_c) = 0$ in a box of size n (for instance, $\bar{a}(n, p_c)$ is the expectation of the effective conductivity between two faces of the box of size n). Does there exist an exponent τ such that $\bar{a}(n, p_c) \simeq n^{-\tau}$? Can the exponents σ and τ be computed? Can we relate σ and τ through the correlation length, in the spirit of the Kesten relations [142]?

Understanding the problem posed by Anantharaman and Le Bris seemed to be a modest but necessary first step towards a better understanding of these problems. However, despite the immense success of the recent work on two-dimensional critical percolation and its critical exponents, it is not at all clear that the exponents σ and τ have a simple expression (although bounds on τ are known, see [141, 71]). Physicists have not come up with such a prediction, despite intensive numerical simulation (see [135] for a review). As an illustration of this possibility, the monochromatic two-arm exponent is not known to have a simple expression. In ongoing work with Christophe Garban, we show that this exponent can be expressed as the principal eigenvalue of a (rather singular) differential operator (see `http://goo.gl/n0hnqf` for a presentation of our contribution).

The third and final set of problems I wish to discuss here is the following. In a sense, the statement of homogenization is similar to a law of large numbers. (This sentence may be confusing, since we saw that homogenization corresponds to a central limit theorem for the diffusion in random environment. Nevertheless, it is fair to say that from a PDE perspective, the statement is a form of law of large numbers. Moreover, even from the probabilistic perspective, the proof of the central limit theorem is ultimately reduced to an application of the ergodic theorem for the environment viewed by the particle; this may be taken as an indication that we are indeed proving a form of law of large numbers after all.) Once sharp estimates are available for the rate of convergence in this law of large numbers, a natural question is to ask if we can actually identify the next-order correction to homogenization. In other words, we look for a CLT-type result for homogenization.

Before proceeding to describe what can be proved in this direction, I wish to explain why such results can be interesting from a practical perspective. To begin with, homogenization per se is appealing because it enables to replace a computationally very expensive problem with rapidly oscillating coefficients by a much simpler, homogeneous problem, described by a few effective parameters. In a sense we will make precise shortly, the same is true of the next-order, fluctuating correction. This fluctuating field is a Gaussian field whose covariance can be described by a new (finite) set of effective parameters, which we may call the *effective fluctuation tensor*.

In this spirit, practitioners have long been interested in the problem of recovering the statistics of the original random field from the observation of the large-scale fluctuations of solutions. However, this problem is ill-defined. The new point of

view outlined above suggests instead to focus on trying to determine this effective fluctuation tensor. This can be done by observing the large-scale fluctuation field of any homogenization problem. Once this effective fluctuation tensor is determined, we gain a powerful approximation of the law of the solution to any other homogenization problem.

In line with the qualitative theory outlined in Subsection 1.1, it is reasonable to focus first on understanding the large-scale behavior of the corrector. We already pointed out the connection between stochastic homogenization and the so-called $\nabla\phi$ interface model of statistical mechanics in relation with the work of Naddaf and Spencer [165, 166] (see also the introduction to [GM16b]). In [165], Naddaf and Spencer identified the scaling limit of this model to be a Gaussian free field. It is therefore natural to conjecture that the same should be true of the corrector (see [35, Conjecture 5]). This is however not correct, although the idea suggests the correct scaling exponent. We will give more precise heuristics and description of the result in Subsection 1.3 below.

An important idea of Naddaf and Spencer is to make use of a representation of correlations due to Helffer and Sjöstrand [134, 186]. This representation is a refinement of the Efron-Stein inequality. It is easiest to describe the inequality in the case of a given family $X = (X_1, \dots, X_N)$ of independent standard Gaussian random variables. In this context, a variant of the Efron-Stein inequality states that for every $F = F(X)$,

$$\mathbb{V}\mathrm{ar}\,(F) \le \sum_{i=1}^{n} \mathbb{E}\Big[(\partial_{x_i} F)^2\Big].$$

The Helffer-Sjöstrand representation of correlation refines this variance estimate into

$$\mathbb{V}\mathrm{ar}\,(F(X)) = \sum_{i=1}^{n} \mathbb{E}\Big[(\partial_{x_i} F)(\mathscr{L}+1)^{-1}(\partial_{x_i} F)\Big], \tag{1.25}$$

where $\mathscr{L}$ is the Laplacian-type operator on the product space

$$\mathscr{L} = \partial^*\partial = \sum_{i=1}^{n} (-\partial_{x_i} + X_i)\partial_{x_i}.$$

(The "divergence" ∂^* is the adjoint to the "gradient" $\partial = (\partial_{x_i})_i$ with respect to the Gaussian measure, hence the additional term X_i above.) Since the formula (1.25) is an identity, it can be polarized into an identity for correlations. In the context of the homogenization of discrete equations, let us assume for simplicity that the random conductances are (nice) functions of standard Gaussian random variables. This is not really necessary for the statement of the Helffer-Sjöstrand representation, but simplifies the application of the chain rule (see [MO14, Remark 2.3]). Under this assumption and in dimension $d \ge 3$, Felix Otto and I [MO14] used this representation of correlations, together with techniques as exposed above, to obtain a description of the asymptotic behavior of the correlations of the corrector $\mathbb{E}[\phi(0)\,\phi(x)]$. Completing the description of the scaling limit of the corrector, Jim Nolen and I proved the asymptotically Gaussian behavior of spatial averages of the corrector in [MN15].

Together with Yu Gu [GM15], we then described the fluctuations of solutions to (1.21). Surprisingly, we showed under similar assumptions that the naive two-scale expansion of u_ε in terms of the correctors does *not* suggest the correct scaling limit for these fluctuations. We also postpone a more precise description of the results obtained there to the next subsection.

1.3. Quantitative homogenization – the additive approach. More recently, I took part in the construction of a new approach to quantitative homogenization [14], [AM16, AKM15, AKM16a, AKM16b]. As was said in the beginning

of Subsection 1.2, in order to make stochastic homogenization quantitative, we need to transfer the assumed mixing properties of the coefficient field into information on the solutions themselves. This is difficult a priori because solutions are nonlinear, nonlocal functions of the coefficients.

The core idea of the new approach we developed is that energy quantities *are* in fact local, additive functions of the coefficient field. We should therefore focus on studying these quantities first, and derive properties of solutions as consequences.

This new approach is interesting for a number of reasons. First, it allows to widen the range of coefficient fields covered by the theory. Indeed, the "nonlinear" approach outlined in the previous subsection, based on the Efron-Stein inequality or its variants, essentially requires that we assume the coefficient field to be a function of a family of independent random variables. This can be weakened in several ways, but as an example, the nonlinear approach cannot be applied to a coefficient field which is only assumed to be of finite range of dependence. Recall that a random field $\mathbf{a}$ is said to be of unit range of dependence if

$$U, V \subseteq \mathbb{R}^d,\ \operatorname{dist}(U,V) \geq 1 \implies (\mathbf{a}(x))_{x\in U} \text{ and } (\mathbf{a}(x))_{x\in V} \text{ are independent.} \tag{1.26}$$

Moreover, the new approach allows to strengthen the stochastic control of the random variables of interest (we are able to show that the fluctuations have Gaussian tails). In my view, the most important point however is to uncover the *additive structure* of stochastic homogenization. By exposing this additive structure, we ultimately reduce the problem to a question involving sums of random variables satisfying the same mixing properties as the underlying coefficient field, up to an error we can neglect. Stronger results naturally follow from the uncovering of this simple structure.

The fact that homogenization can be related to the convergence of energy quantities was first exposed by Dal Maso and Modica [69, 70] (building on previous work of Marcellini [154] in the periodic setting). For a bounded domain $U \subseteq \mathbb{R}^d$ (with Lipschitz boundary) and a vector $p \in \mathbb{R}^d$, denote

$$\nu(U,p) := \inf_{v\in H^1_0(U)} \fint_U \frac{1}{2}(p+\nabla v)\cdot \mathbf{a}(p+\nabla v), \tag{1.27}$$

where $H^1_0(U)$ denotes the closure in $H^1(U)$ of the set of smooth functions with compact support in U. One can check that there exists a unique minimizer for this variational problem, which we denote it by $v(\cdot,U,p)$ (see e.g. [89, Subsection III.8.2]). The Euler-Lagrange equation associated with this variational problem reads

$$\forall \psi \in H^1_0(U), \quad \nabla\psi \cdot \mathbf{a}(p+\nabla v(\cdot,U,p)) = 0. \tag{1.28}$$

This is the weak formulation of the equation (1.8), with null Dirichlet boundary condition on ∂U. In other words, the minimizer $v(\cdot,U,p)$ is a finite-volume approximation of the corrector $\phi(\cdot,p)$.

The quantity $\nu(\cdot,p)$ is subadditive, in the sense that if $U_1,\dots,U_k \subseteq U$ are pairwise disjoint domains such that $|U \setminus \bigcup_i U_i| = 0$ (here $|\cdot|$ denotes the Lebesgue measure), then

$$\nu(U,p) \leq \sum_{i=1}^k \frac{|U_i|}{|U|}\nu(U_i,p). \tag{1.29}$$

Indeed, this holds since we can glue the U_i minimizers together and create a candidate minimizer for $\nu(U,p)$. Denoting by

$$\square_r := \left(-\frac{r}{2},\frac{r}{2}\right)^d, \tag{1.30}$$

it follows from the subadditive ergodic theorem that $\nu(\square_r, p)$ converges almost surely and in L^1 to a deterministic constant as r tends to infinity. Moreover, it is straightforward to check from the characterization (1.28) that the mapping $p \mapsto v(\cdot, U, p)$ is linear, and therefore $p \mapsto \nu(U, p)$ is a quadratic form. This property is preserved by passing to the limit. We deduce that there exists a constant matrix $\bar{\mathbf{a}}$ such that for every $p \in \mathbb{R}^d$,

$$\nu(\square_r, p) \xrightarrow[R\to\infty]{\text{a.s.}} \frac{1}{2} p \cdot \bar{\mathbf{a}} p. \tag{1.31}$$

Our choice of notation implies that the matrix thus defined is the homogenized matrix. Intuitively, this can be understood from the fact that $\frac{1}{2} p \cdot \bar{\mathbf{a}} p$ is the result of the minimization of the functional with constant coefficient $v \mapsto \fint_U \frac{1}{2}(p + \nabla v) \cdot \bar{\mathbf{a}}(p + \nabla v)$ in $H^1_0(U)$, which selects $\bar{\mathbf{a}}$-harmonic functions as minimizers. It roughly matches with the relation (1.14). In a loose sense, Dal Maso and Modica [69, 70] proved that indeed, the convergence (1.31) implies the statement of homogenization as presented in Subsection 1.1 (their precise result is phrased in terms of the Γ-convergence of the energy functionals).

This result already offers a glimpse of the additive structure we wish to uncover. Indeed, the quantity $\nu(\cdot, p)$ is subadditive, and by definition, it is also a local function of the coefficients. Moreover, the result of Dal Maso and Modica strongly suggests to look for a quantification of the statement of homogenization via a quantification of the convergence in (1.31). Roughly speaking, our main task is therefore to show that $\nu(\cdot, p)$ becomes asymptotically additive. If this can be done, then we expect the next order correction to (1.31) to be governed by a central limit theorem.

This idea is however too naive, because the minimizer of $\nu(\square_r, p)$ has a boundary layer, which implies a next-order correction to (1.31) driven by the behavior of the minimizer close to the boundary. This boundary behavior is irrelevant to the properties of solutions we wish to study. In the remainder of this section, we will ignore this problem and make the false assumption that indeed, the next-order correction to (1.31) is governed by a central limit theorem. We will use this simplifying assumption to describe and explain the structure of the scaling limits of the corrector and the fluctuations of general solutions. We will then return to the problem in Chapter II, explain what can be proved about the quantity $\nu(\cdot, p)$, and how to overcome the difficulty caused by boundary layers.

Before we proceed to give heuristics for the scaling limits of the corrector and the fluctuations of solutions, we introduce the object that will arise as a scaling limit [1]. We say that $W = (W_1, \dots, W_d)$ is a *vector white noise* with covariance matrix Q if for every $f = (f_1, \dots, f_d) \in C^\infty_c(\mathbb{R}^d; \mathbb{R}^d)$, the random variable

$$W(f) := W_1(f_1) + \cdots + W_d(f_d)$$

is a centered Gaussian with variance $\int_{\mathbb{R}^d} f \cdot \mathsf{Q} f$. The set of admissible test functions can be extended to $f \in L^2(\mathbb{R}^d; \mathbb{R}^d)$ by density. We say that $W = [W_{ij}]_{1 \le i,j \le d}$ is a matrix white noise if the vector of its entries is a vector white noise. Given a vector white noise W and a positive-definite symmetric matrix $\bar{\mathbf{a}}$, we define the *gradient Gaussian free field*, or gradient GFF for short, as the random d-dimensional distribution $\nabla \boldsymbol{\Psi}$ solving the equation

$$-\nabla \cdot \bar{\mathbf{a}} \nabla \boldsymbol{\Psi} = \nabla \cdot W. \tag{1.32}$$

In other words, $\nabla \boldsymbol{\Psi}$ is the potential part in the Helmholtz-Hodge decomposition $W = -\bar{\mathbf{a}} \nabla \boldsymbol{\Psi} + \mathbf{g}$, where $\mathbf{g}$ is the solenoidal (divergence-free) part. We can interpret

1. The remainder of this section essentially follows a presentation given in Banff in July 2015 which can be watched at http://goo.gl/5bgfpR.

this definition by duality: for every $F \in C_c^\infty(\mathbb{R}^d;\mathbb{R}^d)$, we set

$$(\nabla\mathbf{\Psi})(F) = W\left(\nabla(-\nabla\cdot\bar{\mathbf{a}}\nabla)^{-1}(\nabla\cdot F)\right).$$

The function $\nabla(-\nabla\cdot\bar{\mathbf{a}}\nabla)^{-1}(\nabla\cdot F)$ is the potential part in the Helmholtz-Hodge decomposition of F, and it belongs to $L^2(\mathbb{R}^d;\mathbb{R}^d)$.

This definition is perhaps easier to grasp in dimension $d \geq 3$, where we can define the field $\mathbf{\Psi}$ itself in a canonical manner. Indeed, in this case, we can interpret (1.32) as defining the centered, stationary Gaussian field with two-point correlation function

$$\mathbb{E}[\mathbf{\Psi}(0)\,\mathbf{\Psi}(x)] = \int_{\mathbb{R}^d} \nabla\mathcal{G}(0,y)\cdot\mathsf{Q}\nabla\mathcal{G}(y,x)\,\mathrm{d}y,$$

where $\mathcal{G}(\cdot,\cdot)$ is the Green function associated with the homogeneous operator $-\nabla\cdot\bar{\mathbf{a}}\nabla$. If the matrices Q and $\bar{\mathbf{a}}$ are proportional, then an integration by parts reveals that the right side above is a multiple of $\mathcal{G}(0,x)$. In this case, we therefore recover the more standard definition of the GFF as the centered Gaussian field whose covariance is given by a Green function. When $\bar{\mathbf{a}}$ and Q are not proportional, the field $\mathbf{\Psi}$ is a variant of the more standard GFF with the same scale invariance, but which does not satisfy the spatial Markov property (this was shown in [GM16b], together with the fact that this more general situation does indeed arise in the setting of stochastic homogenization).

Now that our notion of GFF is in place, we can resume our study of stochastic homogenization. As announced, we assume that the quantity ν has been suitably modified (in a way we do not describe here) so that it indeed becomes additive, and the next-order correction to (1.31) is ruled by the central limit theorem. We wish to encode this information in a way that is consistent with changes of the vector p, of the scale r, and of spatial translations. We do so by postulating that

$$\nu(x+\square_r,p) \simeq \frac{1}{2}p\cdot(\bar{\mathbf{a}}+W_r(x))p, \tag{1.33}$$

where W is a matrix white noise, and W_r denotes its spatial average on scale r:

$$W_r(x) := \fint_{x+\square_r} W. \tag{1.34}$$

(In this heuristic argument, we could also think of W_r as the convolution of W with a rescaled bump function: $W_r := W \star \chi^{(r)}$, with $\chi \in C_c^\infty(\mathbb{R}^d,\mathbb{R}_+)$ such that $\int\chi = 1$ and $\chi^{(r)} := r^{-d}\chi(\cdot/r)$.) This encodes in particular the fact that $\nu(x+\square_r,p)$ and $\nu(y+\square_r,q)$ are essentially independent if $x+\square_r$ and $y+\square_r$ are disjoint. By the scaling properties of white noise, each coordinate of $W_r(x)$ is of order $|\square_r|^{-1/2} = r^{-\frac{d}{2}} \ll 1$. We understand that the error in (1.33) is much smaller than $r^{-\frac{d}{2}}$.

Recall that the matrix $\mathbf{a}$ is meant to describe the correspondence between gradients of solutions and their associated fluxes. We interpret (1.33) as displaying the correspondence between *spatial averages* of gradients and fluxes of solutions. More precisely, we interpret (1.33) as meaning that if a function locally minimizes the energy over $x+\square_r$ and has average gradient p, then its flux over $x+\square_r$ is approximately $(\bar{\mathbf{a}}+W_r(x))p$.

We now fix $p \in \mathbb{R}^d$, denote by $\phi = \phi(\cdot,p)$ the corrector in the direction p, and by ϕ_r its spatial average on scale r:

$$\phi_r(x) := \fint_{x+\square_r} \phi.$$

The preceding discussion suggests that $x \mapsto p\cdot x + \phi_r(x)$ is approximately $(\bar{\mathbf{a}}+W_r)$-harmonic. In other words, we expect that the equation

$$-\nabla\cdot(\bar{\mathbf{a}}+W_r)(p+\nabla\phi_r) = 0$$

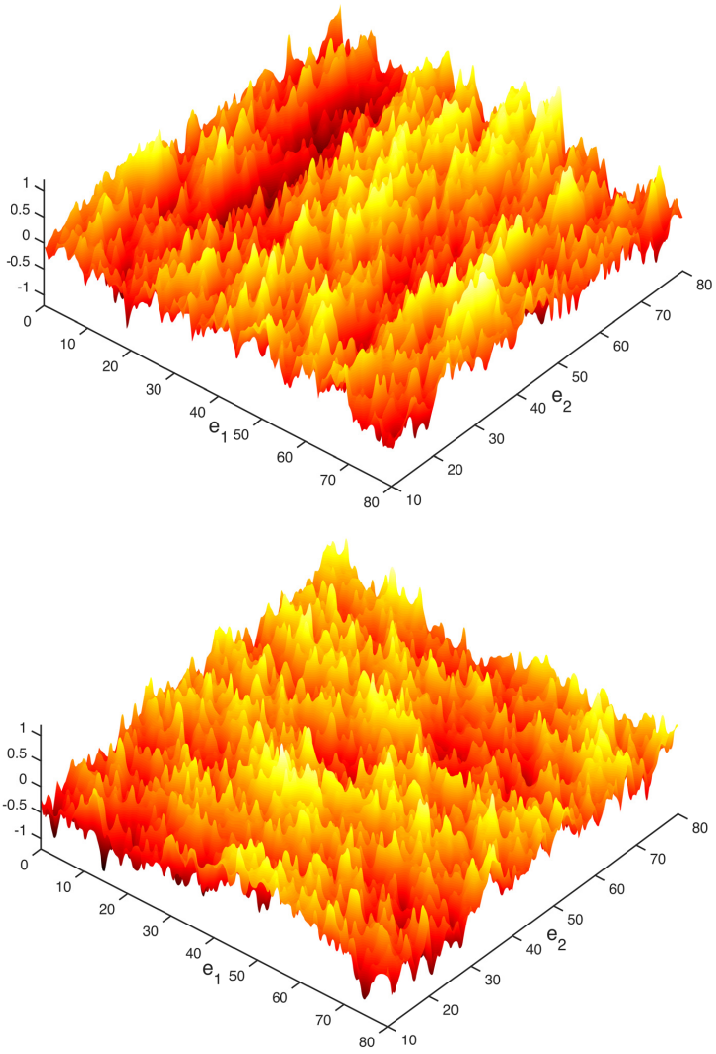

FIGURE 1.2. Graphs of the correctors ϕ_{e_1} and ϕ_{e_2} on the top and bottom, respectively, for a random checkerboard model in $d = 2$ (with the same realization). The random matrix is diagonal with independent entries equidistributed between 1 and 10. Notice that the mountain ranges for ϕ_e seem to line up in the orthogonal direction to e. The image is courtesy of Antti Hannukainen (Aalto University).

holds approximately. Rearranging, we obtain

$$-\nabla \cdot (\bar{\mathbf{a}} + W_r)\nabla\phi_r = \nabla \cdot (W_r p).$$

Since $W_r \ll 1$, we have $\nabla\phi_r \ll 1$. Therefore, the term W_r on the left side can be neglected, and we obtain

$$-\nabla \cdot \bar{\mathbf{a}}\nabla\phi_r = \nabla \cdot (W_r p), \tag{1.35}$$

up to errors that become negligible in the limit $r \to \infty$. Equation (1.35) characterizes a gradient GFF, compare with (1.32). Therefore, our heuristic argument suggests

that for a suitably weak topology,

$$r^{\frac{d}{2}}(\nabla\phi)(r\cdot) \xrightarrow[r\to\infty]{\text{(law)}} \nabla\mathbf{\Psi}, \tag{1.36}$$

where $\nabla\mathbf{\Psi}$ is a gradient GFF. This result was first proved using the nonlinear approach in [MO14, MN15]. The proof requires somewhat restrictive conditions, and as explained above, a core aspect of this approach is to use the Helffer-Sjöstrand representation of correlations and "follow the computations". With Armstrong and Kuusi, we gave a proof of this convergence [AKM16a, AKM16b] which is much closer in spirit to the heuristics we just presented, and holds under a finite range of dependence assumption. Figure 1.2 illustrates the fluctuations captured by this convergence.

Similar heuristics can be devised for solutions to

$$-\nabla\cdot\mathbf{a}\nabla u = f \qquad \text{in } \mathbb{R}^d, \tag{1.37}$$

where $f \in C_c^\infty(\mathbb{R}^d)$ varies on scale $\varepsilon^{-1} \gg 1$. (The function f should be of order ε^2 in order for u to be of order 1.) We consider the spatial average u_r of u over scale r, $1 \ll r \ll \varepsilon^{-1}$. By the same reasoning as above, we expect u_r to satisfy the coarsened equation

$$-\nabla\cdot(\bar{\mathbf{a}} + W_r)\nabla u_r = f.$$

We write $u_r = \overline{u} + \widetilde{u}_r$, where $\overline{u}$ solves

$$-\nabla\cdot\bar{\mathbf{a}}\nabla\overline{u} = f,$$

so that

$$-\nabla\cdot(\bar{\mathbf{a}} + W_r)\nabla\widetilde{u}_r = \nabla\cdot(W_r\nabla\overline{u}).$$

As before, we expect the term W_r on the left side to be negligible, so we obtain

$$-\nabla\cdot\bar{\mathbf{a}}\nabla\widetilde{u}_r = \nabla\cdot(W_r\nabla\overline{u}). \tag{1.38}$$

This approximate identity in law of large-scale spatial averages of solutions to (1.37) was proved in [GM15] using the nonlinear approach, with the important difference that $\widetilde{u}_r$ is replaced by the spatial average of $u - \mathbb{E}[u]$ (instead of $u - \overline{u}$). A proof based on the additive approach has not yet been developed for this problem. As was already pointed out, if we use the formal two-scale expansion $u \simeq \overline{u} + \sum_i \phi^{(i)}\partial_i\overline{u}$ and the large-scale description of the corrector in (1.35), we are led to a different and *incorrect* result.

2. Singular stochastic PDEs

In a broad sense, the elliptic PDEs with random coefficients considered in the previous section are "stochastic PDEs". As is somewhat customary, we will however restrict the use of this name to equations where the randomness appears only as a "thermal" noise, or more precisely, where the part of the equation with derivatives of highest degree is deterministic. In this section, we focus on a particular class of stochastic PDEs, which we may call *singular*. These PDEs contain a non-linearity, and typically a white noise forcing term. They are such that one does not expect the solution to be a proper function, but only to make sense as a distribution, due to the roughness of the noise. This makes the interpretation of the non-linearity problematic.

We will focus our discussion on two such equations: the Kardar-Parisi-Zhang (KPZ) equation

$$\partial_t h = \Delta h + |\nabla h|^2 + \xi, \tag{2.1}$$

and the Φ^4 model

$$\partial_t X = \Delta X - X^3 + aX + \xi, \tag{2.2}$$

where in both cases, ξ is a space-time white noise.

The KPZ equation has played a central role in the development of the theory. It was introduced in [140] as a model for the growth of an interface subject to fluctuations. In this case, the equation is posed on $\mathbb{R}_+ \times \mathbb{R}$, or a subset thereof. We may postulate that on a mesoscopic scale, the growth of the interface is governed by an equation of the form

$$\partial_t h = \Delta h + F(\nabla h) + \xi^\circ, \tag{2.3}$$

where the Laplacian Δ encodes a smoothing mechanism, the function F describes the growth mechanism per se, and ξ° is a smooth noise in space and time with short-ranged correlations. The growth mechanism is assumed to be smooth, symmetric: $F(z) = F(-z)$, and up to a change of time frame, we may assume that $F(0) = 0$ by replacing F by $F - F(0)$ if necessary. For some large parameter $\lambda \gg 1$, we consider the change of scale

$$\hat{h}(t,x) := \lambda^{-\frac{1}{2}} h(\lambda^2 t, \lambda x), \quad \hat{\xi} := \lambda^{\frac{3}{2}} \xi^\circ(\lambda^2 t, \lambda x),$$

so that

$$\partial_t \hat{h} = \Delta \hat{h} + \lambda^{\frac{3}{2}} F\left(\lambda^{-\frac{1}{2}} \nabla \hat{h}\right) + \hat{\xi},$$

and $\hat{\xi}$ converges to a space-time white noise as $\lambda \to \infty$. Since $F(z) \simeq a_2 z^2$ for small z, we see that the non-linear part becomes more and more dominant as we increase the scale. This suggests to "dampen" the non-linearity by replacing F with $\lambda^{-\frac{1}{2}} F$. Using the asymptotics $F(z) \simeq a_2 z^2 + a_4 z^4 + \cdots$, we are led to

$$\partial_t \hat{h} = \Delta \hat{h} + a_2 |\nabla \hat{h}|^2 + \lambda^{-1} a_4 |\nabla \hat{h}|^4 + \cdots + \hat{\xi}. \tag{2.4}$$

So we may expect that the higher-order terms $\lambda^{-1} a_4 |\nabla h|^4 + \cdots$ become negligible in the limit of large λ. This very informal computation in the spirit of [140] therefore suggests a certain universality of the equation (2.1) as a description of the large-scale fluctuations of growing interfaces. Note that in the derivation of the equation, we were forced to modify the non-linearity along the way in order to weaken it.

While these heuristics are very appealing, making sense of them mathematically is very challenging. Indeed, the solution to the linearized version of the equation (2.1), which reads

$$\partial_t h_{\mathrm{lin}} = \Delta h_{\mathrm{lin}} + \xi, \tag{2.5}$$

is such that for each fixed t, the function $x \mapsto h_{\mathrm{lin}}(t,x)$ has the regularity of Brownian motion: it is α-Hölder continuous for every $\alpha < 1/2$, and no more. In particular, the derivative of h_{lin} is a very singular object, and the square of such an object has no canonical meaning. In view of this, the Taylor expansion performed in (2.4) looks very worrisome, and indeed it is misleading, although the non-linearity does become quadratic in the limit — more on this below.

A naive attempt at defining a solution to (2.1) consists in regularizing the noise, e.g. by convolving the white noise field against a smooth bump function of scale ε to get a smooth noise ξ_ε, then solve

$$\partial_t h_\varepsilon = \Delta h_\varepsilon + |\nabla h_\varepsilon|^2 + \xi_\varepsilon,$$

and try to pass to the limit $\varepsilon \to 0$. However, the term ∇h_ε becomes more and more singular as ε tends to zero, and the sequence h_ε diverges. In short, we have to accept the fact that we cannot give a classical meaning to (2.1) as stated. We need to take a step back and find a suitable modification of (2.1) that enables to make sense of the equation, but that is sufficiently minor to ensure that the solution we find is physically relevant to interface growth models. Bertini and Giacomin [36] made

a major step in this direction. First, by a formal Cole-Hopf change of unknown function $Z := \exp(h)$, one arrives at

$$\partial_t Z = \Delta Z + Z\xi. \tag{2.6}$$

This is a linear equation in Z. One can make sense of this equation by writing its mild formulation and interpreting the integral involving $Z\xi$ as an Itô integral (thus avoiding to define the product $Z\xi$ per se). Moreover, Z is strictly positive almost surely [164], so we can decide to define the solution to (2.1) as $h := \log(Z)$. This may seem somewhat ad hoc, but strikingly, Bertini and Giacomin showed that the process thus defined indeed arises as the scaling limit of a certain model of interface growth. This model can be mapped to a particle system on the line, via the identification between the presence/absence of a particle and the slope of the interface being ±1. The dynamics of the particle system is that of simple exclusion, with an asymmetry that is slowly tuned down to zero, in the spirit of the gradual taming of the nonlinearity performed in the heuristic argument leading to (2.4).

If one regularizes the noise in space in (2.6) and writes

$$\partial_t Z_\varepsilon = \Delta Z_\varepsilon + Z_\varepsilon \xi_\varepsilon,$$

an application of Itô's formula reveals that the function $h_\varepsilon := \log(Z_\varepsilon)$ actually solves the modified equation

$$\partial_t h_\varepsilon = \Delta h_\varepsilon + |\nabla h_\varepsilon|^2 - C_\varepsilon + \xi_\varepsilon, \tag{2.7}$$

where $C_\varepsilon \sim c/\varepsilon$ as ε tends to 0, for some constant $c > 0$. By the intermission of the Cole-Hopf transform, we therefore conclude that

(1) for a suitable choice of $C_\varepsilon \sim c/\varepsilon$, the solution h_ε to (2.7) converges to a non-trivial limit h;

(2) this limit h can be obtained as the scaling limit of a natural model of interface growth.

Physicists would say that we added a *counter-term* C_ε to balance the divergence of the term $|\nabla h_\varepsilon|^2$. We may also say that we have "renormalized" our original equation (2.1). The introduction of this additional constant C_ε amounts to a shift in time of the solution, and is therefore very benign from a physical point of view: it simply indicates that the naive guess about the asymptotic speed of growth of the interface has to be corrected by a diverging constant.

While this development was of course a major progress in the mathematical understanding of the KPZ equation, a core aspect of the approach lies in the availability of the Cole-Hopf transformation, which in effect linearizes the equation. A downside is that it completely avoids trying to make sense of the equation (2.1) directly. Moreover, the passage from discrete to continuum also requires a microscopic version of the Cole-Hopf transform (first observed in [98]), and therefore the set of interface growth models for which the approach of [36] applies is rather rigid (see however [77, 65]).

As far as the continuous equation is concerned, these shortcomings were overcome by Hairer, first for the KPZ equation per se [125], and then within a much more general framework covering at once the two equations (2.1) and (2.2) of interest to us here [126]. The key property that the equation needs to satisfy is that of *subcriticality*, or in the language of quantum field theory, the equation must be *superrenormalizable*.

Loosely speaking, a (formal) stochastic PDE is said to be subcritical if the non-linearity is dampened when we zoom in on a (formal) solution. For the KPZ equation in one space dimension, we have already observed that the non-linear term formally blows up when we consider the rescaled solution on larger and larger scales.

Reversing the scaling by letting λ tend to zero instead of infinity corresponds to zooming on the fine details of the solution; this has the opposite effect of reducing the strength of the non-linearity. This equation is therefore subcritical.

We now perform the same analysis for the Φ^4 model (2.2), allowing the space dimension d to be arbitrary. (Background and motivation for this equation will be provided in Chapter III.) The change of scale

$$\hat{X}(t,x) := \lambda^{\frac{d-2}{2}} X(\lambda^2 t, \lambda x), \quad \hat{\xi}(t,x) := \lambda^{\frac{d+2}{2}} \xi(\lambda^2 t, \lambda x)$$

is so that $\hat{\xi}$ and ξ have the same law, and moreover,

$$\partial_t \hat{X} = \Delta \hat{X} - \lambda^{4-d} \hat{X}^3 + \lambda^2 a \hat{X} + \hat{\xi}. \tag{2.8}$$

The non-linearity is tamed down as we let λ tend to zero if $d < 4$: the equation is therefore subcritical in these cases.

In spatial dimension $d = 1$, the solution is expected to be a proper function (as explained above in the context of the KPZ equation), and therefore the stochastic PDE can be made sense of by classical methods. The dimensions of interest to us are thefore $d \in \{2, 3\}$.

When the equation is subcritical, we can hope to obtain an existence theory for the equation by developing a generalized form of "Taylor expansion" of the solution, where the first order of the expansion is described in terms of the solution to the linear equation

$$\partial_t Z = \Delta Z + \xi. \tag{2.9}$$

In spatial dimension $d = 2$, this idea is by now well-understood since the work of Da Prato and Debussche [68] (earlier contributions using different approaches include [138, 3, 177, 161]). The argument proceeds in two steps. First, one can define "renormalized" (or Wick) powers of the solution Z to the linear equation: writing Z_ε for the solution to (2.9) with ξ replaced by the smoothed noise ξ_ε, there exists a constant C_ε which diverges logarithmically as $\varepsilon \to 0$ and such that

$$Z_\varepsilon^2 - C_\varepsilon \xrightarrow[\varepsilon \to 0]{} Z^{:2:}, \qquad Z_\varepsilon^3 - 3C_\varepsilon Z \xrightarrow[\varepsilon \to 0]{} Z^{:3:}, \tag{2.10}$$

and so on with higher order Hermite polynomials. Second, the property of subcriticality suggests that the difference $Y := X - Z$ should have better regularity than X itself, or in other words that we should look for a "Taylor expansion" of X of the form $X = Y + Z$, where Z is hopefully sufficiently regular to enable to write a meaningful equation for it. Indeed, formally starting from (2.2) gives

$$(\partial_t - \Delta) Y = -(Y + Z)^3 + a(Y + Z).$$

Expanding the cubic power leads to undefined terms, but the construction of the renormalized powers of Z suggests the renormalization rules

$$Z^2 \rightsquigarrow Z^{:2:}, \qquad Z^3 \rightsquigarrow Z^{:3:},$$

so that the equation above becomes

$$(\partial_t - \Delta) Y = -Y^3 - 3Y^2 Z - 3Y Z^{:2:} - Z^{:3:} + a(Y + Z). \tag{2.11}$$

As will be explained in more details in Chapter III, in spatial dimension $d = 2$, the renormalized powers of Z belong to every function space of negative regularity, and therefore the equation (2.11) can be solved with Y being almost twice differentiable. This defines a notion of solution to (2.2), and moreover, the solution depends continuously on the triple $(Z, Z^{:2:}, Z^{:3:})$ (for a suitable topology). If we replace the renormalized powers $Z^{:2:}$, $Z^{:3:}$ in (2.11) by their regularized approximations in (2.10),

we obtain a solution Y_ε so that the process $X_\varepsilon := Z_\varepsilon + Y_\varepsilon$ solves the renormalized equation

$$(\partial_t - \Delta)X_\varepsilon = -X_\varepsilon^3 + 3C_\varepsilon X_\varepsilon + aX_\varepsilon + \xi_\varepsilon.$$

Therefore, in this case, the renormalization takes the form of the addition of a large (diverging) linear term to the equation. We will use the informal notation

$$(\partial_t - \Delta)X = -\big(X^3 - 3\infty X\big) + aX + \xi \tag{2.12}$$

to evoke this fact, in the sense that

$$X^3 - 3\infty X := \lim_{\varepsilon\to 0}\big(X_\varepsilon^3 - 3C_\varepsilon X_\varepsilon\big) = Y^3 + 3Y^2 Z + 3YZ^{:2:} + Z^{:3:}.$$

The situation in spatial dimension $d = 3$ is much more intricate, since the equation (2.11) is still ill-posed, and one needs to pursue the "Taylor expansion" of X to higher order. We can reproduce the previous idea and subtract further the solution to the heat equation with right-hand side $Z^{:3:}$, but this is still not sufficient, and then the term $YZ^{:2:}$ no longer allows us to define an easy next-order guess and subtract it from Z.

The theory of regularity structures introduced by Hairer [126] enables to solve this problem, by developing a system of generalized "Taylor expansions" where basic units of an expansion can be not only a monomial, but also a given distribution such as $Z^{:2:}$. This major breakthrough led to an explosion of activity on the subject. We will review in Chapter III the related technique introduced by Gubinelli, Imkeller and Perkowski [120, 57] based on paraproducts. As of now, this second approach is less systematic than the approach based on regularity structures, and operates a less clear separation between the analytic and algebraic aspects of the equation at hand. This limitation also comes with the advantage of being somewhat more "hands-on" in the case of the Φ^4 model. Yet another approach to define solutions to (2.2) is due to Kupiainen [148] and based on "Wilsonian" renormalization group arguments. This last approach enables to show that a sequence of suitably regularized solutions converges to a non-trivial distribution, but the limit is not identified as solving an equation (or a system of equations); nor does it provide us with a pathwise notion of solution.

The assumption of subcriticality is meant to ensure that one can end up with a well-defined equation after a finite number of renormalization steps. Loosely speaking, these iterated renormalization steps ultimately enable to treat the non-linearity as a perturbation of the linear equation. The argument is therefore very robust. This robustness also comes at a cost, since the solution we obtain in this way is only defined up to a strictly positive but possibly very small random explosion time. Also, the approach only applies to compact space domains ($[-1,1]^d$, say).

Since the framework covers at once (2.11) and the same equation with the sign of the non-linearity $-X^3$ reversed, this is unavoidable. Indeed, the equation with the sign reversed is expected to blow up in finite time, with no uniform lower bound on the explosion time. For similar reasons, by its generality, the theory cannot cover unbounded space domains. Indeed, if the solution on the torus blows up after a random time that may be arbitrarily small, then increasing the size of the domain to infinity will bring the explosion time to zero almost surely (since heuristically, we repeat the "blow-up time experiment" many times independently). If one wants to go beyond this limitation, then the approach will have to become model-specific. An important part of my work with Hendrik Weber aims at lifting these limitations in the case of the Φ^4 model. As will be discussed in Chapter III, we showed first that the model in two space dimensions is well-defined for arbitrarily large times and on the full space $\mathbb{R}^2$ [MW15]. We then showed that on the three-dimensional torus, the solution is also well-defined for arbitrarily large times [MW16].

The second important theme of my work with Weber aims at building a robust method to show that certain discrete systems converge to singular stochastic PDEs such as the Φ^4 model. By the very definition of the notion of subcriticality, a solution to the equation is *not* invariant under scaling. Therefore, we need a family of models with a tunable parameter (like the asymmetry for the exclusion process) rather than a single model of which we would take the scaling limit if we want such stochastic PDEs to emerge as limits. For the Φ^4 model, we investigated the Glauber dynamics of a two-dimensional Ising model with long-range interactions. The long-range character of the interaction is technically convenient, but also serves as this required tunable parameter. We show in [MW14] that for temperatures sufficiently close to criticality, the suitably rescaled magnetic field converges in law to the Φ^4 model. This will also be reviewed more precisely in Chapter III. Similarly to the result of [36] for the KPZ equation, this justifies in particular that the chosen definition of the solution to the Φ^4 model is physically relevant. The necessary renormalization translates into a shift of the critical temperature away from its mean-field value. We believe that the technique developped there can be extended to cover other problems, for instance to show a similar result in three space dimensions, but this is of course more difficult.

We now review other works related to these questions. Very recently, the three-dimensional Φ^4 model has been shown to arise as the limit of a variety of other models. The situation closest to our work is probably that investigated by Zhu-Zhu [203] and Hairer-Matetski [129], who show the convergence of certain discretization schemes of the Φ^4 equation to the solution of the continuous equation in three space dimensions. The related work of [132, 184] identifies the Φ^4 model as the limit of a range of suitably scaled stochastic PDEs, in the spirit of (2.3) for the KPZ equation.

Returning to the KPZ equation, global existence in the full space is known via the Cole-Hopf transformation. A more direct proof of global existence was obtained by Gubinelli and Perkowski [122] using the interpretation of the KPZ equation as a control problem. As for universality, Hairer, Quastel and Shen [130, 131] showed that a large class of continuous stochastic PDEs in the spirit of (2.3) converge to the solution to the KPZ equation, after suitable scalings. In particular, while it is true that the non-linearity becomes quadratic in the limit, the constant prefactor depends on *every* coefficient a_2, a_4, etc., contrary to what the heuristic computation in (2.11) suggests. (In fact, it depends only on the quadratic part of the expansion of F if we expand F along the basis of Hermite polynomials.)

For the question of the convergence of discrete systems to the KPZ equation, a very different approach was put forward by Gonçalves and Jara [118] (see also [16]), based on the notion of energy solutions to the stochastic Burgers equation (which is the equation solved by $\partial_x h$). They showed that many discrete systems are attracted to this set of energy solutions. Gubinelli and Perkowski [121, 123] then showed that a suitable modification of the notion of energy solution guarantees uniqueness of the solution. This was also used to study scaling limits of weakly asymmetric PDEs in [124]. However, as of now, this approach only applies for processes started at their equilibrium measure, and this invariant measure must be relatively under control.

We conclude this brief overwiew by mentioning related works falling outside of the realm of subcritical singular stochastic PDEs. First, the KPZ equation is not the only possible meaningful limit of an interface growth model. For instance, the interface model associated with the asymmetric simple exclusion process with a *fixed* strength of asymmetry is expected to converge to a different limit, known as the *KPZ fixed point*. We refer to [137, 19, 176] for the first works in this direction, and to [195, 8, 182, 64, 42] for more recent contributions. This object is expected to describe the large-scale behavior of the solution to the KPZ equation itself.

Paralleling this exploration of the scale-invariant KPZ fixed point is the study of the scaling limit of the critical nearest-neighbor Ising model. In two space dimensions, the understanding of the nearest-neighbor Ising model (on well-chosen graphs) is by now very advanced [187, 50, 61, 62]. As was shown in [110], the Φ^4 measure in two dimensions has a phase transition. It is natural to conjecture that the critical Φ^4 measure rescales to the same limit as that of the nearest-neighbor Ising model.

Finally, we mention the recent work [22] studying the Φ^4 model in the critical four-dimensional case, using perturbative renormalization group arguments, and [55] on a range of models including the heat equation with multipicative noise in the critical two-dimensional case.

3. The Anderson model

Let $(V(x))_{x\in\mathbb{Z}^d}$ be a family of i.i.d. random variables taking values in $[0,\infty]$. In the first part of the discussion, we assume the random variables to be bounded for simplicity. A long-standing open problem, brought forward and understood at a physical level by Anderson [12], is to understand the nature of the spectrum of the operator $H_\beta := -\Delta + \beta V$, where Δ denotes the discrete Laplacian on $\mathbb{Z}^d$, and $\beta \geq 0$ is a free parameter. The operator H_β is self-adjoint on $\ell^2(\mathbb{Z}^d)$. At one extreme ($\beta \to \infty$), the eigenfunctions of the multiplication operator $f \mapsto Vf$ are the Dirac functions at each site of $\mathbb{Z}^d$, and are therefore as localized as can be. At the other extreme, the eigenfunctions of the operator $H_0 = -\Delta$ are sine functions, if we allow for an extended notion of eigenfunction, since these do not belong to $\ell^2(\mathbb{Z}^d)$. They are completely delocalized. The precise notion of localized vs. delocalized eigenfunction correponds to whether the eigenfunction is in $\ell^2(\mathbb{Z}^d)$ or not.

The nature of the spectrum characterizes the evolution of the solution to the associated Schrödinger equation: if we start the Schrödinger equation from an initial condition belonging to an eigenspace of localized eigenfunctions, then the wave packet will remain in a bounded region of space. In the opposite situation of delocalized eigenfunctions, the wave function will spread without bound, and ultimately leave no mass in any finite region of space. The first behavior corresponds to an insulating medium, where electrons do not propagate, while the second behavior corresponds to a conducting medium. Anderson was interested in the conducting or insulating properties of certain materials, and he introduced this model with the belief that random inhomogeneities are of crucial importance.

Anderson's prediction is that in dimension $d \geq 3$, the nature of the spectrum is as described on Figure 3.3. There are by now robust mathematical techniques enabling to detect localized spectrum, based on the analysis of the decay of the Green function at large distances (see [97, 96, 196] for the multiscale method, and [1, 2] for the fractional-moment approach). This permits to justify part of Figure 3.3: the spectrum has been shown to be localized near the edges of the spectrum for each $\beta > 0$, and at every energy level for large β. On the other hand, the existence of a region of delocalized spectrum is a long-standing open question, although very interesting progress on an explicitly solvable model with a particular correlated field $(V(x))$ has been achieved recently, see [82, 159, 81, 178, 179].

It is very desirable to try to describe the position of the critical energy delimiting the transition from localized to delocalized spectrum, in the regime of small $\beta > 0$. Lacking any understanding of the localized phase, we can look for hopefully sharp lower bounds on this critical energy level. In other words, we want to assert that the spectrum of H_β is localized in an interval $[0, E_\beta]$ for E_β as large as possible, and hopefully asymptotically sharp in the limit $\beta \to 0$. My understanding is that physicists expect the arguments below to indeed give a sharp answer for $d \geq 3$ in

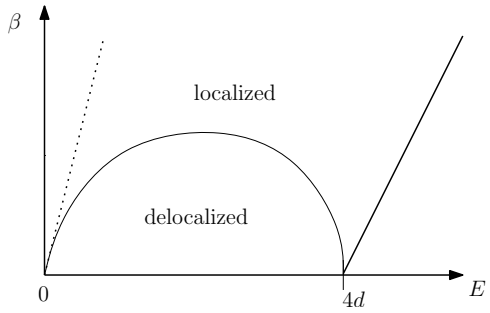

FIGURE 3.3. A cartoon of the conjectured nature of the spectrum of H_β as a function of β, when the support of the law of V is $[0, c]$ and $d \geq 3$. When $\beta = 0$, the spectrum is delocalized between 0 and $4d$. For general β, the spectrum is the interval $[0, 4d + \beta c]$; the right limit of this interval is materialized by the thick line. The mathematical analysis discussed below focuses on the region of small β and E. The region of localized spectrum is known to asymptotically contain $[0, \beta\mathbb{E}[V])$; the right limit of this region is materialized by the dashed line. The existence of the delocalized phase, and a fortiori the asymptotic sharpness of the dashed line, are conjectural.

the limit of small $\beta > 0$, but this is of course a conjecture. The argument allows to identify the region where the spectrum is localized and the density of eigenvalues per unit of volume is exponentially small in β^{-1} (the regime of "Lifshitz tails"). The energy level marking the limit of the regime of Lifshitz tails can be read off the behavior of the Green function. Intuitively, there are almost no eigenvalues in the interval $[0, E_\beta]$, and the Green function decays at large distances as that of $-\Delta + E_\beta$, that is, $G_\beta(x, y) \simeq \exp\left(-\sqrt{2dE_\beta}|y - x|\right)$ (the factor $2d$ comes from the choice of normalization of the discrete Laplacian).

The first mathematical results concerning the evaluation of E_β are due to Wang [198, 199, 197], and are based on representations of the Green function in terms of anti-commuting variables. Wang showed that if the potential V has finite second moment, then, in the limit of large $|x|$ and then small $\beta > 0$,

$$\exp\left(-C\sqrt{\beta}|x|\right) \lesssim G_\beta(0, x) \lesssim \exp\left(-C^{-1}\sqrt{\beta}|x|\right).$$

This estimate was then refined by Klopp and Kosygina, Mountford, Zerner [145, 146]: it was shown that if $\mathbb{E}[V]$ is finite, then, in the limit of large $|x|$ and then small $\beta > 0$,

$$G_\beta(0, x) \simeq \exp\left(-\sqrt{2d\beta\mathbb{E}[V]}|x|\right),$$

and that the spectrum is therefore localized in the region $[0, \beta\mathbb{E}[V])$.

This shows that the region of Lifshitz tails is the "naive" one, that is, the behavior matches with the spectrum of the naively homogenized operator $-\Delta + \beta\mathbb{E}[V]$. (This operator displays obvious Lifshitz tails in the range of energy $[0, \beta\mathbb{E}[V])$, since there is no spectrum at all there.) With Tom Mountford, we wanted to explore richer behaviors, by relaxing the hypothesis of finite moment on the potential. Another possible direction would be to study dilute limits of a fixed potential, for instance a potential taking values in $\{0, 1\}$ and considering the limit of small densities of 1's. The two situations are in fact very similar, and are covered by the same method. For simplicity, we will focus on the first case.

The first step towards this goal was to show the existence of the limit

$$\lim_{r\to\infty} \frac{1}{r} \log G_\beta(0, rx). \tag{3.1}$$

This was previously known only under the assumption that $\mathbb{E}[V^d]$ is finite [202]. Refining the approach to the shape theorem developed for first-passage percolation in [66], I showed in [Mou12a] the validity of suitable intepretations of (3.1) under the weaker assumption that the set $\{x : V(x) < \infty\}$ is a supercritical percolation field. The proof uses the correspondence between the operator $-\Delta + V$ and the simple random walk with killing potential V. As was shown for the first time by Sznitman [193] in the context of Brownian motion in soft Poissonian potential, understanding the limiting behavior of the Green function enables to show a large deviation principle for the random walk conditioned to survive, under the Gibbs measure with killing potential βV. The proof was adapted to potentials possibly taking infinite values in [Mou12a]. As a side remark, we note that these considerations are closely related to the homogenization of a Hamilton-Jacobi equation, see [194, Section 6.1] or [15].

Under the assumption that $V < \infty$ almost surely and for $d \geq 3$, Mountford and I [MM13, MM15b] showed that, roughly speaking,

$$G_\beta(0,x) \simeq \exp\left(-\sqrt{2\beta E_\beta}|x|\right) \qquad (|x| \to \infty),$$

with

$$E_\beta \sim \int \left(\frac{1}{q_d} + \frac{1}{\beta z}\right)^{-1} \mathrm{d}\mu(z). \qquad (\beta \to 0), \tag{3.2}$$

where q_d is the probability that the simple random walk never returns to its starting point, μ is the law of $V(x)$, and (3.2) means that the ratio of the quantities on the left and right sides tends to 1 as $\beta \to 0$. In [MM13, MM15b], the discrete-time random walk is used for the analysis; passing from the Green function of this random walk to that of the operator $-\Delta + V$ (which is the Green function of the continuous-time random walk) requires a minor change of variables, see [MM13, (6.5)].

This result is obtained via a detailed analysis of the best possible strategies for the random walk to survive until reaching a distant point (or hyperplane). It reveals that, on the one hand, sites whose potential is much smaller than β^{-1} are "too numerous" and are therefore sampled by the random walk according to the law of large numbers. At the other extreme, sites whose potential is much larger than β^{-1} are avoided by the walk. The sites with potential of order β^{-1} are not completely avoided, but are visited less frequently. The exact optimization is described more precisely in the introduction to [MM13].

This result enables to understand some aspects of the large deviation behavior of the random walk conditioned to survive. As an example, if we add a constant "pulling" of strength h to the Gibbs measure, then it is known that there exists a critical $h_c(\beta)$ such that the walk is sub-ballistic when $h < h_c(\beta)$, and ballistic when $h > h_c(\beta)$, see [193, 95] and [Mou12a, Remark 1.11]. The result (3.2) enables to derive an asymptotic expansion for $h_c(\beta)$ as $\beta \to 0$. Using the arguments laid down in [145], one can also show that the spectrum of the operator $H_\beta = -\Delta + \beta V$ is localized in the energy band between 0 and (almost) $\int \left(\frac{1}{q_d} + \frac{1}{\beta z}\right)^{-1} \mathrm{d}\mu(z)$, in the limit $\beta \to 0$.

We now conclude this section by outlining some related questions. Assuming that V is bounded for simplicity, one may wish to understand the next-order correction to the upper limit of Lifshitz tail behavior, which we said is $\sim \beta\mathbb{E}[V]$ at first order. This analysis has been carried out in [88] in dimension $d = 3$: the next-order correction is found to be $\beta\mathbb{E}[V] - c\beta^2$, for an explicit constant $c > 0$.

Instead of imposing the potential to be nonnegative, we may consider it to be centered. That is, we simply subtract $\beta\mathbb{E}[V]$ to the operator, so that the regime of

Lifshitz tail is now located in the interval $[-\beta\mathbb{E}[V], -c\beta^2]$. It is natural to expect that a suitable scaling limit of this model then converges to $-\Delta + \xi$, where now Δ is the continuous Laplacian and ξ is a spatial white noise. However, the definition of the operator $-\Delta + \xi$ is not clear a priori, due to the irregularity of the noise. The associated parabolic equation is ill-posed as soon as $d \geq 2$: a suitable, renormalized construction of this equation has been carried out on the two-dimensional torus in [126, 120], on $\mathbb{R}^2$ in [127], and on $\mathbb{R}^3$ in [128, 20]. It would be very interesting to develop an understanding of the spectrum of the continuous Anderson operator, and to show that the continuous and discrete Anderson operators are close to one another under a suitable scaling, including at the level of their spectrum. Some first steps in this direction, for the two-dimensional continuous Anderson operator, are obtained in [6].

4. The pinning model

Let $\omega := (\omega_n)_{n\in\mathbb{N}}$ be i.i.d. random variables, whose law we write $\mathbb{P}$, and let (S_n) be a simple random walk in $\mathbb{Z}$ starting at the origin, with law $\mathbf{P}$, independent of ω. The pinning model with parameters $\beta, h \geq 0$ and length n is the Gibbs measure $\mathbf{P}_n^{\beta,h,\omega}$ with density with respect to $\mathbf{P}_0$ proportional to

$$\exp\left(\sum_{k=0}^{n-1}(\beta\omega_k - h)\mathbf{1}_{\{S_k=0\}}\right). \tag{4.1}$$

The interpretation of the model is that of a directed polymer $(n, S_n)_{n\in\mathbb{N}}$ that receives an energy reward of $(\beta\omega_k - h)$ (or a penalty if this quantity is negative) if it touches the "defect line" $\{(x, y) \in \mathbb{Z}^2 \ : \ y = 0\}$ at the position $(k, 0)$. In fact, the model depends on the random walk only through the law of inter-arrivals on the defect line, and one can therefore easily formulate a more general model defined in terms of the renewal sequence of inter-arrival times only.

There exists a critical parameter $h_c(\beta) \geq 0$ such that the trajectory of the polymer is *localized* if $h < h_c(\beta)$, and delocalized otherwise. By definition, the localized phase corresponds to a situation where the density of contacts with the line of defects grows proportionally with n. The delocalized phase is the complementary region.

Under restrictive assumptions on the law of ω and with respect to the averaged measure $\mathbb{P}\mathbf{P}_n^{\beta,h,\omega}$, it was shown by Giacomin and Toninelli [105] (see also [103, Section 8.2]) that when $h > h_c(\beta)$, the number of points of contact with the defect line is bounded by $C \log n$. In a short note [Mou12b], I showed that one can strengthen this result: no condition on the law of ω is necessary, and the result holds $\mathbb{P}$-almost surely. The result was later refined by Alexander and Zygouras [5]: building upon [Mou12b], they showed that the logarithmic bound is sharp, if one wants a $\mathbb{P}$-almost sure result, and that the number of points of contact is in fact tight in $\mathbb{P}$-probability.

The approach of [Mou12b] (and therefore also [5]) rely on the large deviation results of Birkner, Greven and den Hollander [39]. Unfortunately, Julien Poisat and I discovered an unresolved difficulty in the proof of the main result of [39]. (This problem was communicated to the authors in February 2014; as far as I know, it is still unresolved.) The main result of [39] is a large deviation principle for the operation of cutting "words" out of a sequence of i.i.d. "letters". In the context of the pinning model, the letter sequence is (ω_n), and the polymer obviously operates a cut in this letter sequence. It is very natural to write large deviation principles for sequences of a fixed number of words; in the context of the polymer model, it is more natural to consider that it is the number of letters which is fixed. Passing from one notion to the other would be facilitated by a good understanding of the

effect of truncations of the words. The problem we identified in [39] revolves around an ambiguity of the notion of truncation of a word. On the one hand, it is very natural in the large-deviation framework to ask about rate functions of the process of the first r letters of each word. The rate function of the sequence of full words is obtained as the increasing limit of these rate functions as r tends to infinity, by general arguments. On the other hand, what would be most useful in the context of the pinning model would be to know that if a certain word sequence distribution has some cost, then the same sequence with words cut at length r has roughly the same cost. This is very different from the first notion: we now wonder about the cost of a sequence of words none of which is of length exceeding r, while the first notion, refering to projections, does not constraint the length of words in any way. In [39, (5.8)], the notion referring to projections is used, although it is then applied with the second understanding of the notion of truncation. A clear manifestation of the problem is in the first line of [60, (2.12)]: taking Q to be the reference measure P, we obtain 0 as a supremum of strictly positive quantities (the relative entropy of the truncated word sequence with respect to the untruncated one). In view of this unresolved difficulty, the status of the result of [Mou12b] remains conditional. Poisat and I came to the realization of this problem while thinking about the very interesting conjecture of Derrida and Retaux [80] that the depinning transition is of infinite order.

5. Aging of glassy systems

As a first approximation, a glassy system is a model with many degrees of freedom and a very rough, "random-like" energy landscape, with many different wells and valleys. A mathematical model expected to be glassy is the Edwards-Anderson model. This model is defined like the Ising model on $\mathbb{Z}^d$, with the crucial difference that the coupling between nearest spins is chosen randomly and can be positive of negative. Usually, the couplings are independent centered Gaussians. Typically, it is not possible to arrange the spins so that for every pair of neighboring spins, the spins are aligned if and only if the coupling is positive. Every configuration contains "frustrated pairs".

Experimentally, a signature of glassy materials is that below a certain temperature, the material seems to be forever thermodynamically out of equilibrium, with properties such as stiffness or magnetic susceptibility relaxing over time-scales ranging from seconds to years and more, in a power-law fashion [191, 152]. The material is said to *age*.

The Edwards-Anderson model is very natural, but also extremely difficult to analyse, in particular in terms of dynamical relaxation properties. It is therefore natural to consider simpler, mean-field models. These models depart from the glassy materials one can study in a physics lab (e.g. window glass, plastics, or certain alloys of magnetic materials). On the other hand, the mean-field models are more closely related to some problems arising in computer science or statistics, as for instance random constraint satisfaction problems like "K-SAT", or high-dimensional inference.

The simplest model of a mean-field spin glass is the Random energy model introduced by Derrida [78, 79]. The model is defined on the hypercube $\{-1,1\}^N$. Each site is thought of as representing a spin configuration. To each site $x \in \{-1,1\}^N$ is associated an energy E_x, and the (E_x) are assumed to be i.i.d. Gaussian random variables. This is of course an extremely simplifying assumption, but the analysis of the dynamics remains non-trivial. With Pierre Mathieu, we showed in [MM15a] that Glauber-type dynamics of this model exhibit aging below the critical temperature.

Previous studies concerning the Random energy model had focused on a simpler dynamics called "Random hopping times" [25, 26, 27]. This dynamics is simply a time-change of the simple random walk on the hypercube. This of course simplifies the analysis a lot, since one can sample the trajectory of the dynamics a priori, without reference to the energy landscape. For this same dynamics, the phenomenon of aging could be extended to the p-spin model, which is a correlated mean-field spin-glass model [23, 24, 46]. More recent developments, which aim to extend the class of dynamics considered in [MM15a], but still for the Random energy model, include [99, 58, 100].

During my Ph.D. thesis, I worked on a different toy model, which is very similar to the Random energy model, but is defined on $\mathbb{Z}^d$. The proof of aging developed in [Mou11a] relies on the results of [Mou11b], which was my first contribution to the topic of quantitative homogenization.

The type of aging phenomenon proved in all the mathematical works mentioned so far can be roughly summarized as follows. Asymptotically, the process indicating the Gibbs weight of the configuration where the dynamics stands is close to the following much simpler process: take i.i.d. non-negative random variables $(\tau_i)_{i\in\mathbb{N}}$ such that $\mathbb{P}[\tau_i \geq x] \sim x^{-\alpha}$ as x tends to infinity, where $\alpha \in (0,1)$; then define the process that takes value τ_1 for a length of time τ_1, then τ_2 for a length of time τ_2, and so on. This very simple process exhibits aging in the sense that the typical value of the process at time t does not converge in law as t tends to infinity, but instead takes on larger and larger values, in a power-law fashion. This is related to the fact that one cannot build a stationary renewal sequence out of identically distributed waiting times with infinite mean.

It is debatable whether such a process captures the most physically intersting aspects of aging. This phenomenology was introduced by Bouchaud [43], and is referred to as the "trap model". My understanding is that when physicists refer to aging, they usually have in mind dynamics that mostly go down in the energy landscape, display complex memory effects, etc. The simpler aging picture described mathematically above perhaps takes place only once the last accessible local minima are reached, when the dynamics goes down a local minimum, waits some time there, then exits, mixes, and starts again. The energy landcape in such a scenario has been compared to a golf course [139].

More refined versions of this phenomenology have been proposed and studied, by organizing similar behaviors into a hierarchical structure [44, 47, 180, 181, 101]. This construction is inspired by Parisi's discovery of ultrametric structures within mean-field spin glasses. As of now, the possible emergence of this hierarchical structure for the dynamics of these models is a widely open question.

Different approaches and phenomenology related to aging include the study of short-time dynamics of p-spin models, where the infinite volume limit is taken first and then time is sent to infinity [67, 28, 29, 76], and the study of kinetically constrained models, see for instance [90] and references therein.

6. The contact process

The contact process can be thought of as a model for the spread of an infection in a population. Given a discrete (possibly infinite) connected graph $G = (V, E)$, the contact process on G is a Markov process on $\{0,1\}^V$, which we denote by $(\xi_t)_{t\geq 0}$. If $\xi_t(x) = 0$, then we say that the site x is healthy at time t; otherwise, we say that it is infected. The dynamics is characterized by the infection parameter $\lambda > 0$: every infected individual becomes healthy (or "recovers") with rate 1, and infects each of its neighbors with rate λ.

A very convenient way to show existence of this process is via the so-called graphical construction, see [151, Section I.1]. The idea is to fix a priori every possible event of recovery or infection using Poisson point processes in the space-time domain. This construction provides with a monotone coupling of the dynamics started with different initial configurations. It maps the contact process onto a partially oriented percolation process on $\mathbb{R}_+ \times V$.

We say that the contact process dies out if $\xi_t = 0$ for every t sufficiently large. By monotonicity, there exists $\lambda_1(G) \in [0, \infty]$ such that the process started from a finite configuration dies out almost surely if $\lambda < \lambda_1(G)$, and survives with positive probability if $\lambda > \lambda_1(G)$. The event of survival can itself take different forms: we say that the process survives strongly if for every $x \in V$, we have $\xi_t(x) = 1$ for arbitrarily large times t; we say that it survives weakly if it survives but does not survive strongly. The latter event corresponds to the situation where the infection survives by "spreading to infinity". By monotonicity again, there exists $\lambda_2(G) \in [\lambda_1(G), \infty]$ such that the process started from a finite configuration survives weakly with positive probability but a.s. not strongly if $\lambda_1(G) < \lambda < \lambda_2(G)$, and survives strongly with positive probability if $\lambda > \lambda_2(G)$.

The contact process was first studied on the graph $\mathbb{Z}^d$. Bezuidenhout and Grimmett [38] showed that in this case, we have $\lambda_1(\mathbb{Z}^d) = \lambda_2(\mathbb{Z}^d) \in (0, \infty)$, and moreover, the contact process dies out almost surely at $\lambda = \lambda_2(\mathbb{Z}^d)$.

For every $d \geq 2$, let $\mathbb{T}^d$ denote the infinite tree where each vertex has $(d+1)$ neighbors. Pemantle [175] discovered that on $\mathbb{T}^d$, the contact process exhibits two non-trivial phase transitions. More precisely, it was shown in [175, 150, 190] that $0 < \lambda_1(\mathbb{T}^d) < \lambda_2(\mathbb{T}^d) < \infty$, and moreover, the process dies out when $\lambda = \lambda_1(\mathbb{T}^d)$ and does not survive strongly when $\lambda = \lambda_2(\mathbb{T}^d)$.

If G is a finite graph, then the contact process eventually dies out. However, this does not rule out important changes in the behavior of the contact process as λ varies. Denote by τ the extinction time of the contact process started from the fully infected configuration. If one considers the contact process on increasingly large boxes of $\mathbb{Z}^d$ of side length n, then τ is $(c + o(1)) \log n$ if $\lambda < \lambda_2(\mathbb{Z}^d)$, and is $\exp((c + o(1))n^d)$ when $\lambda > \lambda_2(\mathbb{Z}^d)$ [84, 85, 63, 162, 163]. (The estimates are for $n \to \infty$. In the critical case $\lambda = \lambda_2(\mathbb{Z}^d)$, we only know that $n \ll \tau \ll n^4$ in dimension $d = 1$, see [86].)

The situation for other graphs was not as well-understood. In the case of $(d+1)$-regular trees cut at a given height h, Stacey [189] showed that τ is within multiplicative constants of h when $\lambda < \lambda_2(\mathbb{T}^d)$, and that when $\lambda > \lambda_2(\mathbb{T}^d)$, for every $\sigma < 1$, some $C < \infty$ and h sufficiently large,

$$\exp\left((\sigma d)^h\right) \leq \tau \leq \exp\left(C d^h\right).$$

Recall that $(d+1)^h$ is the number of vertices of the d-ary tree cut at height h.

What triggered our interest into this class of problems is the result of Chatterjee and Durrett [59] concerning the contact process on certain random graphs with power-law degree distribution. These random graphs are known as the configuration model, and are constructed as follows. To each vertex $x \in V := \{1, \dots, n\}$, assign d_x half-edges, where $(d_x)_{x \in V}$ are i.i.d. random variables such that $\mathbb{P}[d_x \geq r] \sim r^{-\alpha}$, with $\alpha \in (0, \infty)$. (We need to condition $\sum d_x$ to be even, but this conditioning is very mild.) Then choose a pairing of half-edges uniformly at random to form the graph.

The striking result of Chatterjee and Durrett is that the contact process is supercritical for every value of $\alpha > 0$ and $\lambda > 0$, in the sense that for each choice of these parameters and each $\delta > 0$, there exists $c > 0$ such that $\tau \geq \exp\left(cn^{1-\delta}\right)$. This is at variance with physicists' prediction of a non-trivial subcritical phase for $\alpha > 2$.

With Mountford, Valesin and Yao [MMVY16], we devised a general method enabling in particular to improve the stretched exponential lower bounds obtained in [59] for the configuration model and in [189] for finite regular trees. The approach is in two steps. The first step consists in showing that the survival time on basically *any* graph is at least exponential in the number of vertices if $\lambda > \lambda_2(\mathbb{Z})$. More precisely, we showed that given an integer d and $\lambda > \lambda_2(\mathbb{Z})$, there exists a constant $c > 0$ such that for every connected graph with n vertices and maximal degree bounded by d, the extinction time is at least $\exp(cn)$ with overwhelming probability. (It is easy to show that the extinction time is bounded by $\exp(Cn)$ with high probability.) Moreover, the random variable $\tau/\mathbb{E}[\tau]$ approaches an exponential random variable as n tends to infinity. This second property is a (weak) way to expose the metastability of the process: conditionally on survival, two processes started from possibly different initial configurations quickly become equal.

This result is interesting in that it shows, in a strong and precise sense, that the worst possible graph for the survival of the contact process is the line. The boundedness assumption on the degree is in practice not restrictive, since one can usually remove some edges of the graph to obtain one of bounded degree while maintaining connectedness.

This result already gives a stronger lower bound compared with the results of [59] or [190], but only under the overly restrictive assumption $\lambda > \lambda_2(\mathbb{Z})$. This is however not how the result is intended to be used. Instead, we should coarse-grain the contact process (as for standard percolation on $\mathbb{Z}^d$) into an effective contact process on a larger scale. During the coarse-graining, we should monitor an increase of the effective infection parameter, in line with the renormalization group picture for percolation. Our first result is useful because (1) it gives us a "target", $\lambda_2(\mathbb{Z})$, up to which it is no longer necessary to enlarge the coarse-graining scale; and (2) it shows that we do not need to keep track of the geometry of the coarse-grained graph, since basically any graph satisfies the bound. This last point is very important since beyond $\mathbb{Z}^d$, the operation of coarse-graining changes the geometry in complicated ways. Since we only need to go past the infection parameter $\lambda_2(\mathbb{Z})$, in practice, the coarse-graining scale should be taken sufficiently large but independent of n. Therefore, the volume of the coarse-grained graph is still a multiplicative constant away from the volume of the original graph.

We successfully implemented this strategy for the configuration model in [MMVY16], and for finite regular trees in [CMMV14]. In fact, we could be more precise in the latter case, and showed that the extinction time τ divided by the height of the tree converges in probability to a constant when $\lambda < \lambda_2(\mathbb{T}^d)$; that $\log \mathbb{E}[\tau]$ divided by the volume of the tree converges to a constant and that $\tau/\mathbb{E}[\tau]$ converges to an exponential random variable when $\lambda > \lambda_2(\mathbb{T}^d)$. More recent developments include [54, 53] for the configuration model and preferential attachment graphs, and [183] for arbitrary graphs.

With Valesin [MV16], we next considered the case of random $(d+1)$-regular graphs. As is well-known, these graphs locally look like $\mathbb{T}^d$. In this case, we discovered that the transition from short to long-time persistence of the infection occurs at the *lower* critical value $\lambda_1(\mathbb{T}^d)$, contrary to the case of regular trees cut at a given height. Intuitively, the infection tries to survive by "spreading to infinity", but in the random regular graph, the few cycles present there maintain this "spreading to infinity" within the graph. On the contrary, on a regular tree cut at a certain height, the infection also starts "spreading to infinity", but then hits the leaves of the tree and dies out.

The best way to understand the result is to realize that while the random $(d+1)$-regular tree locally resembles $\mathbb{T}^d$, the regular tree cut at a given height,

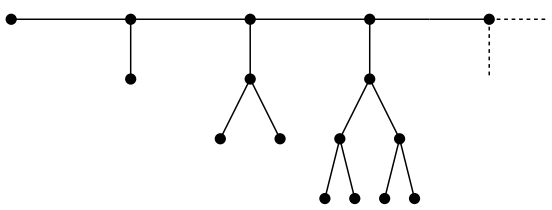

FIGURE 6.4. The binary canopy tree.

and observed locally from a vertex chosen uniformly at random, does *not*. Indeed, with non-vanishing probability, a vertex chosen uniformly at random will be a leaf (that is, a vertex of degree 1 instead of $d+1$); or it will be at distance 1 from a leaf; and so on. As made precise by Benjamini and Schramm [33] (see also [30, Example 5.14]), the local limit of the regular trees cut at increasingly large heights is a graph called the *canopy tree* $\mathbb{CT}^d$ (with a suitably defined random root), see Figure 6.4. Moreover, one can check that $\lambda_1(\mathbb{CT}^d) = \lambda_2(\mathbb{CT}^d) = \lambda_2(\mathbb{T}^d)$ (see the discussion surrounding [MV16, (1.2)]). Therefore, in every situation met so far, the critical infection parameter of the finite graphs is always the *lower* critical value of their local limit. It just so happens that for regular trees cut at a given height, this lower critical value coincides with the upper critical value of $\mathbb{T}^d$. Naturally, it would be very interesting to verify this phenomenon in greater generality. This question is similar to that of the locality of the critical percolation parameter, see [32], [30, Section 5.2], [155, 188].

CHAPTER II

Quantitative homogenization

1. Introduction

The goal of this chapter is to expose the additive approach to quantitative stochastic homogenization, as developed in [14], [AM16, AKM15, AKM16a, AKM16b] and briefly sketched in Subsection I.1.3. This first section presents one of the main results of [AKM16a]. The rest of the chapter aims at explaining the ideas and techniques entering the proof of this result. An important role is played by a regularity theory of solutions on large scales, which goes much beyond what would be true for arbitrary deterministic coefficient fields. In order to best explain the idea, we first review some aspects of the classical, deterministic regularity theory of elliptic equations in Section 2. We then state the large-scale regularity of solutions to equations with random coefficients in Section 3, following [14], [AM16], and reduce it to an error estimate with a possibly very poor rate of convergence but essentially optimal control of the probability that the estimate fails to hold. We then proceed to explain the proof of the latter result in Section 4 (departing slightly from the original argument). Sections 5 and 6 describe our first attempt [AKM15] at proving Theorem 1.1, which only provides with suboptimal results due to the presence of boundary layers. The last section describes the refined approach developed in [AKM16a, AKM16b] allowing to finally reach optimal results.

As was apparent throughout Section I.1, obtaining a fine understanding of the properties of the corrector is central to a quantitative theory of homogenization. In the approach presented here, energy quantities such as that appearing in (I.1.14) or (I.1.27) will play a central role. Once these are understood, it will be relatively straightforward to quantify the weak convergence of gradients and fluxes of the corrector, cf. (I.1.15), (I.1.9) and (I.1.13), and then control the sublinear growth of the corrector and improve on (I.1.10). We now introduce some notation and state the main result of [AKM16a] in this direction.

First, we need to measure spatial averages in a "smooth" way. We define the heat kernel

$$\Phi_{z,r}(x) := (4\pi r^2)^{-\frac{d}{2}} \exp\left(-\frac{|x-z|^2}{4r^2}\right), \tag{1.1}$$

and denote integrals against this heat kernel mask as

$$\int_{\Phi_{z,r}} f := \int_{\mathbb{R}^d} f(x)\,\Phi_{z,r}(x)\,\mathrm{d}x. \tag{1.2}$$

In words, $\int_{\Phi_{z,r}} f$ is a spatial average of f centered at z and of length scale r. We simply write $\int_{\Phi_r} f$ for $\int_{\Phi_{0,r}} f$.

We also introduce some convenient notation to measure the stochastic integrability of random variables. If X is a random variable and $s, \theta \in (0,\infty)$, then we define the statement

$$X \le \mathcal{O}_s(\theta)$$

to mean that

$$\mathbb{E}\left[\exp\left(\left(\theta^{-1}X_+\right)^s\right)\right] \le 2, \tag{1.3}$$

where $X_+ := X \vee 0$ is the positive part of X. This statement implies that the right tail of $\theta^{-1}X$ decay like $\exp(-c|x|^s)$ for some constant $c > 0$; the converse implication also holds with suitably modified constants.

Recall that we assume the coefficient field $\mathbf{a}$ to be $\mathbb{Z}^d$-stationary, to satisfy the uniform ellipticity condition (I.1.3) and the finite range of dependence condition (I.1.26). For convenience, we slightly change notation and denote by $\phi_p = \phi(\cdot, p)$ the corrector in the direction of $p \in \mathbb{R}^d$. Recall that ϕ_p is defined up to an additive constant as the unique function with $\mathbb{Z}^d$-stationary gradient satisfying (I.1.8) and (I.1.9). With Armstrong and Kuusi [AKM16a], we proved the following result.

Theorem 1.1 (Optimal estimates on correctors)**.** *For every $s < 2$, there exists $C < \infty$ such that, for every $r \ge 1$ and $p \in B_1$,*

$$\left| \int_{\Phi_r} \frac{1}{2} (p + \nabla\phi_p) \cdot \mathbf{a} (p + \nabla\phi_p) - \frac{1}{2} p \cdot \bar{\mathbf{a}} p \right| \le \mathcal{O}_s\left(C r^{-\frac{d}{2}}\right), \tag{1.4}$$

$$\left| \int_{\Phi_r} \nabla\phi_p \right| \le \mathcal{O}_s\left(C r^{-\frac{d}{2}}\right), \tag{1.5}$$

and

$$\left| \int_{\Phi_r} \mathbf{a}(p + \nabla\phi_p) - \bar{\mathbf{a}} p \right| \le \mathcal{O}_s\left(C r^{-\frac{d}{2}}\right). \tag{1.6}$$

Moreover, there exist $\varepsilon > 0$ and, for every $s < 2$, a constant $C < \infty$ such that, for each $r \ge 2$ and $p \in B_1$,

$$\left(\fint_{B_r} \left| \phi_p - \left(\fint_{B_r} \phi_p \right) \right|^2 \right)^{\frac{1}{2}} \le \begin{cases} \mathcal{O}_s\left(C \log^{\frac{1}{2}} r\right) & \text{if } d = 2, \\ \mathcal{O}_{2+\varepsilon}(C) & \text{if } d > 2. \end{cases} \tag{1.7}$$

Finally, if $d > 2$, then ϕ_p exists as a $\mathbb{Z}^d$-stationary random field.

The statements (1.4), (1.5) and (1.6) are all consistent with the idea that the fields under consideration (energy, gradient and flux fields respectively) rescale as random fields with a short range of correlation. The limitation $s < 2$ is easily understood in this light: by the central limit theorem, we can at best obtain Gaussian-type tail estimates, which correspond to the critical exponent $s = 2$. However, this intuition is only valid for the energy field. Indeed, refining the proof of Theorem 1.1, we showed in [AKM16b] that a central limit theorem holds for similar energy quantities, but that the law of the large-scale spatial averages of the corrector is instead governed by the gradient Gaussian free field described in (I.1.36). A similar result could be derived for the flux field. The idea that the corrector behaves like a Gaussian free field (regularized at the unit scale) is further corroborated by the characteristic $\log^{\frac{1}{2}}$ divergence of its oscillations in dimension $d = 2$, as displayed in (1.7).

2. Classical regularity theory

A core ingredient of the proof of Theorem 1.1 is a regularity theory for solutions to elliptic equations with random coefficients. For clarity of exposition, we review here the classical regularity theory of deterministic operators. The case of random coefficients and the importance of such regularity theory for our purpose will be discussed in the next sections. In this section and only here, *the coefficient field* $\mathbf{a}$ *is fixed and is only assumed to satisfy the uniform ellipticity condition (I.1.3).*

The simplest statement of regularity of $\mathbf{a}$-harmonic functions is certainly Caccioppoli's inequality. This inequality states that if $u \in H^1(B_{2r})$ satisfies

$$-\nabla \cdot \mathbf{a}\nabla u = 0 \qquad \text{in } B_{2r} \tag{2.1}$$

(in the weak sense), then

$$\int_{B_r} |\nabla u|^2 \leq \frac{C}{r^2} \inf_{c\in\mathbb{R}} \int_{B_{2r}} |u-c|^2, \tag{2.2}$$

where the constant C depends only on the ellipticity constant Λ. In order to justify this, we may assume that $r = 1$, by scaling. Let η be a smooth function which is constant equal to 1 on B_1 and vanishes outside of B_2. Testing the equation (2.1) against $u\eta^2$ and using the chain rule, we obtain

$$\int_{B_2} \eta^2 |\nabla u|^2 \leq \Lambda \int_{B_2} \eta |u| \, |\nabla \eta| \, |\nabla u|.$$

Using Hölder's inequality and $\|\nabla \eta\|_{L^\infty} < \infty$ yields

$$\int_{B_1} |\nabla u|^2 \leq C \int_{B_2} |u|^2,$$

which we can transform into (2.2) by scaling and subtraction of a constant. We chose the ratio of larger to smaller ball in (2.2) to be equal to 2, but this can of course be modified, at the price of a modification of the constant C appearing there. Caccioppoli's inequality is often suggestively called a reverse Poincaré inequality.

When $\mathbf{a}(x) = \bar{\mathbf{a}}$ is constant in space, derivatives of u are themselves $\bar{\mathbf{a}}$-harmonic functions. We can therefore apply Caccioppoli's inequality on the derivatives, and continue inductively to obtain that u is in every interior Sobolev space. By Morrey's embedding, we infer that u is C^∞ in the interior of B_{2r}. The following proposition is related to this statement. For each $k \in \mathbb{N}$, we denote by $\overline{\mathcal{A}}_k$ the space of $\bar{\mathbf{a}}$-harmonic polynomials of degree k.

Proposition 2.1 (Regularity for constant coefficients)**.** *Let $\bar{\mathbf{a}}$ be constant in space, and let u satisfy*

$$-\nabla \cdot \bar{\mathbf{a}}\nabla u = 0 \qquad \text{in } B_r.$$

For each $k \in \mathbb{N}$, there exists a constant $C(k,\Lambda) < \infty$ such that for every $0 < \rho \leq r$,

$$\inf_{p\in\overline{\mathcal{A}}_k} \left(\fint_{B_\rho} |u-p|^2 \right)^{\frac{1}{2}} \leq C \left(\frac{\rho}{r}\right)^{k+1} \inf_{p\in\overline{\mathcal{A}}_k} \left(\fint_{B_r} |u-p|^2 \right)^{\frac{1}{2}}. \tag{2.3}$$

Sketch of proof. We only sketch the proof for $k = 0$. For $\rho \in [r/2, r]$, it suffices to adjust the constant C to ensure that (2.3) holds. We may therefore suppose that $\rho \leq r/2$. By the observations preceding the statement of the proposition and scaling, there exists $C(\Lambda) < \infty$ such that

$$\|\nabla u\|^2_{L^\infty(B_{r/2})} \leq \frac{C}{r^2} \inf_{c\in\mathbb{R}} \fint_{B_r} |u-c|^2. \tag{2.4}$$

Moreover,

$$\inf_{c\in\mathbb{R}} \fint_{B_\rho} |u-c|^2 \leq \fint_{B_\rho} |u-u(0)|^2 \leq \rho^2 \|\nabla u\|^2_{L^\infty(B_\rho)}.$$

The result follows by combining these two inequalities. □

A consequence of the case $k = 0$ of this proposition is Liouville's theorem: if a function is $\bar{\mathbf{a}}$-harmonic on $\mathbb{R}^d$ and bounded, then it is constant. More generally, if a function is $\bar{\mathbf{a}}$-harmonic on $\mathbb{R}^d$ and grows at most polynomially, then it is in fact an $\bar{\mathbf{a}}$-harmonic polynomial.

For a general space-dependent coefficient field $\mathbf{a}(x)$, solutions are not necessarily smooth. Given $\alpha > 0$, one can find a coefficient field $\mathbf{a}(x)$ (with an ellipticity

constant Λ that diverges as $\alpha \to 0$) and an $\mathbf{a}$-harmonic function which is not α-Hölder continuous, see [75, p. 364] or [160, Section 5]. The converse to this statement is the celebrated De Giorgi-Nash-Moser theory [72, 167], which states that for a given ellipticity constant Λ, there exists an exponent $\alpha > 0$ such that every $\mathbf{a}$-harmonic function is α-Hölder continuous. This statement is false for uniformly elliptic systems of equations, see [106, Section 9.1]. We will not make use of the De Giorgi-Nash-Moser theory, and only rely on arguments that also apply to systems of equations.

Despite these negative results, $\mathbf{a}$-harmonic functions do have some regularity if the coefficient field $\mathbf{a}(x)$ does. Before presenting this result, we recall the following characterization of Hölder-continuous functions, due to Campanato [51]. If for some $M < \infty$ and $\alpha \in (0,1)$, a function $u \in L^2(B_1)$ satisfies

$$\inf_{c\in\mathbb{R}} \left(\fint_{B_r(x)} |u-c|^2 \right)^{\frac{1}{2}} \leq M r^\alpha \qquad \text{for every } B_r(x) \subseteq B_1, \tag{2.5}$$

then u is α-Hölder continuous in the interior of B_1 (see [133, Theorem 3.1] for a proof). The result can be generalized to higher regularity. Indeed, let $\mathcal{P}_k$ denote the set of polynomials of degree k. If a function u satisfies

$$\inf_{p\in\mathcal{P}_k} \left(\fint_{B_r(x)} |u-p|^2 \right)^{\frac{1}{2}} \leq M r^{k+\alpha} \qquad \text{for every } B_r(x) \subseteq B_1,$$

then u is k times differentiable with α-Hölder continuous k-th derivative (that is, it is of class $C^{k,\alpha}$).

The following proposition is a Schauder estimate. It is stated in a qualitative form, but it can of course be turned into an estimate of the interior $C^{1,\alpha}$ norm.

Proposition 2.2 (Schauder estimate)**.** *Let $\mathbf{a}(x)$ be an α-Hölder continuous coefficient field, $\alpha \in (0,1)$. If u is an $\mathbf{a}$-harmonic function on B_2, then it is $C^{1,\alpha}$ in the interior.*

Sketch of proof. We give a sketch of proof based on a Campanato iteration [52]. We refer to [133, Theorem 3.13] for more precision.

We fix $x_0 \in B_1$, and for r sufficiently small, we compare u in $B_r(x_0)$ with w satisfying

$$-\nabla \cdot \mathbf{a}(x_0) \nabla w = 0,$$

with $w = v$ on $\partial B_r(x_0)$. Since w solves a homogeneous equation, it satisfies the "improvement of flatness" property given in Proposition 2.1. In order to transfer this regularity property to u, we need to show that the difference $v := u - w$ is small. This difference satisfies

$$-\nabla \cdot \mathbf{a}(x_0) \nabla v = \nabla \cdot (\mathbf{a}(x_0) - \mathbf{a}) \nabla u$$

on $B_r(x_0)$, with null Dirichlet boundary condition. By testing the equation against v, using Young's inequality and the regularity assumption on $\mathbf{a}$, we get

$$\int_{B_r(x_0)} |\nabla v|^2 \leq C r^{2\alpha} \int_{B_r(x_0)} |\nabla u|^2. \tag{2.6}$$

Combining this with Proposition 2.1 and Caccioppoli's inequality, we obtain that

$$\int_{B_\rho(x_0)} |\nabla u|^2 \leq C \left(\frac{\rho}{r}\right)^d \int_{B_r(x_0)} |\nabla u|^2 + C r^{2\alpha} \int_{B_r(x_0)} |\nabla u|^2, \tag{2.7}$$

as well as

$$\inf_{c\in\mathbb{R}} \int_{B_\rho(x_0)} |\nabla u - c|^2 \leq C \left(\frac{\rho}{r}\right)^{d+2} \inf_{c\in\mathbb{R}} \int_{B_r(x_0)} |\nabla u - c|^2 + C r^{2\alpha} \int_{B_r(x_0)} |\nabla u|^2. \tag{2.8}$$

Denoting $F(\rho) := \fint_{B_\rho(x_0)} |\nabla u|^2$, we may rewrite (2.7) as

$$F(\rho) \le C\left[\left(\frac{\rho}{r}\right)^d + r^{2\alpha}\right] F(r).$$

From this inequality, an induction argument (see [133, Lemma 3.4]) shows that for every $\delta > 0$, there exists a constant C such that for every $\rho \le r$,

$$F(\rho) \le C\left(\frac{\rho}{r}\right)^{d-\delta} F(r).$$

Denoting $G(\rho) := \inf_{c\in\mathbb{R}} \fint_{B_\rho(x_0)} |\nabla u - c|^2$, we deduce from this and (2.8) that

$$G(\rho) \le C\left(\frac{\rho}{r}\right)^{d+2} G(r) + Cr^{d+2\alpha-\delta} F(1).$$

Applying the induction argument [133, Lemma 3.4] once more, we obtain that

$$F(\rho) \le Cr^{d+2\alpha-\delta} F(1) \le Cr^{d+2\alpha-\delta} \fint_{B_2} |\nabla u|^2.$$

By the characterization (2.5), we deduce that ∇u is $(\alpha - \frac{\delta}{2})$-Hölder continuous. In particular, we have $F(r) \lesssim r^d$, and repeating the argument with this improved bound shows that ∇u is in fact α-Hölder continuous. □

3. Homogenization and regularity theory

We now resume our study of equations with random coefficients. From now on, we assume that the coefficient field $\mathbf{a}(x)$, in addition to satisfying the uniform ellipticity assumption (I.1.3), is a random, $\mathbb{Z}^d$-stationary field satisfying the finite range of dependence condition (I.1.26).

In a discrete setting with i.i.d. coefficients, Marahrens and Otto [153] proved optimal moment estimates on the derivatives and second mixed derivatives of the Green function, see (I.1.19) and (I.1.20), using a logarithmic Sobolev inequality with respect to the random coefficient field. They then deduced that solutions to random equations are essentially α-Hölder continuous on large scales, for every $\alpha < 1$. (The precise meaning of this will be clarified shortly.) In particular, solutions to random equations are much more regular than solutions to generic equations, which are only expected to be α-Hölder continous for a possibly very small exponent $\alpha > 0$.

Inspired by this and earlier results, Armstrong and Smart [14] opted to reverse this point of view, and to make a regularity theory for general solutions central to the theory. Gradient bounds on the Green function would then be simple consequences of the regularity theory. They were also motivated by the desire to develop quantitative homogenization for possibly nonlinear equations.

The idea to place an improved regularity theory at the centre of a quantitative theory of homogenization echoes eariler work of Avellaneda and Lin [17] in the setting of linear equations and systems with periodic coefficients. There, a compactness argument allows to prove the sought-after regularity theory. This argument does not apply for random coefficients.

An $\mathbf{a}$-harmonic function on a bounded domain U can be seen as the minimizer of the convex integral functional

$$u \mapsto \fint_U \frac{1}{2} \nabla u \cdot \mathbf{a} \nabla u,$$

over functions in $H^1(U)$ with suitable boundary condition. Armstrong and Smart considered the more general setting of homogenizing convex integral functionals (assuming suitable lower and upper bounds on convexity). Although their results

hold in this general situation, we will only state them for linear equations for simplicity. Their first main result is an error estimate for Dirichlet problems.

Theorem 3.1 (Error estimate [14]). *Fix a bounded Lipschitz domain U and an exponent $s \in (0,d)$. There exists an exponent $\delta > 0$ and a random variable $\mathcal{X} \geq 0$ satisfying*

$$\mathcal{X} \leq \mathcal{O}_s(C)$$

such that for every $r \geq \mathcal{X}$, if $u, \overline{u}$ are respectively solutions to

$$-\nabla \cdot (\mathbf{a}\nabla u) = 0 \quad \text{and} \quad -\nabla \cdot (\bar{\mathbf{a}}\nabla \overline{u}) = 0 \quad \text{in } rU,$$

with $u - \overline{u} \in H^1_0(rU)$, then we have the estimate

$$\frac{1}{r}\left(\fint_{rU} |u-\overline{u}|^2\right)^{\frac{1}{2}} \leq r^{-\delta}\left(\fint_{rU} |\nabla \overline{u}|^{2+}\right)^{\frac{1}{2+}}. \tag{3.1}$$

The exponent "2+" in the right side of (3.1) refers to any exponent above 2; the random variable $\mathcal{X}$ depends on this choice of exponent. This is an unimportant technical point necessary to control the behavior of the solution near the boundary. In this theorem and throughout this chapter, we will prefer to depart from the point of view of fixing a domain and shrinking the coefficient field as in (I.1.4), and consider instead that the coefficient field is fixed and that we are blowing up the domain. Of course, the two points of view are mathematically equivalent. In the former point of view, the theorem above can be roughly restated as

$$\mathbb{P}\left[\left(\fint_U |u_\varepsilon - \overline{u}|^2\right)^{\frac{1}{2}} \geq \varepsilon^\delta\right] \lesssim \exp\left(-\varepsilon^{-s}\right). \tag{3.2}$$

The exponent δ measuring the error is possibly very small; in contrast, the exponent s measuring the stochatic integrability is essentially optimal. Indeed, if the law of the coefficient field is that of a random checkerboard as on Figure I.1.1, then no error estimate can be expected on the event that "every square is black". This event has probability $\simeq \exp\left(-\varepsilon^{-d}\right)$.

Theorem 3.1 allows to deduce a regularity theory on large scales, for reasons that are best explained by analogy with the classical Campanato approach to Schauder estimates explained in the previous section. There, we used the Hölder continuity of the coefficient field to compare the solution to the heterogeneous equation with that of a homogeneous one. We then relied on the regularity properties of the latter to infer "improvement of flatness" properties of the solution to the heterogeneous equation, up to an error which becomes smaller as we consider smaller and smaller scales. An inductive argument based on this estimate then enabled to prove the Hölder continuity of the derivative, through Campanato's characterization (2.5).

In the context of stochastic homogenization, the situation is similar, except that we compare the solution to the heterogeneous equation to that of the homogenized equation. However, this time the error we make becomes smaller on *larger* scales, as homogenization improves. Theorem 3.1 quantifies this error. In the Schauder estimate (Proposition 2.2), any small exponent of Hölder regularity suffices to obtain Lipschitz regularity of the solution. By analogy, the smallness of $\delta > 0$ will not affect the proof of Lipschitz regularity of solutions on large scales. However, it is indeed crucial that the rate of convergence be *some* power of ε (more precisely, the iteration requires that the sum of errors on dyadic scales be finite). The iteration stops at the random scale $\mathcal{X}$ below which the error estimate (3.1) no longer holds. The Lipschitz regularity on large scales roughly states the following.

Theorem 3.2 ($C^{0,1}$ regularity [14]). *Let $s < d$. There exists a random variable $\mathcal{X}$ satisfying*

$$\mathcal{X} \leq \mathcal{O}_s(C)$$

such that if $R \geq \mathcal{X}$ *and* u *is a solution to*

$$-\nabla \cdot \mathbf{a} \nabla u = 0 \qquad \text{in } B_R,$$

then for every $r \in [\mathcal{X}, R]$,

$$\inf_{c \in \mathbb{R}} \left(\fint_{B_r} |u-c|^2 \right)^{\frac{1}{2}} \leq C \left(\frac{r}{R} \right) \inf_{c \in \mathbb{R}} \left(\fint_{B_R} |u-c|^2 \right)^{\frac{1}{2}}.$$

By Caccioppoli's and Poincaré's inequalities, we can rephrase the latter inequality as

$$\fint_{B_r} |\nabla u|^2 \leq C \fint_{B_R} |\nabla u|^2.$$

In other words (or by comparing with (2.5)), the solution u is essentially Lipschitz, down to the random scale $\mathcal{X}$, which is essentially of unit size.

We may again rephrase the statement in terms of a rescaled coefficient field in a fixed domain, as in (3.2). By Proposition 2.2, if the coefficient field is α-Hölder continuous, then an $\mathbf{a}$-harmonic function satisfies

$$|\nabla u(0)| \leq C \left(\int_{B_1} |\nabla u|^2 \right)^{\frac{1}{2}}. \tag{3.3}$$

However, if we rescale the coefficient field from $\mathbf{a}$ to $\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)$, then the estimate will deteriorate further and further. In contrast, Theorem 3.2 shows that this Lipschitz estimate remains correct for solutions to random equations, down to a random scale of size essentially ε.

4. Subadditive quantities

As mentioned in Subsection I.1.3, the approach to the proof of Theorem 3.1 is inspired by the work of Dal Maso and Modica [69, 70], and emphasizes the role played by the energy quantity ν. We slightly reformulate the definition in (I.1.27) as

$$\nu(U,p) := \inf_{v \in \ell_p + H^1_0(U)} \fint_U \frac{1}{2} \nabla v \cdot \mathbf{a} \nabla v, \tag{4.1}$$

where ℓ_p denotes the affine function $x \mapsto p \cdot x$. We denote by $v(\cdot, U, p)$ the unique minimizer of this variational problem. Recall that this quantity is subadditive in the sense of (I.1.29), and therefore, by the subadditive ergodic theorem,

$$\nu(\square_r, p) \xrightarrow[R \to \infty]{\text{a.s.}} \frac{1}{2} p \cdot \bar{\mathbf{a}} p. \tag{4.2}$$

The main difficulty consists in quantifying this convergence. The idea of [14] is to introduce another energy-type quantity, which in a sense is dual to ν. For every bounded Lipschitz domain U and $q \in \mathbb{R}^d$, we let

$$\mu(U,q) := \inf_{u \in H^1(U)} \fint_U \left(\frac{1}{2} \nabla u \cdot \mathbf{a} \nabla u - q \cdot \nabla u \right). \tag{4.3}$$

Standard arguments ensure the existence and uniqueness of the minimizer of this variational problem, up to an additive constant which we fix by requiring the minimizer to have zero mean. We denote by $u(\cdot, U, q)$ this minimizer. This minimizer is the harmonic function with affine Neumann boundary condition given by q.

The quantity μ is superadditive. Indeed, a minimizer on a large domain provides us with a minimizer candidate on a subdomain, by restriction. Moreover, using $v(\cdot, U, p)$ as a minimizer candidate for μ, we obtain that for every $p, q \in \mathbb{R}^d$,

$$\mu(U,q) \leq \nu(U,p) - p \cdot q. \tag{4.4}$$

In other words, the quantity $-\mu(U, \cdot)$ is always larger than the convex dual of the (convex) function $\nu(U, \cdot)$. Moreover, if $\mathbf{a}$ were constant, then the minimizers for μ and ν would both be affine functions, and $-\mu(U, \cdot)$ would be equal to the convex

dual of $\nu(U,\cdot)$. It is therefore reasonable to expect that $-\mu(U,\cdot)$ asymptotically approaches the convex dual of $\nu(U,\cdot)$, and therefore that

$$-\mu(\square_r,p) \xrightarrow[R\to\infty]{\text{a.s.}} \frac{1}{2} q\cdot \bar{\mathbf{a}}^{-1} q, \tag{4.5}$$

since the right side is the convex dual of the right side of (4.2). The proof of Theorem 3.1 consists in showing that as μ and ν start to stabilize, their minimizers become asymptotically flatter and flatter on large scales, so that $-\mu$ and ν do become asymptotically dual to one another, and therefore must be close to their limits. If they are close to one another, then they are additive up to a small error, and we can use standard tools of probability to bound their fluctuations optimally.

We now give a more precise sketch of Theorem 3.1. While we stay relatively close to the proof exposed in [14], the details differ and are informed by subsequent work. The main work consists in checking that the expectation of μ and ν converge sufficiently fast. For each $n\in\mathbb{N}$, define

$$\tau_n := \sup_{q\in B_1} \left(\left|\mathbb{E}[\mu(\square_{3^{n+1}},q)] - \mathbb{E}[\mu(\square_{3^n},q)]\right| + \left|\mathbb{E}[\nu(\square_{3^{n+1}},q)] - \mathbb{E}[\nu(\square_{3^n},q)]\right| \right).$$

By the ergodic theorem, we know that τ_n tends to 0 as n tends to infinity. Note that τ_n measures the size of the difference between two scales, not the distance to the limit. Roughly speaking, the idea of the proof of Theorem 3.1 consists in the following steps.

(1) For $z+\square_{3^n}\subseteq\square_{3^{n+1}}$, the quantity τ_n controls the quantity

$$\mathbb{E}\left[\fint_{z+\square_{3^n}} |\nabla u(\cdot,\square_{3^{n+1}},q) - \nabla u(\cdot,z+\square_{3^n},q)|^2 \right],$$

by uniform convexity. Using independence (see [14, Lemma 3.2]), we deduce that the variance of the spatial average

$$\fint_{\square_{3^n}} \nabla u(\cdot,\square_{3^{n+1}},q)$$

is controlled by τ_n. We denote the expectation of this random variable by $\overline{P}_n q$ (the mapping $q\mapsto \overline{P}_n q$ is linear).

(2) We use the same argument to assert that the spatial average of $\nabla v(\cdot,\square_{3^{n+1}},p)$ has small fluctuations. Its expectation is essentially p.

(3) We use the multiscale Poincaré inequality described below to deduce that the minimizer $u(\cdot,\square_{3^{n+1}},q)$ itself is close to the affine function $x\mapsto (\overline{P}_n q)\cdot x$; and that the minimizer $v(\cdot,\square_{3^{n+1}},p)$ is close to the affine function $x\mapsto p\cdot x$. Here, "close" involves a weighted average of the $(\tau_k)_{k\le n}$, which essentially behaves like τ_n.

(4) We deduce that for each given $q\in B_1$,

$$3^{-2n}\mathbb{E}\left[\fint_{\square_{3^n}} \left| u(\cdot,\square_{3^{n+1}},q) - v(\cdot,\square_{3^{n+1}},\overline{P}_n q)\right|^2 \right] \tag{4.6}$$

is small ("small" in the same sense as "close" above, essentially τ_n). By Caccioppoli's inequality, we infer that the gradients of $u(\cdot,\square_{3^{n+1}},q)$ and $v(\cdot,\square_{3^{n+1}},\overline{P}_n q)$ are close to one another, and therefore that

$$\inf_{p\in\mathbb{R}^d} \left(\mathbb{E}[\nu(\square_{3^n},p)] - p\cdot q - \mathbb{E}[\mu(\square_{3^n},q)] \right)$$

is small. Roughly speaking, we thus obtained a control of the difference between ν and its limit (and of the difference between μ and its limit) in terms of the difference of ν's between two scales. Write $s_n := \mathbb{E}[\nu(\square_{3^n},p)]$,

which decreases to its limit s_∞. The recursive relation we obtain is a slightly more involved version of the relation

$$0 \le s_n - s_\infty \le C(s_n - s_{n+1}).$$

This is similar to the differential inequality $y' \le -y/C$, with $y(n) = s_n - s_\infty$. We therefore deduce that $\tau_n \simeq s_n - s_\infty \lesssim 3^{-\delta n}$ for some $\delta > 0$ (with $\delta \simeq C^{-1}$). In other words, we conclude that there exists $\delta > 0$ and $C < \infty$ such that for every $r \ge 1$ and $p \in B_1$,

$$0 \le \mathbb{E}[\nu(\square_r, p)] - \frac{1}{2} p \cdot \bar{\mathbf{a}} p \le C r^{-\delta}, \tag{4.7}$$

and for every $r \ge 1$ and $q \in B_1$,

$$0 \le -\mathbb{E}[\mu(\square_r, q)] - \frac{1}{2} q \cdot \bar{\mathbf{a}}^{-1} q \le C r^{-\delta}. \tag{4.8}$$

An important ingredient in the outline above is a "multiscale" Poincaré inequality. This inequality estimates the oscillation of a function in terms of the spatial averages of its gradient. For a bounded domain U, and we write

$$(u)_U := \fint_U u$$

to denote the spatial average of u over the domain U.

Proposition 4.1 (Multiscale Poincaré inequality)**.** *Fix $m \in \mathbb{N}$ and, for each $n \in \mathbb{N}$, $n \le m$, define $\mathcal{Z}_n := 3^n \mathbb{Z}^d \cap \square_{3^m}$. There exists a constant $C(d) < \infty$ such that, for every $u \in H^1(\square_{3^m})$,*

$$\left(\fint_{\square_{3^m}} |u - (u)_{\square_{3^m}}|^2 \right)^{\frac{1}{2}} \le C \left(\fint_{\square_{3^m}} |\nabla u|^2 \right)^{\frac{1}{2}} + C \sum_{n=0}^{m-1} 3^n \left(|\mathcal{Z}_n|^{-1} \sum_{y \in \mathcal{Z}_n} \left| (\nabla u)_{y + \square_{3^n}} \right|^2 \right)^{\frac{1}{2}}.$$

This inequality first appeared in [AKM15, Proposition 6.1]. Recall that the usual Poincaré inequality states that

$$\left(\fint_{\square_{3^m}} |u - (u)_{\square_{3^m}}|^2 \right)^{\frac{1}{2}} \le C 3^m \left(\fint_{\square_{3^m}} |\nabla u|^2 \right)^{\frac{1}{2}}.$$

This estimate is optimal when u is affine. Yet, Proposition 4.1 is a significant improvement over the standard Poincaré inequality when the large-scale spatial averages of the gradient are small. In the outline above, we apply this proposition to the functions $x \mapsto u(x, \square_{3^{n+1}}, q) - (\overline{P}_n q) \cdot x$ and $x \mapsto v(x, \square_{3^{n+1}}, p) - p \cdot x$.

The proof of Proposition 4.1 rests on the fact that the right side of the inequality in the statement controls the H^{-1} norm of ∇u. This can be intuitively understood from the definitions of Besov spaces based on wavelet or Fourier decompositions, see (III.4.7) with $q = 2$ and $\alpha = -1$. (To dispel the possible confusion over the sign of α, note that in Proposition 4.1, the length scale increases with n, while it decreases with k in (III.4.7).)

Once the expectations are controlled, we can apply standard probabilistic tools to control sums of independent random variables and conclude the argument. We refer for instance to [AKM15, Theorem 5.1] for a general statement allowing to deduce strong stochastic control of μ and ν from (4.7) and (4.8). Incidentally, this provides with a nice converse to the result of Alexander [4], which states that for a subadditive sequence, the size of the bias (the difference between the expectation and its limit) is controlled by the size of the fluctuations. The key point of our argument is that we are given *two* quantities, one subadditive ($\nu(\cdot, p)$) and the other superadditive ($\mu(\cdot, \bar{\mathbf{a}} p) + p \cdot \bar{\mathbf{a}} p$), which are ordered ((4.4)) and asymptotically converge to the same limit.

The fact that the multiscale Poincaré inequality can be used to simplify part of the arguments of [14], [AM16] was developed in the master's thesis of Alexandre Bordas (under my supervision).

As was pointed out, the approach of [14] covers possibly nonlinear equations derived from convex minimization problems. However, in the linear setting, it seems a priori restricted to the case of random fields $\mathbf{a}$ taking values in symmetric matrices. Yet, Armstrong and I [AM16] could extend the approach to general quasilinear equations and systems in divergence form, in particular covering the case of nonsymmetric matrices. Our approach relies on a perhaps underapreciated variational representation for arbitrary maximal monotone maps due to Fitzpatrick [94, 102]. Moreover, we substantially relaxed the independence assumption on the coefficients, allowing instead for stretched exponential or polynomial mixing conditions. This modifies the stochastic integrability of the "minimal radius" $\mathcal{X}$ in Theorems 3.1 and 3.2 accordingly. One motivation for extending the result to possibly nonsymmetric matrices is that it opens the possibility to apply the approach to parabolic equations. Indeed, the time derivative essentially acts as an antisymmetric matrix. The fact that parabolic equations admit a variational structure was discovered by Brézis and Ekeland [48].

It is straightforward to deduce gradient Green function estimates from Theorem 3.2. Indeed, it suffices to observe that the function $x \mapsto G(0,x)$ is $\mathbf{a}$-harmonic outside of the origin, apply the regularity result in $x + B_{|x|/2}$ and use the quenched upper bound $G(0,x) \lesssim |x|^{-(d-2)}$ to obtain the bound (I.1.19) with a much better, exponential control of the random variable, instead of polynomial moments. (Strictly speaking, we need to average the gradient over a unit ball due to the possible local irregularity of the Green function. Alternatively, one may assume that the coefficient field $\mathbf{a}$ is Hölder continuous and use the Schauder estimate, see (3.3).) Estimates on the second mixed derivative of the Green function (I.1.20) are obtained similarly, since the function $x \mapsto \nabla_1 G(0,x)$ is $\mathbf{a}$-harmonic outside of the origin, and its size is estimated by the previous argument.

The work of [14] inspired Gloria, Neukamm and Otto to develop their own approach to the regularity theory of linear equations and systems with random coefficients [114]. Their result differs from that of [14], [AM16] in that their $C^{1,1}$ regularity theory is formulated in a more intrinsic way, using correctors (as was already the case of the work of [17] in the periodic setting). We will return to this below. This allows them to define a finite minimal radius similar to the random variable $\mathcal{X}$ above under the mere assumption of ergodicity (without mixing). Using this construction, they obtain that there are no subquadratic $\mathbf{a}$-harmonic functions besides the functions $x \mapsto p \cdot x + \phi_p(x) + c$, assuming only ergodicity. Next, assuming sufficient structure on the correlations of the coefficient field, they resort to "nonlinear" concentration inequalities to estimate the stochastic integrability of the minimal radius, and deduce exponential-type estimates on the corrector (with a suboptimal exponent). The "intrinsic" point of view to regularity theory was then generalized to higher order by Fischer and Otto [93].

5. Higher-order regularity

The works [14], [AM16] were a promising first step towards a theory of quantitative homogenization devoid of "nonlinear" concentration inequalities. With Armstrong and Kuusi, we then focused on completing this program. Our first attempt rested on the following $C^{k,1}$ higher regularity result, which can be stated as follows (see [AKM15, Theorem 3.1]). Recall that $\overline{\mathcal{A}}_k$ denotes the set of $\bar{\mathbf{a}}$-harmonic polynomials of degree k.

Theorem 5.1 (Extrinsic higher-order regularity)**.** *For every $s < d$, there exists an exponent $\delta > 0$ and a random variable $\mathcal{X}$ satisfying*

$$\mathcal{X} \le \mathcal{O}_s(C)$$

such that if $R \ge \mathcal{X}$ and u is a solution to

$$-\nabla \cdot \mathbf{a} \nabla u = 0 \qquad \text{in } B_R,$$

then for every $r \in [\mathcal{X} \vee R^{\frac{k}{k+\delta}}, R]$,

$$\inf_{p \in \overline{\mathcal{A}}_k} \left(\fint_{B_r} |u-p|^2 \right)^{\frac{1}{2}} \le C \left(\frac{r}{R} \right)^{k+1} \left(\fint_{B_R} |u|^2 \right)^{\frac{1}{2}}.$$

This result is obtained along similar lines as Theorem 3.2, combining Theorem 3.1 and Proposition 2.1 through a Campanato iteration. Note that contrary to the case $k = 0$, this higher regularity is only shown to hold from the mesoscopic scale $R^{\frac{k}{k+\delta}}$ upwards. This is due to the fact that an element p of $\overline{\mathcal{A}}_k$ is not $\mathbf{a}$-harmonic for $k \ge 1$, so we cannot subtract it and continue the induction with u replaced by $u - p$. The hope was to use this result to deduce optimal error estimates allowing to replace δ by 1 in Theorem 3.1, which in turn allows to deduce higher regularity with $R^{\frac{k}{k+\delta}}$ replaced by the optimal scale $R^{\frac{k}{k+1}}$.

This point of view was more fruitful for our first attempt [AKM15], which we will review shortly. Our final proof of Theorem 1.1 [AKM16a] relied rather on "intrinsic" regularity theory, similarly to [17, 113, 93]. Let

$$\mathcal{A}_k := \left\{ u \in H^1_{\mathrm{loc}}(\mathbb{R}^d) \text{ is } \mathbf{a}\text{-harmonic and such that } \lim_{r \to \infty} r^{-(k+1)} \left(\fint_{B_r} |u|^2 \right)^{\frac{1}{2}} = 0 \right\}.$$

In words, $\mathcal{A}_0$ is the set of sublinear $\mathbf{a}$-harmonic functions, $\mathcal{A}_1$ the set of subquadratic $\mathbf{a}$-harmonic functions, etc.

Theorem 5.2 (Intrinsic higher-order regularity)**.** *For every $s < d$, there exists a random variable $\mathcal{X}$ satisfying*

$$\mathcal{X} \le \mathcal{O}_s(C)$$

such that for each $k \in \mathbb{N}$, if $R \ge \mathcal{X}$ and u is a solution to

$$-\nabla \cdot \mathbf{a} \nabla u = 0 \qquad \text{in } B_R,$$

then for every $r \in [\mathcal{X}, R]$,

$$\inf_{\psi \in \mathcal{A}_k} \left(\fint_{B_r} |u-\psi|^2 \right)^{\frac{1}{2}} \le C \left(\frac{r}{R} \right)^{k+1} \inf_{\psi \in \mathcal{A}_k} \left(\fint_{B_R} |u-\psi|^2 \right)^{\frac{1}{2}}.$$

Theorems 5.1 and 5.2 contain different information regarding the behavior of solutions. (In particular, Theorem 5.2 tells us nothing about the large-scale $C^{k,1}$ regularity of elements of $\mathcal{A}_j$ if $j \le k$.) [AKM16a, Proposition 3.1] is slightly stronger than each of these theorems, and is proved using similar ideas. In particular, it also shows that the vector spaces $\mathcal{A}_k$ and $\overline{\mathcal{A}}_k$ have the same dimension.

6. The master quantity J

Let us start by summarizing and making slightly more precise what has been achieved at this point. We have argued that for every $s < d$, there exists $\alpha > 0$ and a constant $C < \infty$ such that for every $r \ge 1$ and $p, q \in B_1$,

$$0 \le \nu(\square_r, p) - \frac{1}{2} p \cdot \bar{\mathbf{a}} p \le \mathcal{O}_s\left(C r^{-\alpha}\right) \tag{6.1}$$

and

$$0 \le -\mu(\square_r, q) - \frac{1}{2} q \cdot \bar{\mathbf{a}}^{-1} q \le \mathcal{O}_s\left(C r^{-\alpha}\right). \tag{6.2}$$

From this, we deduced the error estimate given by Theorem 3.1 and then the regularity results of Theorems 5.1 and 5.2. The goal now is to use these additional ingredients to improve on the value of the exponent α in (6.1) and (6.2). We allow ourselves to lower the value of the exponent s as we increase α, in coherence with the simple fact that for every $s < s' \in (0,\infty)$ and every random variable X,

$$\left\{ \begin{array}{l} X \text{ takes values in } [0,1] \\ X \le \mathcal{O}_s(\theta) \end{array} \right. \quad \Longrightarrow \quad X \le \mathcal{O}_{s'}\left(\theta^{\frac{s}{s'}}\right), \tag{6.3}$$

see [AKM16a, Remark 2.1]. (It is elementary to check that $\nu(\cdot,p)$ and $\mu(\cdot,q)$ are bounded by a deterministic constant, uniformly over $p,q \in B_1$.)

In view of the outline of the proof of Theorem 3.1, it is natural to focus our study on the quantity

$$J(U,p,q) := \nu(U,p) - \mu(U,q) - p\cdot q \ge 0. \tag{6.4}$$

Probably the most important contribution of our first attempt to prove Theorem 1.1 is the realization that the quantity J admits a variational representation itself. Denote by $\mathcal{A}(U)$ the set of $\mathbf{a}$-harmonic functions in U:

$$\mathcal{A}(U) := \left\{ u \in H^1_{\rm loc}(U) \ : \ \forall \psi \in H^1_0(U), \ \int_U \nabla\psi\cdot\mathbf{a}\nabla u = 0 \right\}.$$

By definition, the set $\mathcal{A}(U)$ is the orthogonal complement of $H^1_0(U)$ for the scalar product $(u,v) \mapsto \int_U \nabla u\cdot\mathbf{a}\nabla v$. The minimizer of $\nu(U,p)$ is the element of $\ell_p + H^1_0(U)$ with minimal norm for this scalar product, see (4.1). By duality, it is therefore also the element of $\mathcal{A}(U)$ closest to ℓ_p. That is, maximizing

$$u \mapsto -\fint_U \frac12 \nabla(u-\ell_p)\cdot\mathbf{a}\nabla(u-\ell_p) = \fint_U \left(-\frac12\nabla u\cdot\mathbf{a}\nabla u + p\cdot\mathbf{a}\nabla u\right) + \text{cst}$$

over $u \in \mathcal{A}(U)$ selects $u(\cdot,U,p)$ as the maximizer. It is not difficult to check that in fact,

$$\nu(U,p) = \sup_{u\in\mathcal{A}(U)} \fint_U \left(-\frac12\nabla u\cdot\mathbf{a}\nabla u + p\cdot\mathbf{a}\nabla u\right),$$

and the maximizer equals $u(\cdot,U,p)$. An interesting aspect of this representation is that the boundary condition is encoded in a more implicit way. The crucial point is that a similar analysis can be carried out for the quantity J, and reveals that

$$J(U,p,q) = \sup_{w\in\mathcal{A}(U)} \fint_U \left(-\frac12\nabla w\cdot\mathbf{a}\nabla w - p\cdot\mathbf{a}\nabla w + q\cdot\nabla w\right), \tag{6.5}$$

see [AKM15, Lemma 3.1] for a complete proof. Moreover, the maximizer of this variational problem (unique up to an additive constant), which we denote by $w(\cdot,U,p,q)$, satisfies

$$w(\cdot,U,p,q) = u(\cdot,U,q) - v(\cdot,U,p). \tag{6.6}$$

(Recall that $-v(\cdot,U,p) = v(\cdot,U,-p)$ to connect with previous observations.) The quantity J is the cornerstone of our analysis. It is nonnegative and subadditive. We aim to show that it becomes almost additive as we reach to larger and larger scales. The quantity J contains the quantities μ and ν as subcases, fixing $p=0$ or $q=0$ respectively. By the first variation (see [AKM15, Lemma 3.2]), we have

$$\forall v \in \mathcal{A}(U), \ \fint_U \nabla v\cdot\mathbf{a}\nabla w(\cdot,U,p,q) = \fint_U \left(p\cdot\mathbf{a}\nabla v - q\cdot\nabla v\right). \tag{6.7}$$

Applying this to $v = w(\cdot, U, p, q)$ yields

$$J(U,p,q) = \fint_U \frac{1}{2}\nabla w(\cdot,U,p,q)\cdot \mathbf{a}\nabla w(\cdot,U,p,q) = \frac{1}{2}\fint_U \left(-p\cdot \mathbf{a}\nabla w(\cdot,U,p,q) + q\cdot \nabla w(\cdot,U,p,q)\right). \tag{6.8}$$

Since $(p,q)\mapsto J(U,p,q)$ is a quadratic form, we see that we can easily recover the knowledge of the spatial averages of fluxes and gradients of maximizers from the knowledge of J itself, by polarization (see [AKM15, (3.10)] or [AKM16a, (4.8)]).

As before, in order to improve upon (6.1)-(6.2), we can focus on studying the rate of convergence of the expectations of μ and ν. The conjunction of the statements

$$0 \le \mathbb{E}[\nu(\square_r,p)] - \frac{1}{2}p\cdot \bar{\mathbf{a}}p \le Cr^{-\alpha}$$

and

$$0 \le -\mathbb{E}[\mu(\square_r,q)] - \frac{1}{2}q\cdot \bar{\mathbf{a}}^{-1}q \le Cr^{-\alpha},$$

is equivalent to

$$0 \le \mathbb{E}[J(\square_r,p,\bar{\mathbf{a}}p)] \le Cr^{-\alpha}, \tag{6.9}$$

up to an inconsequential change of the constant C. We focus on improving the latter property. This allows to study the function $w(\cdot,\square_r,p,\bar{\mathbf{a}}p)$, which is already small a priori, contrary to $u(\cdot,\square_r,\bar{\mathbf{a}}p)$ or $v(\cdot,\square_r,p)$ taken separately, see the first equality in (6.8). We also make crucial use of the elementary observation that if we perturb the maximizer $w(\cdot,\square_r,p,q)$ slightly and compute

$$\fint_U \left(-\frac{1}{2}\nabla w\cdot \mathbf{a}\nabla w - p\cdot \mathbf{a}\nabla w + q\cdot \nabla w\right)$$

with $w(\cdot,\square_r,p,q)$ replaced by $w(\cdot,\square_r,p,q)+v$, then the obtained result changes *quadratically* in the size of v. This is a general phenomenon in the minimization of uniformly convex functions (or rather, due to our sign convention, the maximization of uniformly concave functions), which we will refer to as the *quadratic response*. In the context of linear equations, the quadratic response can be computed exactly (see [AKM15, Lemma 3.2]); it is equal to $\fint_U \frac{1}{2}\nabla v\cdot \mathbf{a}\nabla v$.

A sketch of the argument used in [AKM15] for the improvement of the exponent α is presented in [AKM15, Subsection 4.1]. In short, we study the problem on a large box $\square_R$, which we decompose into subboxes $(y+\square_r)_y$. By the first equality of (6.8), the problem is to show that the expectation of

$$\fint_{\square_R}(-p\cdot \mathbf{a}\nabla w + \bar{\mathbf{a}}p\cdot \nabla w) = \left(\frac{r}{R}\right)^d \sum_y \fint_{y+\square_r}(-p\cdot \mathbf{a}\nabla w + \bar{\mathbf{a}}p\cdot \nabla w)$$

is much smaller than $R^{-\alpha}$, where we write $w = w(\cdot,\square_R,p,\bar{\mathbf{a}}p)$ for concision. By the first variation (6.7), the latter sum is equal to

$$\sum_y \fint_{y+\square_r} \nabla w_y\cdot \mathbf{a}\nabla w, \tag{6.10}$$

where $w_y := w(\cdot,y+\square_r,p,\bar{\mathbf{a}}p)$. The assumption of (6.9) and the Lipschitz estimate (Theorem 3.2) imply that

$$\mathbb{E}\left[\fint_{y+\square_r}|\nabla w_y|^2\right] \lesssim r^{-\alpha} \quad \text{and} \quad \mathbb{E}\left[\fint_{y+\square_r}|\nabla w|^2\right] \lesssim R^{-\alpha}. \tag{6.11}$$

Using Hölder's inequality and these estimates in (6.10), we essentially recover an estimate of the order of $R^{-\alpha}$ (for $r \simeq R$). No improvement results. This is hardly surprising, since we did not take advantage of the decorrelation properties of the coefficient field.

If w is sufficiently close to an affine function of slope p_y in $y+\square_r$ (in the sense that w almost has this affine boundary condition), then (since w_y is a solution to the equation) we expect that up to a small error, we can replace the sum in (6.10) by

$$\sum_y p_y \cdot \fint_{y+\square_r} \mathbf{a}\nabla w_y. \tag{6.12}$$

Essentially, by looking at the parameter $q = \bar{\mathbf{a}}p$, we have enforced that

$$\mathbb{E}\Big[\fint_{\square_R} \nabla w\Big] \simeq 0,$$

and therefore $\sum_y \mathbb{E}[p_y] \simeq 0$. The expectation of (6.12) is therefore essentially a correlation between the vector of spatially averaged gradients (p_y) and the vector of spatially averaged fluxes $\fint_{y+\square_r} \mathbf{a}\nabla w_y$. Clearly, the latter is essentially a vector of i.i.d. random variables. The point of the higher-order regularity (Theorem 5.1) is to guarantee that on the other hand, the vector (p_y) exists and *varies slowly as we move from one box to the neighboring one*. In other words, it cannot "conspire" against our trying to leverage on CLT cancellations and line up with the independent positive and negative deviations of the vector of averaged fluxes. If we denote by s a mesoscale with $r \ll s \ll R$ on which the p_y's are approximately constant, then we can expect to bound the expectation of (6.12) by

$$R^{-\alpha/2} r^{-\alpha/2} \Big(\frac{s}{r}\Big)^{-d/2}. \tag{6.13}$$

Indeed, by (6.11), $R^{-\alpha/2}$ is the expected size of each p_y and $r^{-\alpha/2}$ is the expected size of the spatially averaged flux. The factor of $(s/r)^{-d/2}$ comes from the CLT scaling, as $(s/r)^d$ is the number of cubes of size r in each larger cube of size s.

The expression in (6.13) can be made smaller than $R^{-\alpha}$ by choosing the mesoscales r and s appropriately, provided that $\alpha < d$. This suggests that the correct estimate for $\mathbb{E}\,[J(\square_R, p, \bar{\mathbf{a}}p)]$ should be (almost) R^{-d}.

Unfortunately, the actual implementation of this argument cannot perform so well. Indeed, recall from (6.6) that w is the difference of the μ minimizer, solving a Neumann problem with affine flux condition, and of a ν minimizer, solving a Dirichlet problem with affine boundary condition. Each of these minimizers is expected to display a boundary layer of size at least $O(1)$ near the boundary (and these will not magically cancel out), which necessarily create a contribution to $\mathbb{E}[J(\square_R, p, \bar{\mathbf{a}}p)] \gtrsim R^{-1}$. Technically, Theorem 5.1 only gives *interior* regularity estimates, so we need to trim a boundary layer before implementing the sketch of argument above; this boundary layer contributes an error term of size R^{-1}.

To sum up, we have managed to improve on the exponent α in (6.1)-(6.2) and show that the inequalities hold with α arbitrarily close to 1 instead of a tiny, unknown $\alpha > 0$. The result comes with essentially optimal control of the stochastic integrability. Although suboptimal, the argument presented in [AKM15] is definite evidence that an "additive" quantitative theory of homogenization is possible. In fact, it already brings us arbitrarily close to the optimal exponent for the convergence of μ and ν in dimension $d = 2$. It would also give optimal estimates if the coefficient field was sufficiently correlated to force the exponent of convergence to be stuck below $\alpha = 1$. It also makes it clear that the optimal estimates for the stochastic integrability of the corrector will be of the form of the Gaussian estimates of Theorem 1.1, a proprety that is currently out of reach of "nonlinear" methods.

7. Revealing the additive structure

This final section outlines the approach to quantitative homogenization developed in [AKM16a]. This approach refines the arguments of the previous section and finally allow to prove Theorem 1.1. While we were writing [AKM16a], Gloria and Otto posted a preprint [117] which also explores the possibility to prove error estimates without using "nonlinear" concentration inequalities. with an approach based on the analysis of the parabolic equation. As of now, the approach of [117] does not allow to reach the optimal exponent $\frac{d}{2}$ in (1.4)-(1.5)-(1.6). Being *at* the critical exponent instead of almost there is crucial for the subsequent more refined description of the fluctuations in terms of white noise (for the energy) and gradient GFF (for the gradients), as described informally in Subsection I.1.3 and rigorously obtained in [AKM16b]. As will be clear below, our approach allows to prove estimates of fluctuations at the critical exponent $\frac{d}{2}$, because we prove first that the energy quantity J is additive *way past this exponent*, up to *exponent* d. We also mention another very recent preprint of Duerinckx, Gloria and Otto [83] where the convergences of appropriate quantities to white noise or the GFF similar to those obtained in [MO14, MN15, GM15] (see Subsection I.1.3) were reproved using different arguments still based on "nonlinear" concentration inequalities. (See the introduction to [AKM16b] for a more precise discussion.)

Our first attempt [AKM15] to prove Theorem 1.1 was hampered by boundary layers because in the definition of J (6.5), the variational problem abruptly stops at the boundary of the domain. It is therefore natural to introduce smooth masks to replace indicator functions of domains. Using masks defined in terms of a heat kernel is convenient, due to their semigroup property. Recall the notation introduced in (1.2) for integrals against a heat kernel mask. It is therefore natural to try to replace the definition of $J(z+\square_r,p,\bar{\mathbf{a}}q)$ by

$$\sup_{w\in\mathcal{A}}\int_{\Phi_{z,r}}\left(-\frac{1}{2}\nabla w\cdot\mathbf{a}\nabla w-p\cdot\mathbf{a}\nabla w+\bar{\mathbf{a}}q\cdot\nabla w\right),$$

where now the supremum is over all possible $\mathbf{a}$-harmonic functions on $\mathbb{R}^d$. Formulated as such, it is however very difficult to control the resulting maximizer and exclude wild oscillations where the heat kernel mask is small. This suggests to prescribe a growth condition on the set of candidates. We found most convenient to make this idea precise by restricting the set of maximizer candidates to the set $\mathcal{A}_1$ of subquadratically growing $\mathbf{a}$-harmonic functions. Recall that this set is a $(d+1)$-dimensional vector space spanned by constants and functions of the form $p\mapsto p\cdot x+\phi_p(x)$, $p\in\mathbb{R}^d$. For convenience, we will slightly alter our terminology and allow ourselves to use the word "correctors" to refer to elements of $\mathcal{A}_1$, instead of using this word for the functions ϕ_p. We therefore define, for every $z\in\mathbb{R}^d$, $r\geq 1$ and $p,q\in\mathbb{R}^d$,

$$J(z,r,p,q):=\sup_{w\in\mathcal{A}_1}\int_{\Phi_{z,r}}\left(-\frac{1}{2}\nabla w\cdot\mathbf{a}\nabla w-p\cdot\mathbf{a}\nabla w+\bar{\mathbf{a}}q\cdot\nabla w\right). \tag{7.1}$$

We denote the maximizer with null $\Phi_{z,r}$-average by $w(\cdot,z,r,p,q)$. (Note that we now write $\bar{\mathbf{a}}q$ instead of q in the integral. This is of course a simple change of variables, so that J is small when $p\simeq q$ instead of $p\simeq\bar{\mathbf{a}}q$. It is purely a matter of convenience, especially useful for writing (7.2) below.) The variational problem is really posed over the d-dimensional space $\mathcal{A}_1$ moded out by constants. A very convenient aspect of this formulation is that this space and the natural spaces for the variables p and q have the same dimension. In particular, for a fixed $q\in\mathbb{R}^d$ and r past a minimal radius, the mapping which to a given $p\in\mathbb{R}^d$ associates the maximizing element of $\mathcal{A}_1/\{\text{constants}\}$ in (7.1) is bijective (see the proof of [AKM16a, Lemma 8.1]).

In [AKM16a], we found it fruitful to consider higher-order versions of the quantity J defined in (7.1), where for each p, q in the space $\overline{\mathcal{A}}_k$ of $\bar{\mathbf{a}}$-harmonic polynomials of degree k, we set

$$J_k(z,r,p,q) := \sup_{w \in \mathcal{A}_k} \int_{\Phi_{z,r}} \left(-\frac{1}{2} \nabla w \cdot \mathbf{a} \nabla w - \nabla p \cdot \mathbf{a} \nabla w + \bar{\mathbf{a}} \nabla q \cdot \nabla w \right). \tag{7.2}$$

This preserves the bijective property mentioned above. The definition (7.1) corresponds to the choice of $k = 1$. In this case, it is best to modify the definition (1.1) and use the $\bar{\mathbf{a}}$-harmonic heat kernel instead. We believe that these more general quantities can be used to understand refined properties of solutions. As of now, we use these to strengthen the localization statements we are able to prove on J and J_k. However, for the purpose of proving Theorem 1.1, we can entirely work with the simpler quantity $J = J_1$ defined in (7.1), so we will not discuss the case of general k any further. The following summarizes the main properties of the quantity J.

Theorem 7.1 (Additive structure of J)**.** *For every $s < 1$, there exists a constant $C < \infty$ such that the following statements hold.*

(i) Additivity. *For every $R > r \geq 1$ and $p, q \in B_1$,*

$$\left| J(0,R,p,q) - \int_{\Phi_{\sqrt{R^2-r^2}}} J(\cdot,r,p,q) \right| \leq \mathcal{O}_s\left(Cr^{-d}\right).$$

(ii) Control of the expectation. *For every $r \geq 1$ and $p, q \in B_1$,*

$$\left| \mathbb{E}\left[J(0,r,p,q)\right] - \frac{1}{2}(q-p) \cdot \bar{\mathbf{a}}(q-p) \right| \leq Cr^{-d}.$$

(iii) CLT scaling of the fluctuations. *For every $r \geq 1$ and $p, q \in B_1$,*

$$\left| J(0,r,p,q) - \mathbb{E}\left[J(0,r,p,q)\right] \right| \leq \mathcal{O}_{2s}\left(Cr^{-\frac{d}{2}}\right).$$

(iv) Localization. *For every $\delta, \varepsilon > 0$, there exist $C < \infty$ and, for every $r \geq 1$ and $p, q \in B_1$, an $\mathcal{F}(B_{r^{1+\delta}})$-measurable random variable $J^{(\delta)}(0,r,p,q)$ such that, for every $\gamma \in \left(0, \frac{d}{2s} \wedge \left(\frac{d}{2}(1+\delta) + \delta\right) - \varepsilon\right]$,*

$$\left| J(0,r,p,q) - J^{(\delta)}(0,r,p,q) \right| \leq \mathcal{O}_{2s}\left(Cr^{-\gamma}\right).$$

Each of the estimates of Theorem 7.1 is optimal, with the exception of the localization statement. While any exponent larger than $\frac{d}{2}$ is satisfactory in the sense that it shows that the localization error is of strictly lower order compared to the CLT scaling, we expect that

$$\left| J(0,r,p,q) - J^{(\delta)}(0,r,p,q) \right| \leq \mathcal{O}_s\left(Cr^{-d}\right).$$

In dimension $d = 2$, we can in fact prove this estimate with the exponent d replaced by any $\gamma < d$, using the higher-order energy quantities J_k. Theorem 1.1 is a consequence of Theorem 7.1. Indeed, up to the identification of the corrector in the direction of p as the maximizer for $J(0,r,-p,0)$, we readily obtain (1.4). Moreover, recalling the second identity in (6.8) (or rather the similar identity for the newly defined J), we can infer from J the spatial averages of gradients and fluxes of maximizers and obtain (1.5)-(1.6). The control of the correctors themselves is deduced via a heat-kernel version of the multiscale Poincaré inequality.

The CLT scaling of the fluctuations of J (property (iii) above) is obtained as a consequence of the additivity and localization properties. One can in fact be more precise and prove an actual central limit theorem for J, see [AKM16b]. The control of the expectation (property (ii) above) is derived from the fact that

$(p,q) \mapsto \mathbb{E}[J(0,r,p,q)]$ is a quadratic form, which we understand up to $r^{-\delta}$ by (4.7)-(4.8), and additivity.

For clarity, we will temporarily stop to keep track of the value of the exponent s describing the stochastic integrability of the random variables. Denote by $\mathsf{Add}(\alpha)$ the "additivity" statement that there exists a constant $C < \infty$ such that for every $z \in \mathbb{R}^d$, $1 \le r \le R$ and $p, q \in B_1$,

$$\left| J(0,R,p,q) - \int_{\Phi_{\sqrt{R^2-r^2}}} J(\cdot,r,p,q) \right| \le \mathcal{O}\left(Cr^{-\alpha}\right).$$

Denote by $\mathsf{Fluc}(\alpha)$ the "fluctuation" statement that there exists a constant $C < \infty$ such that for every $z \in \mathbb{R}^d$, $1 \le r \le R$ and $p, q \in B_1$,

$$\left| J(0,r,p,q) - \mathbb{E}\left[J(0,r,p,q)\right] \right| \le \mathcal{O}\left(Cr^{-\alpha}\right).$$

Assuming that $J(z,r,p,q)$ is essentially localized in a ball of radius r centered around z, we can summarize the structure of the argument by the following two heuristic implications:

$$\mathsf{Add}(\alpha) \quad \Longrightarrow \quad \mathsf{Fluc}\left(\alpha \wedge \tfrac{d}{2}\right), \tag{7.3}$$

and

$$\mathsf{Fluc}(\alpha) \quad \Longrightarrow \quad \mathsf{Add}(2\alpha). \tag{7.4}$$

Since we can deduce from (6.1)-(6.2) that $\mathsf{Add}(\alpha)$ holds for some tiny $\alpha > 0$, iterating on the implications above indeed leads to Theorem 7.1. The implication (7.3) is a classical CLT scaling for sums of independent random variables; the approximation by a sum is controlled by the quality of the additivity statement, that is, by the exponent α. The idea for the implication (7.4) is as follows. First, by subadditivity, we have

$$J(0,R,p,q) \le \int_{\Phi_{\sqrt{R^2-r^2}}} J(\cdot,r,p,q). \tag{7.5}$$

If $\mathsf{Fluc}\,(\alpha)$ holds, then we are able to identify the spatial averages of the maximizers up to $r^{-\alpha}$. This essentially allows us to know which corrector is picked in the maximization problem for $J(z,r,p,q)$, up to an error of $r^{-\alpha}$. In other words, we essentially know that in each maximization problem for $J(z,r,p,q)$, the chosen maximizer is the corrector with slope $q-p$, up to an error of $r^{-\alpha}$. If we line up these slopes and choose the corrector of slope exactly $q-p$, then on the one hand, we have a maximizer candidate for the large-scale problem $J(0,R,p,q)$; on the other hand, computing

$$\int_{\Phi_{z,r}} \left(-\frac{1}{2} \nabla w \cdot \mathbf{a} \nabla w - p \cdot \mathbf{a} \nabla w + q \cdot \nabla w \right)$$

with w replaced by this corrector instead of the $J(z,r,p,q)$ maximizer produces an error of about $r^{-2\alpha}$, *by quadratic response*. Therefore

$$J(0,R,p,q) \ge \int_{\Phi_{\sqrt{R^2-r^2}}} J(\cdot,r,p,q) - \mathcal{O}(Cr^{2\alpha}).$$

Combining this with (7.5) yields the implication (7.4).

This heuristic derivation has to be taken with a grain of salt. First, if we do keep track of the exponent s of stochastic integrability, we see that squaring the error in additivity replaces s by $s/2$. More precisely, if the initial error in additivity is $\mathcal{O}_s(r^{-\alpha})$, we have argued that we can essentially upgrade this error and replace it by its square $(\mathcal{O}_s(r^{-\alpha}))^2$, which is $\mathcal{O}_{s/2}(r^{-2\alpha})$. Since we start the inductive argument with a possibly very tiny exponent $\alpha > 0$, if we simply argue like this, we will need to make a possibly very large number of iterations and end up with a possibly very small exponent s of stochastic integrability. Note that due to the boundedness of

J, we can at each step move back from the squared estimate to the original one: indeed, by (6.3),

$$\mathcal{O}_{s/2}(r^{-2\alpha}) \wedge 1 = \mathcal{O}_s(r^{-\alpha}).$$

The key to preserving the quality of the stochastic integrability is to use this principle with the a priori information that J is additive up to an error of $r^{-\delta}$ with overwhelming probability, where $\delta > 0$ is possibly very small, as we learn from (6.1)-(6.2). This essentially allows us to use instead the principle that

$$\mathcal{O}_{s/2}(r^{-2\alpha}) \wedge r^{-\delta} = r^{-\delta}\left(\mathcal{O}_{s/2}(r^{-2\alpha+\delta}) \wedge 1\right) = \mathcal{O}_s\left(r^{-\alpha-\frac{\delta}{2}}\right).$$

A more careful analysis shows that this principle allows indeed to preserve the optimal exponent of stochastic integrability along the iterations.

A second point that was ignored in the heuristic presentation of the argument concerns the proof of the localization property of J. The definition (7.1) is non-local for two reasons: first, because the heat kernel has unbounded support. Since the candidate maximizers have controlled growth, it is easy to chop the tails and get rid of this problem. More importantly, the set of maximizer candidates $\mathcal{A}_1$ (that is, the set of correctors) is a priori very non-local. The key to showing localization properties of J such as property (iv) of Theorem 7.1 is to use the intrinsic regularity of Theorem 5.2. Indeed, we first define an auxiliary maximization problem similar to that of the right side in (7.1), but which is measurable with respect to the coefficient field in a ball of radius $r^{1+\delta}$ by definition. Using the current control of the properties of J (and therefore of the correctors) up to exponent α, we can infer that this local function is close to a corrector up to exponent α, over the larger ball of radius $r^{1+\delta}$. Now, Theorem 5.2 allows us to deduce that this maximizer is *closer than exponent α to some corrector* in a ball of smaller radius, and therefore to show localization in a ball of radius $r^{1+\delta}$ up to an error of order $r^{-\alpha(1+\delta)-\delta}$. See [AKM16a, Lemma 9.1] for more precision.

A relatively obscure aspect of the heuristic argument in favor of (7.4) is the statement that identifying the spatial average of the gradient of a maximizer essentially allows us to "recognize" it as the corrector with this slope. This loose statement hides a fundamental difficulty. (Note that if we could really perform this identification exactly, we would obtain white noise scaling limit for the spatial averages of the gradient of the corrector!) We now discuss it more precisely.

Observe first that $\mathsf{Add}(\alpha)$ and $\mathsf{Fluc}(\alpha)$ imply that

$$J(z,r,p,q) = \frac{1}{2}(q-p)\cdot\bar{\mathbf{a}}(q-p) + \mathcal{O}(Cr^{-\alpha}).$$

As was seen, this allows us to read off the spatial averages of gradients and fluxes of maximizers: we deduce that for every $z \in \mathbb{R}^d$, $r \geq 1$ and $p, q \in B_1$,

$$\int_{\Phi_{z,r}} \nabla w(\cdot,z,r,p,q) = (q-p) + \mathcal{O}(Cr^{-\alpha})$$

and

$$\int_{\Phi_{z,r}} \mathbf{a}\nabla w(\cdot,z,r,p,q) = \bar{\mathbf{a}}(q-p) + \mathcal{O}(Cr^{-\alpha}).$$

Moreover, provided that r is larger than a minimal radius described by the regularity theory, we know that the mapping $p \mapsto w(\cdot,z,r,p,0)$ spans $\mathcal{A}_1/\{\text{constants}\}$ (recall the discussion below (7.1)). This shows that the correspondence between spatial averages of gradients and fluxes of correctors is described by $\bar{\mathbf{a}}$, up to a suitable error, cf. [AKM16a, Lemma 8.1]. (In Subsection I.1.3, we saw that describing the next-order correction to this mapping was the key idea to unveil the precise nature of the large-scale fluctuations of correctors.) From this, we infer that if $v \in \mathcal{A}_1$, then the mapping $x \mapsto \int_{\Phi_{x,r}} v$ is essentially $\bar{\mathbf{a}}$-harmonic, up to a suitable error; cf. [AKM16a,

Lemma 8.5] (again, the ideas of Subsection I.1.3 are similar but keep track of the precise form of the next-order correction). We can then compare maximizers on two different scales ([AKM16a, Lemma 8.8]) using arguments similar to those employed to derive (4.6), and this essentially justifies the heuristic argument in favor of the implication (7.4).

CHAPTER III

The Φ^4 model

The goal of this chapter is to review my work with Hendrik Weber [MW14, MW15, MW16], which focuses on properties of the Φ^4 model in two and three dimensions. Our main results are the convergence of a discrete Ising-type model to Φ^4 in two dimensions, and the proof of global well-posedness of the equation on the full two-dimensional plane, as well as on the three-dimensional torus. These works will be reviewed in Sections 3, 5 and 6 respectively. The other sections are introductory.

1. Heuristic link between the Ising and Φ^4 models

We start by a heuristic discussion supporting the idea that the Φ^4 model is a continuous version of the Ising model. We argue in two steps, first modifying the Ising model on a finite graph to allow the spins to take values in the continuum, and then passing to the continuous limit in the graph itself.

The first step is elementary: we simply want to define a one-body potential that has two minima, for instance using the potential

$$V : \begin{cases} \mathbb{R} & \to & \mathbb{R} \\ \phi & \mapsto & \frac{1}{4}(\phi^2 - 1)^2. \end{cases} \tag{1.1}$$

Given a finite graph $G = (V, E)$, we define the Gibbs measure

$$\frac{1}{Z} \exp\left(- \sum_{x \in V} V(\phi(x)) - \frac{1}{2} \sum_{x \sim y} (\phi(y) - \phi(x))^2 \right) \prod_{x \in V} \mathrm{d}\phi(x), \tag{1.2}$$

where Z is such that the expression above is a probability measure on $\mathbb{R}^V$. Naturally, we could add tunable multiplicative parameters to the model. The potential V favors configurations where spins concentrate around the values ± 1, while the two-body interactions tend to favor the alignment of spins. If we take the graph G to be a finite box of $\mathbb{Z}^d$, and assume that the sums in (1.2) are suitably normalized to remain of order 1, then formally passing to the limit leads to a measure proportional to

$$\exp\left(- \int \left[\frac{1}{4}(\phi^2 - 1)^2 + \frac{1}{2} |\nabla \phi|^2 \right] \right) \mathcal{D}\phi, \tag{1.3}$$

where $\mathcal{D}\phi$ is an infinite-dimensional product of Lebesgue measures. Of course, this object is not well-defined mathematically, but we could modify the expression (1.3) into

$$\exp\left(-\frac{1}{4} \int (\phi^2 - 1)^2 \right) \mathrm{d}\mu(\phi),$$

where μ is the law of the Gaussian free field. In dimensions two and higher, the Gaussian free field only makes sense as a distribution, so this expression is still meaningless. Nevertheless, in dimension $d = 2$, Nelson [168] showed that this construction can be carried if one performs a Wick renormalization of the polynomial $(\phi^2 - 1)^2$. This measure is often called the two-dimensional Φ^4 Euclidean field theory. It is also possible to build the Φ^4 measure rigorously in three dimensions

[107, 87, 108, 92, 91], but the procedure is much more complicated and the resulting measure is singular with respect to the law of the Gaussian free field.

Returning to the setting of the finite graph, we can also consider dynamics for which the Gibbs measure in (1.2) is reversible. Indeed, it suffices to perturb the L^2 gradient flow associated with the energy by the corresponding noise. Namely, we consider the dynamics given by

$$\mathrm{d}\phi(t,x) = \left[-V'(\phi(t,x)) + \sum_{y\sim x}(\phi(t,y)-\phi(t,x))\right]\mathrm{d}t + \frac{1}{\sqrt{2}}\,\mathrm{d}B(t,x),$$

where $(B(\cdot,x))_{x\in V}$ are independent standard Brownian motions. Formally passing to the continuous-space limit as above leads to

$$\partial_t X = \Delta X - X^3 + aX + \xi, \tag{1.4}$$

where ξ is a space-time white noise and a is a constant. (Every equation of this form with constants in front of each term can be reduced to this equation for a suitable value of a, by changes of scales. We changed notation from Φ to X to be more consistent with the rest of the document.) In this derivation, the choice of double-well potential (1.1) was relatively arbitrary. However, we should in fact think of the Φ^4 model as a scaling limit in the sense described in Section I.2 for the KPZ equation (see the discussion around (I.2.4)): in this case, the higher powers of an arbitrary potential become negligible, and only the first non-vanishing contribution remains. For a generic symmetric double-well potential, the first non-quadratic contribution is a fourth power. (Quadratic contributions preserve the Gaussian character of the measure, they simply introduce a "massive" term.)

2. Perspective from quantum field theory

The previous section and Section I.2 motivate the Φ^4 model and the KPZ equation as continuous versions of models of statistical mechanics. In the 60's and 70's, the Φ^4 model was the focus of an intense research effort for a different reason, namely as a fundamental test-bed for constructive quantum field theory. The goal of this section is to briefly review this perspective.

Let us first recall the original motivation for quantum field theory itself. Classical quantum mechanics was hugely successful to explain a variety of phenomena such as the photoelectric effect or the black-body radiation, and to give accurate predictions for instance on the spectrum of light waves emitted by excited atoms. However, very fine experiments of atomic spectroscopy reveal small deviations between classical quantum theory predictions and observations.

Besides, classical quantum mechanics is not compatible with special relativity. It is desirable from a theoretical perspective to devise a theory that combines quantum and relativistic effects in a unified framework. This is (what I understand to be) the overarching goal of quantum field theory. On a more experimental side, physicists now understand well how to build a quantum field theory of atomic spectroscopy, and this more refined theory agrees perfectly with experiments, up to measurement errors [109, Chapter 15]. The need for quantum field theory as a predictive tool becomes much more central when one moves to nuclear and elementary particles physics, where short distances and high energies give rise to large relativistic effects. Physicists have devised a very successful quantum field model, called the *Standard model*, accounting for the electromagnetic, weak and strong interactions, and giving a classification of subatomic particles. This model is complicated and very far from being mathematically understood.

The Φ^4 model was introduced as a simplified model of quantum field theory. At this stage, we have not explained what a quantum field theory is as a mathematical

object. For the Φ^4 model, physicists would write the classical Lagrangian

$$-\frac{1}{2}|\partial_t X|^2 + \frac{1}{2}|\nabla X|^2 + \frac{1}{2}\mu^2 X^2 + \frac{1}{4}\lambda X^4, \tag{2.1}$$

or the associated action, which is the Lagrangian integrated over space-time, and proceed to "quantize" this action, e.g. by means of a path integral formulation. It is not easy to make sense of this procedure mathematically. In the 50's, Wightman proposed an axiomatic formulation of what a quantum field theory should be. It was realized very early on that moving to imaginary time is technically very convenient (this replaces the minus sign in front of the time derivative in (2.1) by a plus sign). In particular, it maps the action described above to that of the *Euclidean* Φ^4 model as introduced in (1.3). Osterwalder and Schrader [172, 171] showed that a Euclidean field theory satisfying a certain number of properties, now called Osterwalder-Schrader axioms, can be mapped back into the corresponding quantum field theory, and vice versa. This was the starting point of a large number of studies devoted to the rigorous construction of Euclidean field theories satisfying the Osterwalder-Schrader axioms. Note that in this point of view, imaginary time is one of the coordinates we otherwise think of as a space coordinate. The time variable of (1.4) is an *additional* time variable, sometimes called the stochastic time to distinguish it from the physical (albeit imaginary!) time. Perhaps the most important of the Osterwalder-Schrader axioms are invariance under Euclidean isometries and reflection positivity (see [40] for a review on reflection positivity).

As far as I understand, the construction of a three-dimensional Euclidean Φ^4 theory satisfying both Euclidean invariance and reflection positivity has only been achieved under the additional assumption of small coupling constant. A Euclidean field satisfying reflection positivity alone has been constructed for all couplings. References and more precise explanations can be found in [109, Section 20.1].

The idea of viewing Euclidean field theories as invariant measures of stochastic dynamics such as (1.4) was brought forward by Parisi and Wu [174]. If one could show the existence and uniqueness of an invariant measure for (1.4), then this would provide us with a Euclidean field theory satisfying Euclidean invariance. As was pointed out recently [136], the law of the solution to (1.4) at a given finite time is *not* reflection positive, and it seems difficult to recover reflection positivity by passing to the limit.

One way to recover reflection positivity would consist in showing that the invariant measure of the continuous equation is the limit of invariant measures of Ising-type models. It is standard to check that the invariant measure of an Ising-type model is reflection positive (see [40, Corollary 5.4]), and this property is preserved by passing to the limit. This is close to the results of [MW14], [49, 129].

3. Convergence of the two-dimensional Ising-Kac model to Φ^4

This section reviews the work [MW14], which shows that a certain two-dimensional Ising-type model rescales to Φ^4.

The local solution theory for the Φ^4 model on the two-dimensional torus was briefly sketched in Section I.2. The precise understanding of this solution theory is based on Besov spaces and their multiplicative structure. At the risk of logical inconsistency, we postpone a precise description of these spaces and their properties to Section 4, which is more oriented towards the solution theory in three space dimensions. For now, we content ourselves with the fact that there exists a family of distribution spaces $\mathcal{C}^\alpha$ on the torus, where $\alpha \in \mathbb{R}$ captures the degree of regularity of the distribution. For $\alpha \in (0,1)$, the space $\mathcal{C}^\alpha$ is the space of Hölder-continuous functions of exponent α. For $\alpha < 0$, we have

$$f \in \mathcal{C}^\alpha \quad \Longrightarrow \quad \forall x \in [-1,1]^2,\ \big(f, \varepsilon^{-2}\phi(\varepsilon^{-1}(\cdot - x))\big) \lesssim \varepsilon^\alpha,$$

where ϕ is a smooth test function and $(\cdot,\cdot)$ is the duality pairing. The local solution theory also borrows from Section 4 that the multiplication operator $(f,g) \mapsto fg$ extends to a continuous bilinear map from $\mathcal{C}^\alpha \times \mathcal{C}^\beta$ to $\mathcal{C}^\alpha$ if $\alpha < 0 < \beta$ satisfy $\alpha + \beta > 0$, as well as regularizing properties of the heat semigroup.

We now introduce the ferromagnetic spin model that will ultimately be shown to converge to the Φ^4 model.

Recall that by the very definition of the notion of subcriticality, a solution to the Φ^4 equation is not invariant under scaling. Therefore, the Φ^4 equation cannot be obtained as the scaling limit of a fixed system. Rather, we need a family of models with a tunable parameter that we will adjust as we take the scaling limit (just as we need to tune the asymmetry down to zero for the asymmetric exclusion process to converge to the KPZ equation). In particular, the nearest-neighbor critical Ising model does *not* rescale to the Φ^4 model.

Our additional tunable parameter is the range of the interactions between spins. We work in the discrete torus $\Lambda_N := \{-N,\dots,N\}^2$. The space of configurations is $\Sigma_N := \{\pm 1\}^{\Lambda_N}$, and we write $\sigma = (\sigma(k))_{k\in\Lambda_N}$ for a generic element of Σ_N.

We link every quantity sent to zero or infinity in the scaling limit with a single small parameter $\gamma > 0$ such that $N = \gamma^{-2}$. The interactions between spins are described by a kernel κ_γ of unit ℓ^1 norm and with range γ^{-1}:

$$\kappa_\gamma(k) := \gamma^2 K(\gamma k),$$

where K is a fixed nonnegative, smooth, compactly supported, rotationally invariant bump function such that $\int_{\mathbb{R}^2} K = 1$. Note that we have already enforced a particular relationship between the length scale γ^{-2} of the space and the range of the interaction γ^{-1}. One can check that this choice is the only one that gives rise to the Φ^4 model in the limit. The Hamiltonian of our ferromagnetic model is given by

$$H_\gamma(\sigma) := -\frac{1}{2}\sum_{j,k} \kappa_\gamma(j-k)\sigma(j)\sigma(k) \tag{3.1}$$

$$= -\frac{1}{2}\sum_{k} \sigma(k) h_\gamma(\sigma,k), \tag{3.2}$$

where we have set

$$h_\gamma(\sigma,k) := \kappa_\gamma \star \sigma(k),$$

and $\star$ denotes the discrete convolution. The associated Gibbs measure at inverse temperature $\beta \in (0,\infty)$ is given by

$$\lambda_\gamma(\sigma) := \frac{1}{Z_\gamma}\exp\left(-\beta H_\gamma(\sigma)\right),$$

where Z_γ is the partition function. We consider dynamics that change one spin at a time. Denoting by $\sigma^j \in \Sigma_N$ the configuration obtained from σ by flipping the spin at position $j \in \Lambda_N$, we can write the infinitesimal generator of the dynamics in the general form

$$\mathcal{L}_\gamma f(\sigma) := \sum_j c_\gamma(\sigma,j)\left(f(\sigma^j) - f(\sigma)\right), \tag{3.3}$$

where $c_\gamma(\sigma,j)$ denotes the jump rate from σ to σ^j. The Glauber dynamics corresponds to the choice

$$c_\gamma(\sigma,j) := \frac{\lambda_\gamma(\sigma^j)}{\lambda_\gamma(\sigma) + \lambda_\gamma(\sigma^j)}.$$

This rate clearly satisfies the detailed balance condition

$$\lambda_\gamma(\sigma)c_\gamma(\sigma,j) = \lambda_\gamma(\sigma^j)c_\gamma(\sigma^j,j).$$

The Glauber dynamics is a popular scheme for numerical purposes since the jump rate depends only on the difference $H_\gamma(\sigma^j) - H_\gamma(\sigma)$. Indeed, an explicit computation reveals that

$$c_\gamma(\sigma, j) = \frac{1}{2}\left[1 - \sigma(j)\tanh\left(\beta h_\gamma(\sigma, j)\right)\right]. \tag{3.4}$$

We denote the Markov process associated with these jump rates by $(\sigma(t))_{t\geq 0}$, and by a slight abuse of notation, we write $h_\gamma(t,k) := h_\gamma(\sigma(t),k)$.

Our goal is to show that after a suitable scaling, the locally averaged magnetic field $h_\gamma(t,k)$ converges in law to the Φ^4 model. We are interested in situations where the initial condition is relatively close to equilibrium, in the sense that $h_\gamma(t=0,k)$ is small. (If we start this Ising-type model far from equilibrium, say with all spins equal to 1 in half of the torus and all spins equal to -1 in the other half, then we will monitor interface dynamics, which in the presence of a small external magnetic field can be expected to be in the KPZ universality class, and in any case have a very different behavior than what we want to describe here.) In particular, we expect the jump rate c_γ to stay relatively close to $\frac{1}{2}$.

In the spirit of numerical analysis, we want to justify that our spin system is a "discretization scheme" of the continuous equation. The starting point is to write the integrated PDE-like identity

$$h_\gamma(t,k) = h_\gamma(0,k) + \int_0^t \mathcal{L}_\gamma h_\gamma(s,k)\,\mathrm{d}s + m_\gamma(t,k), \tag{3.5}$$

where $m_\gamma(\cdot,k)$ is a martingale. The predictable quadratic variation of this martingale can be computed explicitly as the integral of the "carré du champ" along the trajectory of the process, see [MW14, (2.11)]. As for the integral in (3.5), observing that

$$h_\gamma(\sigma^j,k) - h_\gamma(\sigma,k) = -2\kappa_\gamma(k-j)\sigma(j)$$

and using (3.3)-(3.4), we get

$$\mathcal{L}_\gamma h_\gamma(\sigma,k) = -h_\gamma(\sigma,k) + \kappa_\gamma \star \tanh\left(\beta h_\gamma(\sigma,\cdot)\right)(k).$$

Since we are interested in the regime of small h_γ, we can write

$$\tanh\left(\beta h_\gamma\right) \simeq \beta h_\gamma - \frac{(\beta h_\gamma)^3}{3} + \cdots,$$

and a rearrangement yields

$$\mathcal{L}_\gamma h_\gamma(\sigma,k) = (\kappa_\gamma \star h_\gamma - h_\gamma) + (\beta-1)\kappa_\gamma \star h_\gamma - \frac{\beta^3}{3}\kappa_\gamma \star h_\gamma^3 + \mathsf{Err}, \tag{3.6}$$

where Err is a hopefully small error term. The combination of (3.5) and (3.6) looks very promising: the first term on the right of (3.6) is an approximation of a multiple of the Laplacian, the second term is linear, the third term displays the expected cubic behavior signaling that the effective potential is bimodal, and the martingale in (3.5) is akin to an integrated white noise. Our goal is to find the richest possible scaling limit for the field h_γ, that is, one where the diffusive part, the non-linearity and the randomness all contribute to the limit. In other words, we want to find a rescaling in time, space and magnitude for h such that after scaling, the first term on the right side of (3.6) converges to a Laplacian, the non-linearity remains of order 1, and so does the rescaled quadratic variation of m_γ. There is only one way to satisfy all these constraints: it is to set

$$X_\gamma(t,x) := \gamma^{-1} h_\gamma\left(\gamma^{-2}t, \gamma^{-2}x\right).$$

(This is explained in more details in [MW14]. See also [MW14, Remark 2.2] for the adequate change of scale in arbitrary dimension, and its similarity with (I.2.8). The rescaling superficially looks non-diffusive, but since the model is long range, the space

is really measured in units of γ instead of γ^2, and one should read $\gamma^{-2}x = \gamma^{-1}(\gamma^{-1}x)$.) This change of scale combined with (3.5)-(3.6) leads to

$$\partial_t X_\gamma = \Delta_\gamma X_\gamma + \frac{\beta - 1}{\gamma^2} X_\gamma - \frac{\beta^3}{3} X_\gamma^3 + \partial_t M_\gamma + \mathsf{Err}', \tag{3.7}$$

where we wrote the relation in differential form for clarity, and where

$$M_\gamma(t,x) := \gamma^{-1} m_\gamma\left(\gamma^{-2}t, \gamma^{-2}x\right)$$

has quadratic variation of order 1. Equation (3.7) makes it intuitively clear that if β is far from the critical inverse temperature of the mean-field (Curie-Weiss) model $\beta = 1$, then the limit of X_γ as $\gamma \to 0$ will degenerate to 0 (β small, high temperature) or will diverge to infinity (β large, small temperature). A naive guess would be to tune the temperature according to $\beta = 1 + a\gamma^2$, where $a \in \mathbb{R}$ is a free parameter. However, our understanding of the behavior of the continuous equation suggests that X_γ^3 needs to be renormalized in order to remain of order 1. Our model performs a sort of regularization at scale γ, and therefore we expect that we need to subtract a counter-term of the form $(c_0 \log \gamma^{-1})X_\gamma$ to the equation, for a suitable choice of constant c_0. In other words, we fix the inverse temperature according to

$$\beta_\gamma := 1 + c_0\gamma^2 \log \gamma^{-1} + a\gamma^2, \tag{3.8}$$

where c_0 is a fixed constant, and $a \in \mathbb{R}$ is a free parameter. The main result of [MW14] can be informally stated as follows.

Theorem 3.1. *Let the inverse temperature be fixed according to (3.8). If $X_\gamma(t = 0) \to X^\circ$ in $\mathcal{C}^{-\alpha}$ for a sufficiently small $\alpha > 0$, then X_γ converges in law to X solution to*

$$\begin{cases} \partial_t X = \Delta X - \frac{1}{3}(X^3 - 3\infty X) + aX + \sqrt{2}\xi, \\ X(t = 0) = X^\circ, \end{cases} \tag{3.9}$$

for the topology of the space of caldag processes taking values in $\mathcal{C}^{-\alpha}$.

In (3.9), we used the informal notation with an ∞ sign to denote the renormalized equation, compare with (I.2.8). The distribution ξ is the standard space-time white noise. If desired, one can recover from this result the convergence of the field itself $\gamma^{-1}\sigma(\gamma^{-2}t, \gamma^{-2}k)$, up to a suitable weakening of the topology considered (so as to compensate for the missing convolution by κ_γ). The fact that we are interested in configurations close to equilibrium is encoded in the assumption that $X_\gamma(t = 0)$ converges: this corresponds to the assumption that $\gamma^{-1}\sigma(t = 0, \gamma^{-2}k)$ converges to a finite limit, in a sufficiently weak topology.

The fact that the critical temperature of the Ising-Kac model is shifted by a multiple of $\gamma^2 \log \gamma^{-1}$ was already shown in [56]. Theorem 3.1 was stated as a conjecture in [104].

We now describe the strategy of proof of Theorem 3.1. Recall that the solution X to the continuous equation is defined as $Z + Y(Z, Z^{:2:}, Z^{:3:})$, where we now emphasize in the notation that the function Y defined as the solution to (I.2.11) is a continuous function of $(Z, Z^{:2:}, Z^{:3:})$ for the topology of $(L^\infty([0,T], \mathcal{C}^{-\alpha}))^3$. The proof of convergence can be decomposed into three steps.

(1) We define Z_γ to be the solution to the discrete linearized equation

$$\partial_t Z_\gamma = \Delta_\gamma Z_\gamma + \partial_t M_\gamma, \tag{3.10}$$

compare with (3.7) and (I.2.9). We also define a notion of renormalized powers $Z_\gamma^{:2:}$ and $Z_\gamma^{:3:}$. We show that these quantities remain bounded in $L^\infty([0,T], \mathcal{C}^{-\alpha})$, as well as some time regularity.

(2) We show that the triple $(Z_\gamma, Z_\gamma^{:2:}, Z_\gamma^{:3:})$ converges in law to $(Z, Z^{:2:}, Z^{:3:})$.

(3) Starting from (3.7) and arguing similarly to the continuous equation, we write X_γ in the form $Z_\gamma + Y_\gamma(Z_\gamma, Z_\gamma^{:2:}, Z_\gamma^{:3:})$, up to a small error. Here the quantity Y_γ solves an equation similar to (I.2.11). We argue that the mappings

$$(Z, Z^{:2:}, Z^{:3:}) \mapsto Y(Z, Z^{:2:}, Z^{:3:}) \quad \text{and} \quad (Z, Z^{:2:}, Z^{:3:}) \mapsto Y_\gamma(Z, Z^{:2:}, Z^{:3:}) \tag{3.11}$$

are close to one another, and use their continuity properties and the previous step to conclude.

We see in this proof outline the power of having a pathwise notion of solution to the continuous equation: we can decouple the probabilistic questions concerning the convergence of the renormalized powers and the numerical-analytic question of showing that the mappings (3.11) are close to one another and continuous. This is a feature originating in Lyons's theory of rough paths, and one of the great successes of the theories built on regularity structures or on paraproducts is to preserve this feature for much more general singular stochastic PDEs than the one we consider here.

Part (3) of the outline above can be technical at places, but is perhaps not surprising. We will not discuss it further, and focus on the other parts, beginning with part (1). In order to explain the difficulty, let us briefly review how one typically argues about the existence of the renormalized square $Z^{:2:}$ of the continuous, linear equation (I.2.9). Denote by Z_ε the solution to (I.2.9) with the white noise ξ regularized on scale $\varepsilon > 0$. For convenience, we will drop the time variable in the notation, which is kept fixed in this argument. We may first want to test Z_ε^2 against a smooth function ϕ, and compute the second moment of the resulting quantity. We find

$$\mathbb{E}\left[\left(\int Z_\varepsilon^2(x)\phi(x)\,\mathrm{d}x\right)^2\right] = \int \mathbb{E}[Z_\varepsilon^2(x)Z_\varepsilon^2(y)]\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y.$$

Since $(Z_\varepsilon(x), Z_\varepsilon(y))$ is a Gaussian vector, the expectation above can be expressed in terms of its covariance matrix. We can compute this using Wick's product formula for Gaussian random variables, which expresses the expectation of a product of centered Gaussians as a sum over {all possible pairings of the variables} of products of the covariances between each pair. In our case, this gives

$$\mathbb{E}[Z_\varepsilon^2(x)Z_\varepsilon^2(y)] = 2\mathbb{E}[Z_\varepsilon(x)Z_\varepsilon(y)]^2 + \mathbb{E}[Z_\varepsilon^2(x)]\,\mathbb{E}[Z_\varepsilon^2(y)].$$

The quantity Z_ε locally resembles a Gaussian free field, and therefore

$$\lim_{\varepsilon\to 0} \mathbb{E}[Z_\varepsilon(x)Z_\varepsilon(y)] \sim \log|x-y|^{-1} \qquad (|x-y| \to 0). \tag{3.12}$$

On the other hand, $\mathbb{E}[Z_\varepsilon^2(x)] = \mathbb{E}[Z_\varepsilon^2(y)] = C_\varepsilon \simeq \log\varepsilon^{-1}$ diverges as $\varepsilon \to 0$. The renormalization is meant to compensate this divergence, and indeed, one can check that

$$\mathbb{E}\left[\left(\int (Z_\varepsilon^2(x) - C_\varepsilon)\phi(x)\,\mathrm{d}x\right)^2\right] = \int \mathbb{E}[Z_\varepsilon(x)Z_\varepsilon(y)]^2\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y.$$

In view of (3.12), this suggests that indeed $Z_\varepsilon^2(x) - C_\varepsilon$ converges to a non-trivial distribution. In order to justify this in a sufficiently strong sense, it is necessary to control the boundedness of higher moments of $Z_\varepsilon^2(x) - C_\varepsilon$. This is usually obtained by invoking Nelson's estimate, which guarantees that for each p, the p-th moment of any random variable in a fixed Wiener chaos is controlled by a constant times its second moment. The argument for the third and higher renormalized powers is identical, and displays a nice interplay between Wick's product formula for moments of products of Gaussians and Hermite's polynomials.

While these algebraic properties of Gaussian random variables are remarkable, they will no longer hold in the discrete setting, where the driving noise is not Gaussian. Similarly, classical proofs of Nelson's estimate are very rigid to the Gaussian assumption (see e.g. [169, Section 1.4.3]). The strategy to overcome these difficulties is to reinterpret the renormalized powers as iterated stochastic integrals, and to rely on the martingale structure inherited from the Markovianity of the discrete process. For simplicity, I will start to outline the idea at the level of the continuous equation. Denoting by ξ_ε the white noise regularized in space at scale ε, and by $p(t,x)$ the heat semi-group, we have

$$Z_\varepsilon(t,x) = \int_{s=0}^{t} \int_{y\in[-1,1]^2} p(t-s,x-y)\,\xi_\varepsilon(s,y)\,\mathrm{d}y\,\mathrm{d}s.$$

We introduce an auxiliary variable and denote

$$R_{\varepsilon,t}(s,x) := \int_{r=0}^{s} \int_{y} p(t-r,x-y)\,\xi_\varepsilon(r,y)\,\mathrm{d}y\,\mathrm{d}r,$$

so that $(R_{\varepsilon,t}(\cdot,x))_{0\le s\le t}$ is a martingale, and $R_{\varepsilon,t}(t,x) = Z_\varepsilon(t,x)$. Defining

$$R^{:2:}_{\varepsilon,t}(s,x) := 2\int_{r=0}^{s} R_{\varepsilon,t}(r,x)\,\mathrm{d}R_{\varepsilon,t}(r,x),$$

one can check by Itô's formula that

$$R^{:2:}_{\varepsilon,t}(s,x) = \left(R_{\varepsilon,t}(s,x)\right)^2 - \left\langle R_{\varepsilon,t}(\cdot,x)\right\rangle_s,$$

where $\langle\cdot\rangle$ denotes the predictable quadratic variation. More generally, we define recursively on n

$$R^{:n:}_{\varepsilon,t}(s,x) := n\int_{r=0}^{s} R^{:n-1:}_{\varepsilon,t}(r,x)\,\mathrm{d}R_{\varepsilon,t}(r,x),$$

and we can check by Itô's formula that

$$R^{:n:}_{\varepsilon,t}(s,x) = H_n\left(R_{\varepsilon,t}(s,x), \left\langle R_{\varepsilon,t}(\cdot,x)\right\rangle_s\right),$$

where H_n is the n-th homogeneous Hermite polynomial, see [MW14, (3.5)]. In the setting of the continuous equation, the quadratic varation $\langle R_{\varepsilon,t}(\cdot,x)\rangle$ is deterministic, but this is no longer the case when we turn to the discrete system. We set

$$Z^{:n:}_\varepsilon(t,x) := R^{:n:}_{\varepsilon,t}(t,x).$$

The formulation in terms of iterated stochastic integrals enables to estimate this quantity very efficiently. Indeed, stochastic cancellations are clearly displayed by the formulation. Moreover, The Burkholder-Davis-Gundy (BDG) inequality provides us with a suitable replacement of Nelson's estimate. Indeed, we can bound the measure $\mathrm{d}\left\langle R_{\varepsilon,t}(\cdot,x)\right\rangle_s$ by a deterministic measure (using that $C_\gamma \le 1$ in [MW14, (2.15)]), and then estimate every finite moment of $R^{:n:}_{\varepsilon,t}$ inductively on n by the BDG inequality.

We now briefly discuss part (2) of the outline above, concerning the convergence in law of $(Z_\gamma, Z^{:2:}_\gamma, Z^{:3:}_\gamma)$ to $(Z, Z^{:2:}, Z^{:3:})$. Recall that Z is defined as solving (3.10). One difficulty is that the quadratic variation of the martingale M_γ is not known exactly. It depends on the "true" field X_γ, which we know nothing about at this stage. The dependence on the field X_γ is only through the jump rate C_γ, see again [MW14, (2.15)]. In order to solve this problem, we observe that the quadratic variation is known with sufficient precision as long as the field X_γ does not become very large. That is, we choose a very large constant $\mathfrak{m}$ and define the stopping time

$$\tau_{\gamma,\mathfrak{m}} := \inf\left\{t\ge 0 : \|X_\gamma(t,\cdot)\|_{\mathcal{C}^{-\alpha}} \ge \mathfrak{m}\right\}.$$

We then consider a modified process $Z_{\gamma,\mathfrak{m}}$ which coincides with X_γ before time $\tau_{\gamma,\mathfrak{m}}$, and otherwise behaves "just right" (the precise definition is not important). We show that $Z_{\gamma,\mathfrak{m}}$ converges in law to the solution to the limiting equation, using the martingale characterization of the latter. We use the estimates proved in the first

step to ensure that the renormalized powers $Z^{:2:}_{\gamma,\mathfrak{m}}$ and $Z^{:3:}_{\gamma,\mathfrak{m}}$ also converge in law to $Z^{:2:}$ and $Z^{:3:}$ (the convergence holds jointly). We then continue the argument through step (3) of the outline using this modified process, and conclude that a similarly modified process $X_{\gamma,\mathfrak{m}}$ converges in law to the solution X to the nonlinear SPDE. In order to conclude the proof of Theorem 3.1, we need to remove the artificial stopping-time constraint we have introduced, that is, we need to assert that for every $M < \infty$,

$$\lim_{\mathfrak{m}\to\infty} \liminf_{\gamma\to 0} \mathbb{P}\left[\tau_{\gamma,\mathfrak{m}} \geq M\right] = 1.$$

This can be deduced from the convergence in law stated above, provided that for every $M < \infty$,

$$\lim_{\mathfrak{m}\to\infty} \mathbb{P}\left[\tau_{\mathfrak{m}} \geq M\right] = 1, \tag{3.13}$$

where

$$\tau_{\mathfrak{m}} := \inf\left\{t \geq 0 : \|X(t,\cdot)\|_{\mathcal{C}^{-\alpha}} \geq \mathfrak{m}\right\}$$

is the corresponding stopping time for the continuous equation. The estimate (3.13) is a formulation of the fact that the solution to the continuous equation does not blow up in finite time; we will return to this in Section 5.

4. Besov spaces and paraproducts

In this section, we review the definition and multiplicative properties of Besov spaces. Our presentation will be mostly informal; we refer to [18, Chapter 2] for a rigorous treatment.

We write $\mathscr{F}f$ or $\hat{f}$ for the Fourier transform (and by $\mathscr{F}^{-1}$ its inverse), which is well-defined for any Schwartz distribution f on $\mathbb{R}^d$, and reads, for $f \in L^1(\mathbb{R}^d)$,

$$\mathscr{F}f(\zeta) = \hat{f}(\zeta) = \int e^{-ix\cdot\zeta} f(x)\,\mathrm{d}x.$$

The definition of Besov spaces rests on a decomposition of the Fourier transform of a function along dyadic annuli: we think of splitting $\hat{f}$ into

$$\hat{f}\mathbf{1}_{B(0,1)} + \sum_{k=0}^{+\infty} \hat{f}\mathbf{1}_{B(0,2^{k+1})\setminus B(0,2^k)}.$$

The term of the series associated with a high k measures the fast oscillations of the function; the general Besov norm can be thought of as a weighted average of the L^p norm of these summands. However, it is much better for analytical purposes to use smoothened versions of these indicator functions. That is, we construct $\widetilde{\chi} \in C_c^\infty$ supported in a ball and $\chi \in C_c^\infty$ supported in an annulus, both taking values in $[0,1]$, and such that

$$\forall \zeta \in \mathbb{R}^d,\ \widetilde{\chi}(\zeta) + \sum_{k=0}^{+\infty} \chi(\zeta/2^k) = 1. \tag{4.1}$$

We write

$$\chi_{-1} = \widetilde{\chi}, \qquad \chi_k = \chi(\cdot/2^k) \quad (k \geq 0). \tag{4.2}$$

We can make sure that the supports of χ_k and $\chi_{k'}$ overlap only if $|k-k'| \leq 1$. For every $f \in C_c^\infty$ and every $k \geq -1$, we let

$$\delta_k f = \mathscr{F}^{-1}\left(\chi_k \hat{f}\right),$$

so that $f = \sum_{k\geq -1} \delta_k f$, and define the Besov norm with regularity exponent $\alpha \in \mathbb{R}$ and integrability exponents (∞,∞) as

$$\|f\|_{\mathcal{C}^\alpha} := \sup_{k\geq -1} 2^{\alpha k} \|\delta_k f\|_{L^\infty}. \tag{4.3}$$

It is easy to check that this quantity is finite for $f \in C_c^\infty$. The space $\mathcal{C}^\alpha$ is the completion of C_c^∞ with respect to this norm. This space can be realized as a subspace of the space of Schwartz distributions.

Letting

$$\eta_k = \mathscr{F}^{-1}(\chi_k), \qquad \eta = \eta_0, \tag{4.4}$$

so that for $k \geq 0$, $\eta_k = 2^{kd}\eta(2^k \cdot)$, we have for every k,

$$\delta_k f = \eta_k \star f, \tag{4.5}$$

where $\star$ denotes the convolution. In particular, for $f \in \mathcal{C}^\alpha$, we have

$$|\langle f, \eta_k \rangle| \leq \|f\|_{\mathcal{C}^\alpha} 2^{-\alpha k}.$$

More generally, for $\alpha < 0$, one can show that and for any $\phi \in C_c^\infty$,

$$\left|\left\langle f, 2^{kd}\phi(2^k \cdot)\right\rangle\right| \leq C\|f\|_{\mathcal{C}^\alpha} 2^{-\alpha k},$$

where the constant $C(\phi)$ does not depend on k. For $\alpha \in (0,1)$, one needs to ask in addition that ϕ be of zero mean. For $\alpha < 0$, the space $\mathcal{C}^\alpha$ therefore provides with a natural extension of Hölder spaces to negative exponents of regularity.

Perhaps the most important property of Besov spaces for our purpose is their multiplicative structure. We have already mentioned the following statement in the beginning of Section 3.

Proposition 4.1. *Let $\alpha < 0 < \beta$ be such that $\alpha + \beta > 0$. The multiplication $(f, g) \mapsto fg$ extends to a continuous bilinear map from $\mathcal{C}^\alpha \times \mathcal{C}^\beta$ to $\mathcal{C}^\alpha$.*

The proof of this fact rests on the decomposition

$$fg = \sum_{k<l-1} \delta_k f\, \delta_l g + \sum_{|k-l|\leq 1} \delta_k f\, \delta_l g + \sum_{k>l+1} \delta_k f\, \delta_l g,$$

which we will write suggestively in the form

$$fg = f \varolessthan g + f \varominus g + f \varogreaterthan g.$$

This is often called *Bony's decomposition* into the *paraproducts* $f \varolessthan g$ and $f \varogreaterthan g = g \varolessthan f$ and the *resonant term* $f \varominus g$.

In order to prove Proposition 4.1, it suffices to show that each of these terms extends to a continuous bilinear map from $\mathcal{C}^\alpha \times \mathcal{C}^\beta$ to $\mathcal{C}^\alpha$. However, it will be crucial for the development of future arguments to be precise about the behavior of each term separately. We thus simply assume that $\alpha < 0 < \beta$ and $f \in \mathcal{C}^\alpha$, $g \in \mathcal{C}^\beta$ to begin with, and see whether and how we can estimate each of the terms in Bony's decomposition. We start with

$$f \varolessthan g = \sum_{k<l-1} \delta_k f\, \delta_l g. \tag{4.6}$$

Recall that the spectrum of $\delta_k f$ (resp. $\delta_l g$) is supported on an annulus of radius about 2^k (resp. 2^l). The spectrum of the product is supported on the convolution of these two annuli. Since $k < l-1$, this convolution is an annulus of radius about 2^l. Moreover,

$$\|\delta_k f\|_{L^\infty} \leq 2^{-\alpha k}\|f\|_{\mathcal{C}^\alpha},$$

and similarly for $\|\delta_l g\|_{L^\infty}$. Therefore, the part of the double sum on the right side of (4.6) with spectrum in an annulus of about 2^l has L^∞ norm bounded by

$$\sum_{k=-1}^{l-2} 2^{-\alpha k - \beta l}\|f\|_{\mathcal{C}^\alpha}\|g\|_{\mathcal{C}^\beta} \leq C 2^{-(\alpha+\beta)l}\|f\|_{\mathcal{C}^\alpha}\|g\|_{\mathcal{C}^\beta},$$

where we used that $\alpha < 0$ in the last step, and where C does not depend on f or g. This suggests that

$$\|f \varolessthan g\|_{\mathcal{C}^{\alpha+\beta}} \leq C\|f\|_{\mathcal{C}^\alpha}\|g\|_{\mathcal{C}^\beta}.$$

	$f \circledcirc g$	$g \circledcirc f$	$f \circledominus g$	fg
Regularity	$\alpha+\beta$	α	$\alpha+\beta$	α
Needs $\alpha+\beta>0$	No	No	Yes	Yes

TABLE 4.1. Summary of the regularity properties of paraproducts for $\alpha < 0 < \beta$, $f \in \mathcal{C}^\alpha$ and $g \in \mathcal{C}^\beta$.

The same analysis applies for the term $g \circledcirc f$, except that since we have $\beta > 0$, we get

$$\sum_{l=-1}^{k-2} 2^{-\alpha k - \beta l} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta} \le C 2^{-\alpha k} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta},$$

which suggests that

$$\|g \circledcirc f\|_{\mathcal{C}^\alpha} \le C \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

This analysis holds irrespectively of the assumption $\alpha + \beta > 0$. The term $f \circledcirc g$ can be interpreted as a modulation of the high frequency modes of g by the low frequency modes of f. This is in agreement with the fact that $f \circledcirc g$ is more regular than $g \circledcirc f$ under our hypothesis $\alpha < \beta$.

We now turn to the resonant term, which for simplicity we think of as being

$$\sum_{k=-1}^{+\infty} \delta_k f \, \delta_k g.$$

The crucial difference with the previous computations is that the spectrum of the summand indexed by k, which is the convolution of the annulus of radius about 2^k by itself, results in a *ball* of radius 2^k, as opposed to an annulus. Therefore, the Fourier modes of magnitude 2^l receive contributions from every summand indexed by $k \ge l$. The L^∞ norm of each summand is bounded by

$$2^{-(\alpha+\beta)k} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

If we want this to be summable over $k \ge l$, we need to assume $\alpha + \beta > 0$. In this case, the sum is of order $2^{-(\alpha+\beta)l} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}$, which suggests that

$$\|f \circledominus g\|_{\mathcal{C}^{\alpha+\beta}} \le C \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

These computations can all be made rigorous, and we summarize them in Table 4.1.

We now quantify the idea that the heat kernel has a regularizing effect on distributions, using the spaces $\mathcal{C}^\alpha$.

Proposition 4.2. *if $\alpha \ge \beta \in \mathbb{R}$, then there exists $C < \infty$ such that for every $t > 0$,*

$$\|e^{t\Delta} f\|_{\mathcal{C}^\alpha} \le C \, t^{\frac{\beta-\alpha}{2}} \, \|f\|_{\mathcal{C}^\beta}.$$

Sketch of proof. The Laplacian Δ is a multiplication operator in Fourier space. As a consequence, we have $\delta_k(e^{t\Delta} f) = e^{t\Delta}(\delta_k f)$, and since $\hat{\Delta}(\zeta) = -|\zeta|^2$, roughly speaking, we have $e^{t\Delta}(\delta_k f) \simeq e^{-t2^{2k}} \delta_k f$. This suggests that

$$\|\delta_k(e^{t\Delta} f)\|_{L^\infty} \le C \exp\left(-t2^{2k}\right) \|\delta_k f\|_{L^\infty} \le C \exp\left(-t2^{2k}\right) 2^{-\beta k} \|f\|_{\mathcal{C}^\beta},$$

and therefore

$$2^{k\alpha} \|\delta_k(e^{t\Delta} f)\|_{L^\infty} \le C \left[(2^{2k}t)^{\frac{\alpha-\beta}{2}} \exp\left(-2^{2k}t\right) \right] t^{\frac{\beta-\alpha}{2}} \|f\|_{\mathcal{C}^\alpha}.$$

Since the term between square brackets is bounded uniformly over k and t, the proof is complete. □

Equipped with Propositions 4.1 and 4.2, we can see why the equation (I.2.11) for the remainder Y is well-posed locally. Indeed, Proposition 4.2 indicates that the singularity in time produced by the increase of exponent of regularity is integrable provided that the difference of exponents is below 2. We can therefore expect the heat semi-group to induce an increase of regularity of (almost) two units. On the right side of (I.2.11), the processes $Z, Z^{:2:}$ and $Z^{:3:}$ are in $C(\mathbb{R}_+, \mathcal{C}^{-\alpha})$ for $\alpha > 0$ arbitrarily small. We can therefore expect to find Y as a continuous process taking values in $\mathcal{C}^{2-\alpha}$. If this is so, then we can indeed define the products $Y^2 Z$ and $Y Z^{:2:}$. This informal string of arguments can be turned into a proof via a classical fixed point argument.

The definition of Besov spaces can be extended, by replacing the choice of L^∞ and the supremum in (4.3) by more general L^p and ℓ^q norms. That is, for every $p, q \in [1, \infty]$ and $f \in C_c^\infty$, we can define

$$\|f\|_{\mathcal{B}^\alpha_{p,q}} := \left\| \left(2^{\alpha k} \|\delta_k f\|_{L^p} \right)_{k \geq -1} \right\|_{\ell^q}, \tag{4.7}$$

and then take $\mathcal{B}^\alpha_{p,q}$ as the completion of C_c^∞ with respect to this norm. This space can also be realized as a subspace of the space of Schwartz distributions, and satisfies properties very similar to those outlined above. The parameter $q \in [1, \infty]$ is a fine tuning of this scale of function spaces, which can be absorbed by varying α ever so slightly. Therefore, for the rest of this chapter we use the convention that

$$\mathcal{B}^\alpha_p := \mathcal{B}^\alpha_{p,\infty}. \tag{4.8}$$

The Besov spaces can also be defined on the torus $[-1, 1]^d$, with essentially no effect on the results presented above. For the stochastic processes we will consider (e.g. white noise or solutions of corresponding stochastic PDEs), Besov norms on the full space are infinite, for reasons unrelated to the regularity of the objects, but simply because they are stationary and unbounded. In this case, it becomes necessary to introduce weighted versions of the Besov spaces (see [MW15, Sections 2-4]).

5. Global well-posedness of the Φ^4 model in the plane

In this section, we discuss the result of [MW15], which shows that the Φ^4 equation is well-defined for all times and on the full two-dimensional plane.

As discussed on page 20, in our case, showing global existence on compact domains is essentially a precondition for being able to address even the local existence of the solution on the full space. We start by explaining how to show global existence of the solution on the unit torus.

Recall that the solution to the Φ^4 equation on the two-dimensional torus is defined as $Z + Y$, where Z solves the linearized equation (I.2.6), and Y solves

$$(\partial_t - \Delta) Y = -Y^3 - 3Y^2 Z - 3Y Z^{:2:} - Z^{:3:} + a(Y + Z). \tag{5.1}$$

The argument for non-explosion is entirely deterministic: we assume that we are given processes Z, $Z^{:2:}$ and $Z^{:3:}$ with suitable regularity, and show non-explosion of the corresponding solution. The probabilistic part of the argument is entirely contained in the statement the actual processes Z, $Z^{:2:}$ and $Z^{:3:}$ of interest to us satisfy these suitable regularity properties with probability one, and the proof of this was alluded to in Section 3. As was briefly sketched above, we have a local solution theory for equation (5.1). Moreover, this theory ensures local existence (and uniqueness) of the solution to (5.1) up to a random explosion time which can be bounded from below by a constant depending only on suitable norms of $(Z, Z^{:2:}, Z^{:3:})$ and $\|Y(t = 0)\|_{L^p}$, for some large $p < \infty$ (see [MW15, Theorem 6.2]).

Denote by $T^\star \in (0, \infty]$ the explosion time of the solution to (5.1). In order to ensure that this explosion time is infinite, it would suffice to show that for any given

$T_{\mathsf{max}} < \infty$, we can find a suitably large constant $C < \infty$ (depending on $(Z, Z^{:2:}, Z^{:3:})$ and on the initial condition $Y(t=0)$) such that

$$\sup_{t\in[0,T^\star\wedge T_{\mathsf{max}})} \|Y(t)\|_{L^p} \le C. \tag{5.2}$$

Indeed, if this bound holds, then we can iterate on the local solution theory with a time step that is uniformly bounded away from 0, and therefore conclude that $T^\star \ge T_{\mathsf{max}}$.

We now fix $t < T^\star \wedge T_{\mathsf{max}}$ and proceed to explain the argument for (5.2). For simplicity, we drop a few terms in (5.1) and consider only the simpler equation

$$(\partial_t - \Delta)Y = -Y^3 - 3YZ^{:2:} - Z^{:3:}. \tag{5.3}$$

The true equation (5.1) is handled in the same way. We need to devise an argument that leverages on the fact that the cubic non-linearity Y^3 comes with the right negative sign, since otherwise the equation is indeed expected to blow up in finite time. Multiplying the equation by Y, integrating in space and formally integrating by parts yields

$$\frac{1}{2}\partial_t\|Y(t)\|_{L^2}^2 + \|\nabla Y(t)\|_{L^2}^2 + \|Y(t)\|_{L^4}^4 = -\left\langle Y, 3YZ^{:2:} + Z^{:3:}\right\rangle(t),$$

where $\langle f, g\rangle = \int_{[-1,1]^2} fg$ for smooth f, g and is extended by continuity. This is not very rigorous, since the mapping $t \mapsto \|Y(t)\|_{L^2}$ is not differentiable. However, it is Hölder continuous for every exponent below 1, and in particular for some exponent above 1/2. Arguing as in the classical Young theory of integration against Hölder integrands (which is nothing but a simpler version of Proposition 4.1), we can justify the estimate above in the integrated sense of

$$\begin{aligned}\frac{1}{2}\left(\|Y(t)\|_{L^2}^2 - \|Y(0)\|_{L^2}^2\right) + \int_0^t \left(\|\nabla Y(s)\|_{L^2}^2 + \|Y(s)\|_{L^4}^4\right) \mathrm{d}s \\ = -\int_0^t \left\langle Y, 3YZ^{:2:} + Z^{:3:}\right\rangle(s)\,\mathrm{d}s.\end{aligned} \tag{5.4}$$

Roughly speaking, the mapping $(f, g) \mapsto \langle f, g\rangle$ extends to a continuous bilinear form on $\mathcal{B}_p^\alpha \times \mathcal{B}_{p'}^{-\alpha}$ for every $\alpha \in \mathbb{R}$ and conjugate exponents $p, p' \in [1, \infty]$. In fact, this is slightly inaccurate because we should also send the silenced additional exponent "q" to its dual (recall our convention (4.8)), but it suffices to augment one of the exponents α or $-\alpha$ by an arbitrarily small amount to make the statement correct. Since $Z^{:2:}$ and $Z^{:3:}$ are continuous processes taking values in $\mathcal{B}_\infty^{-\alpha}$ for $\alpha > 0$ arbitrarily small, we obtain

$$-\int_0^t \left\langle Y, 3YZ^{:2:} + Z^{:3:}\right\rangle(s)\,\mathrm{d}s \le C_0 \int_0^t \left(\|Y^2(s)\|_{\mathcal{B}_1^\alpha} + \|Y(s)\|_{\mathcal{B}_1^\alpha}\right)\mathrm{d}s,$$

for some constant $C_0 < \infty$ (depending on $Z^{:2:}$ and $Z^{:3:}$ up to time T_{max}). Moreover, one can show (see [MW15, Proposition 3.25]) that for each given $\alpha \in (0, 1)$,

$$\|f\|_{\mathcal{B}_1^\alpha} \lesssim \|\nabla f\|_{L^1}^\alpha \|f\|_{L^1}^{1-\alpha} + \|f\|_{L^1} \lesssim \|\nabla f\|_{L^1} + \|f\|_{L^1},$$

and therefore

$$\begin{aligned}\|Y^2(s)\|_{\mathcal{B}_1^\alpha} + \|Y(s)\|_{\mathcal{B}_1^\alpha} &\lesssim \|Y\nabla Y\|_{L^1} + \|Y^2\|_{L^1} + \|\nabla Y\|_{L^1} + \|Y\|_{L^1} \\ &\le \frac{1}{2C_0}\|\nabla Y(s)\|_{L^2}^2 + C + C\|Y(s)\|_{L^2}^2,\end{aligned}$$

for a large constant C. The last line is obtained by applying Hölder's and Young's inequalities. Using this in (5.4) and Jensen's inequality yields, after a redefinition

of the constant C,

$$\frac{1}{2}\left(\|Y(t)\|_{L^2}^2 - \|Y(0)\|_{L^2}^2\right) + \int_0^t \left(\|\nabla Y(s)\|_{L^2}^2 + \|Y(s)\|_{L^4}^4\right) \mathrm{d}s \le \frac{1}{2}\int_0^t \|\nabla Y(s)\|_{L^2}^2 \,\mathrm{d}s + C + C\left(\int_0^t \|Y(s)\|_{L^4}^4 \,\mathrm{d}s\right)^{\frac{1}{2}}$$

and therefore,

$$\|Y(t)\|_{L^2}^2 - \|Y(0)\|_{L^2}^2 + \int_0^t \left(\|\nabla Y(s)\|_{L^2}^2 + \|Y(s)\|_{L^4}^4\right) \mathrm{d}s \lesssim 1. \tag{5.5}$$

In other words, we used the good terms $\|\nabla Y(s)\|_{L^2}^2$ and $\|Y(s)\|_{L^4}^4$ to control the multiplication of Y by $Z^{:2:}$ and the multiplication of Y^2 by Z. Note that we needed to use the term with ∇Y due to the irregularity of $Z^{:2:}$ and $Z^{:3:}$. Had they been processes taking values in L^∞, we would not have needed to use this term to our advantage. The estimate (5.5) implies (5.2) for $p = 2$, while we wanted to have it for a very large p. This stronger estimate can be obtained in the same way by testing the equation with Y^{p-1}, where p is a large even integer.

This concludes the proof of global existence of the solution on the torus. In order to show existence of a solution on the full space, we then proceed as follows. First, we think of functions on a given torus $[-M, M]^2$ as $(2M)\mathbb{Z}^2$-periodic functions on the full space $\mathbb{R}^2$. For each M, we periodize the noise ξ over $[-M, M]^2$, define the associated renormalized powers, and solve the corresponding equation for the remainder, whose solution we denote by Y_M. We introduce Besov and L^p spaces with polynomially decreasing weights, and show that we can reproduce the argument above with the weighted spaces. The key point is that we can obtain a priori weighted L^p bounds on Y_M that hold uniformly over M. Using the mild formulation of the equation, we show next that we can upgrade the a priori L^p to an a priori estimate with respect to a weighted Besov norm. This enables to pass to the limit in the sequence of equations for Y_M, and therefore to ensure the existence of a solution on the full space.

The uniqueness of the solution on the full space also requires justification. This is a priori difficult because we cannot write down a self-contained Gronwall-type inequality bounding some weighted norm of the solution by a convolution of the same norm, due to the non-linearity. We solve this problem by using an infinite scale of weights, controlling the growth of the norms when the weights become flatter and flatter, and iterating a non-self-contained Gronwall-type bound infinitely many times. See [MW15, (9.15)] for the unfolding of a Besov norm with polynomial weights into an infinite scale of Besov norms with stretched exponential weights and controlled growth, and [MW15, Proposition 9.3] for the Gronwall-type inequality we iterate over. Working with stretched exponential weights requires to be more careful than usual when defining the corresponding Besov spaces, since we need to choose particularly smooth, almost analytic functions with compact support in the definition of the spaces, see [MW15, Definition 2.1].

6. Global well-posedness of the Φ^4 model in three dimensions

In this section, we review [MW16], which considers the same question of global existence of the solution to the Φ^4 equation, but now for the three-dimensional torus. In this case, the renormalization procedure is much more involved. In order to better keep track of it, we change notation and now write the renormalized powers of the solution to the linear equation (I.2.6) as

$$\uparrow := Z, \qquad \vee := Z^{:2:}, \qquad \Psi := Z^{:3:}.$$

In these symbols, a dot stands for an instance of white noise, a line stands for an integration against the heat kernel, and the confluence of lines stands for (renormalized) multiplication. We will perform formal manipulations ("renormalizations") on the equation until we reach an equation that has a chance of being mathematically meaningful. In view of (5.1), our starting point is to postulate an expansion of the solution to the Φ^4 equation (I.2.2) as $X = \,\uparrow - \,\Psi + u$, where Ψ solves

$$(\partial_t - \Delta)\Psi = \,\Psi.$$

Then u formally solves

$$(\partial_t - \Delta)u = -(u + \,\uparrow - \Psi)^3 + m(u + \,\uparrow - \Psi) - \,\Psi \tag{6.1}$$

$$= -u^3 - 3(u - \Psi)\,\mathcal{V} + Q(u), \tag{6.2}$$

where we introduced the notation

$$Q(u) = b_0 + b_1 u + b_2 u^2,$$

with

$$b_0 = m(\,\uparrow - \,\Psi) + (\Psi)^3 - 3\,\uparrow (\Psi)^2,$$
$$b_1 = m + 6\,\Psi\,\uparrow - 3(\Psi)^2,$$
$$b_2 = -3\,\uparrow + 3\,\Psi.$$

We wrote an equality sign when going from (6.1) to (6.2), but note that we used the "renormalization rule" $(\uparrow)^3 \rightsquigarrow \,\Psi$, so we should rather understand this step as a first renormalization of divergences. Now $\uparrow$ has regularity $(-\frac{1}{2})^-$ (since the white noise has regularity $(-\frac{5}{2})^-$, and integration allows to gain two degrees of regularity), while $\mathcal{V}$ has regularity $(-1)^-$ and Ψ has regularity $(\frac{1}{2})^-$. In view of (6.2), we cannot expect the regularity of u to be better than 1^-. Therefore, and in contrast to the two-dimensional situation, equation (6.2) is still ill-defined, since the product $u\,\mathcal{V}$ has no canonical interpretation. Moreover, we cannot continue applying the same trick of subtracting a suitable diagram and continuing the expansion, since this ill-defined term involves the function u itself. The main point of the theories based on regularity structures [126] or on paraproducts [120] is to explain how to continue the expansion nonetheless. The idea based on paraproducts is that the term $u \varodot \mathcal{V}$ is always well-defined and carries the most irregular part of any potential candidate for the product $u\,\mathcal{V}$, see Table 4.1. We therefore postulate an expansion of u into $v + w$, where v and w solve the system of equations

$$(\partial_t - \Delta)v = -3(v + w - \Psi) \varodot \mathcal{V}, \tag{6.3}$$

$$(\partial_t - \Delta)w = -(v + w)^3 - 3(v + w - \Psi) \circledcirc \mathcal{V} + Q(v + w), \tag{6.4}$$

where we write $\circledcirc = \varodot + \circleddash$ for concision. Now (recall Table 4.1), we expect w to be of regularity $(\frac{3}{2})^-$ instead of 1^-. On the right side of (6.4), only the terms $v \circleddash \mathcal{V}$ and $\Psi \circleddash \mathcal{V}$ are ill-defined. We can construct a suitable version of the second term, which we denote by $\Psi \varodot \mathcal{V}$, in much the same way as for renormalized Wick powers (that is, by an argument based on explicit covariance computations and Gaussianity). The proper definition of the term $v \circleddash \mathcal{V}$ brings about certain commutator terms, which we denote by $\mathsf{com}(v, w)$. We refer to [MW16, Section 1] for a more thorough discussion. We end up with the following system of equations:

$$\begin{cases} (\partial_t - \Delta)v = F(v + w), \\ (\partial_t - \Delta)w = G(v, w), \end{cases} \tag{6.5}$$

where F and G are defined by

$$F(v+w) := -3(v+w-\Vdots)\circledcirc\Vvee, \tag{6.6}$$

$$G(v,w) := -(v+w)^3 - 3\mathsf{com}(v,w) - 3w\circledcirc\Vvee - 3(v+w-\Vdots)\circledcirc\Vvee + P(v+w), \tag{6.7}$$

with

$$P(v+w) = a_0 + a_1(v+w) + a_2(v+w)^2, \tag{6.8}$$

and where a_0, a_1, a_2 are defined explicitly in terms of diagrams. The precise definition of these coefficients a_0, a_1, a_2 is irrelevant, the only important point being that they all have regularity $(-\frac{1}{2})^-$. The system of equations in (6.5) can now be solved locally in time by means of a fixed point argument, as was shown in a slightly different formulation by Catellier and Chouk [57].

As in the two-dimensional case, in order to show global existence of a solution, we must leverage on the fact that the cubic non-linearity has the right sign. Our tool for doing so is to test the equation for w against w itself (or more generally, against $w|w|^{p-2}$ for arbitrary p). This can only succeed if we have left (as we did) the most irregular part of the sum $v+w$ in v. Indeed, as was explained in the previous section, the testing argument brings about the time integral of $\|\nabla w\|_{L^2}^2$, and this term would not be finite if w was replaced by v.

When testing the equation for w against w itself, several non-linear terms do not appear with an obvious sign, an instance of this being the term $-wv^3$. We do not know how to control this term. In order to circumvent the difficulty, we modify the system of equations (6.5) into

$$\begin{cases} (\partial_t - \Delta)v = F(v+w) - cv, \\ (\partial_t - \Delta)w = G(v,w) + cv, \end{cases} \tag{6.9}$$

for a sufficiently large constant $c \geq 0$. While this changes the definitions of the functions v and w, it does *not* alter the sum $v+w$, as can be seen by considering a smooth approximation of the equation. Roughly speaking, we can then show that if c is sufficiently large, then a suitable norm of v is controlled by a very small multiple of a suitable norm of w. How large c must be chosen depends on the size of the diagrams $\Vvee$ and $\Vdots$ up to our target time T_{max}. This technique effectively allows us to reduce the system of equations (6.9) to a single equation on w.

The last sentence must however be taken with a grain of salt. Indeed, the system (6.9) behaves very differently from the two-dimensional equation (5.1) in the following sense. In the two-dimensional case, by being slightly more careful in the derivation of (5.5), one can obtain an integrated form of the inequality

$$\partial_t \|Y(t)\|_{L^2}^2 + \|Y(t)\|_{L^2}^4 \lesssim 1,$$

where the implicit constant depends only on suitable norms of $\I$, $\Vvee$ and $\Vdots$. By comparing with the ordinary differential equation $\dot{y} + y^2 \lesssim 1$, we see that the solution to the two-dimensional equation "goes down from infinity", in the sense that we can write down an a priori estimate on $\|Y(1)\|_{L^2}$ that depends on $\I$, $\Vvee$ and $\Vdots$, but that does *not* depend on the initial condition $Y(t=0)$.

While this property of coming down from infinity may very well be true for the solution to the three-dimensional Φ^4 equation, it is *not* true of the pair (v,w) solving (6.9). Indeed, there is no non-linear mechanism that can bring v down from infinity. In other words, if we start the system (6.9) with very large initials conditions for v and w, then v will at best relax only exponentially fast to a moderate value, while w will first "go down from infinity" to ensure that $v+w$ is moderate, and then relax at the pace of v while preserving this property. From a technical point of view, this

is reflected in the fact that the estimates we obtain when testing the equation for w against w itself will always contain a norm of the initial condition for v with a cubic non-linearity.

From now on, we ignore this point, and proceed to explain how to show that w does not explode, pretending that it solves a closed equation not involving v. We have explained how to control the cubic non-linearity appearing in G. The main difficulty with the commutator term is that in order to show that it is well-defined, and then to estimate its size, we need to use information on the time regularity of w itself (as well as v and $\mathbb{V}$). The next step of our approach consists in using the mild formulation of the equation to estimate the size of the time increments of w. We give a bound on $(w(t) - w(s))/|t - s|^{\frac{1}{8}}$ in terms of various norms of w, see [MW16, Theorem 4.1]. This can then be used to estimate the commutator term $\mathsf{com}(v, w)$.

The quadratic non-linearity in G taking the form $a_2(v + w)^2$ is surprisingly challenging, because of the low regularity of the term a_2. In order to control the term $a_2 w^2$, we need to appeal to the "good term" obtained when testing the equation against w^3, which takes the form of a time integral of

$$\|w^2|\nabla w|^2\|_{L^1} = \|\nabla(w^2)\|_{L^2}^2.$$

The cross-term $a_2 vw$ also requires attention, since it cannot be controlled by a combination of $a_2 v^2$ and $a_2 w^2$.

Another important difficulty is caused by the presence of the term $w \circledcirc \mathbb{V}$ appearing in G. Indeed, recall that $\mathbb{V}$ has regularity $(-1)^-$, hence defining (and estimating) the resonant term $w \circledcirc \mathbb{V}$ requires that w be of regularity strictly above 1. In order to control the regularity of w, the testing argument can only produce "good terms" of the form $\|w^{p-2}|\nabla w|^2\|_{L^1}$, and is therefore bound to fail, since these good terms only control one derivative of w, while we need a little more. At this stage, we simply "give up" on this term and accept to have a non-self-contained estimate. What the testing argument does provide us is a bound of some norm of w^3 in terms of a (stronger) norm of w: that is, it allows us to estimate a non-linear quantity in terms of a linear one.

Once this is achieved, we have effectively reduced our non-linear equation for w into a linear one. A Gronwall-type argument allows us to conclude that w (and therefore v) does not explode in finite time. We need to take care of the fact that our bounds depend relatively strongly on initial conditions; we integrate the estimate over the initial time to control this aspect.

At this point, we have finally obtained self-contained bounds on some norms of v and w. Since we had to leverage on the good terms $\|\nabla w\|_{L^2}^2$ and $\|\nabla(w^2)\|_{L^2}^2$, we had to work with spaces of fairly low exponent of integrability (unlike in the two-dimensional case). In order to complete the argument and link with the requirements of the local existence theory, we need to upgrade these estimates. We do this by using the mild formulation of the equation, iteratively improving on the quality of the norms controlling v and w.

In work in preparation, Hendrik Weber and I plan to revisit this string of arguments to obtain that v and w are tight as time tends to infinity. By the Krylov–Bogolyubov argument, this would provide us with a dynamical construction of the Φ^4 measure in three dimensions.

Bibliography presented for the HDR

[AKM15] S. N. Armstrong, T. Kuusi, and J.-C. Mourrat. Mesoscopic higher regularity and subadditivity in elliptic homogenization. *Comm. Math. Phys.*, in press, arXiv:1507.06935.

[AKM16a] S. N. Armstrong, T. Kuusi, and J.-C. Mourrat. The additive structure of elliptic homogenization, preprint, arXiv:1602.00512.

[AKM16b] S. N. Armstrong, T. Kuusi, and J.-C. Mourrat. Scaling limits of energies and correctors, preprint, arXiv:1603.03388.

[AM16] S. N. Armstrong and J.-C. Mourrat. Lipschitz regularity for elliptic equations with random coefficients. *Arch. Ration. Mech. Anal.*, 219(1):255–348, 2016.

[CMMV14] M. Cranston, T. Mountford, J.-C. Mourrat, and D. Valesin. The contact process on finite homogeneous trees revisited. *ALEA Lat. Am. J. Probab. Math. Stat.*, 11(1):385–408, 2014.

[dBM15] P. de Buyer and J.-C. Mourrat. Diffusive decay of the environment viewed by the particle. *Electron. Commun. Probab.*, 20:no. 23, 12, 2015.

[EGMN15] A.-C. Egloffe, A. Gloria, J.-C. Mourrat, and T. N. Nguyen. Random walk in random environment, corrector equation and homogenized coefficients: from theory to numerics, back and forth. *IMA J. Numer. Anal.*, 35(2):499–545, 2015.

[GM12] A. Gloria and J.-C. Mourrat. Spectral measure and approximation of homogenized coefficients. *Probab. Theory Related Fields*, 154(1-2):287–326, 2012.

[GM13] A. Gloria and J.-C. Mourrat. Quantitative version of the Kipnis-Varadhan theorem and Monte Carlo approximation of homogenized coefficients. *Ann. Appl. Probab.*, 23(4):1544–1583, 2013.

[GM14] Y. Gu and J.-C. Mourrat. Pointwise two-scale expansion for parabolic equations with random coefficients. *Probab. Theory Related Fields*, in press, arXiv:1410.2157.

[GM15] Y. Gu and J.-C. Mourrat. Scaling limit of fluctuations in stochastic homogenization. *Multiscale Model. Simul.*, in press, arXiv:1503.00578.

[GM16a] A. Giunti and J.-C. Mourrat. Quantitative homogenization of degenerate random environments, preprint, arXiv:1602.04467.

[GM16b] Y. Gu and J.-C. Mourrat. On generalized Gaussian free fields and stochastic homogenization, preprint, arXiv:1601.06408.

[MM13] T. Mountford and J.-C. Mourrat. Lyapunov exponents of random walks in small random potential: the lower bound. *Comm. Math. Phys.*, 323(3):1071–1120, 2013.

[MM15a] P. Mathieu and J.-C. Mourrat. Aging of asymmetric dynamics on the random energy model. *Probab. Theory Related Fields*, 161(1-2):351–427, 2015.

[MM15b] T. Mountford and J.-C. Mourrat. Lyapunov exponents of random walks in small random potential: the upper bound. *Electron. J. Probab.*, 20:no. 49, 18, 2015.

[MMVY16] T. Mountford, J.-C. Mourrat, D. Valesin, and Q. Yao. Exponential extinction time of the contact process on finite graphs. *Stochastic Process. Appl.*, 126(7):1974–2013, 2016.

[MN15] J.-C. Mourrat and J. Nolen. Scaling limit of the corrector in stochastic homogenization. *Ann. Appl. Probab.*, in press, arXiv:1502.07440.

[MO16] J.-C. Mourrat and F. Otto. Anchored Nash inequalities and heat kernel bounds for static and dynamic degenerate environments. *J. Funct. Anal.*, 270(1):201–228, 2016.

[MO14] J.-C. Mourrat and F. Otto. Correlation structure of the corrector in stochastic homogenization. *Ann. Probab.*, in press, arXiv:1402.1924.

[Mou10] J.-C. Mourrat. Principal eigenvalue for the random walk among random traps on $\mathbb{Z}^d$. *Potential Anal.*, 33(3):227–247, 2010.

[Mou11a] J.-C. Mourrat. Scaling limit of the random walk among random traps on $\mathbb{Z}^d$. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(3):813–849, 2011.

[Mou11b] J.-C. Mourrat. Variance decay for functionals of the environment viewed by the particle. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(1):294–327, 2011.

[Mou12a] J.-C. Mourrat. Lyapunov exponents, shape theorems and large deviations for the random walk in random potential. *ALEA Lat. Am. J. Probab. Math. Stat.*, 9:165–211, 2012.

[Mou12b] J.-C. Mourrat. On the delocalized phase of the random pinning model. In *Séminaire de Probabilités XLIV*, volume 2046 of *Lecture Notes in Math.*, pages 401–407. Springer, Heidelberg, 2012.

[Mou12c] J.-C. Mourrat. A quantitative central limit theorem for the random walk among random conductances. *Electron. J. Probab.*, 17:no. 97, 17, 2012.

[Mou13] J.-C. Mourrat. On the rate of convergence in the martingale central limit theorem. *Bernoulli*, 19(2):633–645, 2013.

[Mou14] J.-C. Mourrat. Kantorovich distance in the martingale CLT and quantitative homogenization of parabolic equations with random coefficients. *Probab. Theory Related Fields*, 160(1-2):279–314, 2014.

[Mou15a] J.-C. Mourrat. First-order expansion of homogenized coefficients under Bernoulli perturbations. *J. Math. Pures Appl. (9)*, 103(1):68–101, 2015.

[Mou15b] J.-C. Mourrat. A tightness criterion in local Hölder spaces of negative regularity, preprint, arXiv:1502.07335.

[MV16] J.-C. Mourrat and D. Valesin. Phase transition of the contact process on random regular graphs. *Electron. J. Probab.*, 21:no. 31, 17, 2016.

[MW14] J.-C. Mourrat and H. Weber. Convergence of the two-dimensional dynamic Ising-Kac model to Φ^4_2. *Comm. Pure Appl. Math.*, in press, arXiv:1410.1179.

[MW15] J.-C. Mourrat and H. Weber. Global well-posedness of the dynamic Φ^4 model in the plane. *Ann. Probab.*, in press, arXiv:1501.06191.

[MW16] J.-C. Mourrat and H. Weber. Global well-posedness of the dynamic Φ^4_3 model on the torus, preprint, arXiv:1601.01234.

Bibliography

[1] M. Aizenman and S. Molchanov. Localization at large disorder and at extreme energies: an elementary derivation. *Comm. Math. Phys.*, 157(2):245–278, 1993.

[2] M. Aizenman, J. H. Schenker, R. M. Friedrich, and D. Hundertmark. Finite-volume fractional-moment criteria for Anderson localization. *Comm. Math. Phys.*, 224(1):219–253, 2001. Dedicated to Joel L. Lebowitz.

[3] S. Albeverio and M. Röckner. Stochastic differential equations in infinite dimensions: solutions via Dirichlet forms. *Probab. Theory Related Fields*, 89(3):347–386, 1991.

[4] K. S. Alexander. Approximation of subadditive functions and convergence rates in limiting-shape results. *Ann. Probab.*, 25(1):30–55, 1997.

[5] K. S. Alexander and N. Zygouras. Path properties of the disordered pinning model in the delocalized regime. *Ann. Appl. Probab.*, 24(2):599–615, 2014.

[6] R. Allez and K. Chouk. The continuous Anderson hamiltonian in dimension two, preprint, arXiv:1511.02718.

[7] Y. Almog. Averaging of dilute random media: a rigorous proof of the Clausius-Mossotti formula. *Arch. Ration. Mech. Anal.*, 207(3):785–812, 2013.

[8] G. Amir, I. Corwin, and J. Quastel. Probability distribution of the free energy of the continuum directed random polymer in 1+1 dimensions. *Comm. Pure Appl. Math.*, 64(4):466–537, 2011.

[9] A. Anantharaman and C. Le Bris. Homogénéisation d'un matériau périodique faiblement perturbé aléatoirement. *C. R. Math. Acad. Sci. Paris*, 348(9-10):529–534, 2010.

[10] A. Anantharaman and C. Le Bris. A numerical approach related to defect-type theories for some weakly random problems in homogenization. *Multiscale Model. Simul.*, 9(2):513–544, 2011.

[11] A. Anantharaman and C. Le Bris. Elements of mathematical foundations for numerical approaches for weakly random homogenization problems. *Commun. Comput. Phys.*, 11(4):1103–1143, 2012.

[12] P. W. Anderson. Absence of diffusion in certain random lattices. *Phys. Rev.*, 109(5):1492, 1958.

[13] S. Andres, M. T. Barlow, J.-D. Deuschel, and B. M. Hambly. Invariance principle for the random conductance model. *Probab. Theory Related Fields*, 156(3-4):535–580, 2013.

[14] S. N. Armstrong and C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér.*, in press, arXiv:1406.0996.

[15] S. N. Armstrong and H. V. Tran. Stochastic homogenization of viscous Hamilton-Jacobi equations and applications. *Anal. PDE*, 7(8):1969–2007, 2014.

[16] S. Assing. A pregenerator for Burgers equation forced by conservative noise. *Comm. Math. Phys.*, 225(3):611–632, 2002.

[17] M. Avellaneda and F.-H. Lin. Compactness methods in the theory of homogenization. *Comm. Pure Appl. Math.*, 40(6):803–847, 1987.

[18] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Heidelberg, 2011.

[19] J. Baik, P. Deift, and K. Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12(4):1119–1178, 1999.

[20] I. Bailleul, F. Bernicot, and D. Frey. Higher order paracontrolled calculus, 3d-PAM and multiplicative burgers equations, preprint, arXiv:1506.08773.

[21] M. T. Barlow. Random walks on supercritical percolation clusters. *Ann. Probab.*, 32(4):3024–3084, 2004.

[22] R. Bauerschmidt, D. C. Brydges, and G. Slade. Scaling limits and critical behaviour of the 4-dimensional n-component $|\phi|^4$ spin model. *J. Stat. Phys.*, 157(4-5):692–742, 2014.

[23] G. Ben Arous, A. Bovier, and J. Černý. Universality of random energy model-like ageing in mean field spin glasses. *J. Stat. Mech. Theory Exp.*, (4):L04003, 8, 2008.

[24] G. Ben Arous, A. Bovier, and J. Černý. Universality of the REM for dynamics of mean-field spin glasses. *Comm. Math. Phys.*, 282(3):663–695, 2008.

[25] G. Ben Arous, A. Bovier, and V. Gayrard. Glauber dynamics of the random energy model. I. Metastable motion on the extreme states. *Comm. Math. Phys.*, 235(3):379–425, 2003.

[26] G. Ben Arous, A. Bovier, and V. Gayrard. Glauber dynamics of the random energy model. II. Aging below the critical temperature. *Comm. Math. Phys.*, 236(1):1–54, 2003.

[27] G. Ben Arous and J. Černý. The arcsine law as a universal aging scheme for trap models. *Comm. Pure Appl. Math.*, 61(3):289–329, 2008.

[28] G. Ben Arous, A. Dembo, and A. Guionnet. Aging of spherical spin glasses. *Probab. Theory Related Fields*, 120(1):1–67, 2001.

[29] G. Ben Arous, A. Dembo, and A. Guionnet. Cugliandolo-Kurchan equations for dynamics of spin-glasses. *Probab. Theory Related Fields*, 136(4):619–660, 2006.

[30] I. Benjamini. *Coarse geometry and randomness*, volume 2100 of *Lecture Notes in Mathematics*. Springer, Cham, 2013. Lecture notes from the 41st Probability Summer School held in Saint-Flour, 2011.

[31] I. Benjamini, H. Duminil-Copin, G. Kozma, and A. Yadin. Disorder, entropy and harmonic functions. *Ann. Probab.*, 43(5):2332–2373, 2015.

[32] I. Benjamini, A. Nachmias, and Y. Peres. Is the critical percolation probability local? *Probab. Theory Related Fields*, 149(1-2):261–269, 2011.

[33] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13 pp. (electronic), 2001.

[34] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*, volume 5 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam-New York, 1978.

[35] N. Berger and M. Biskup. Quenched invariance principle for simple random walk on percolation clusters. *Probab. Theory Related Fields*, 137(1-2):83–120, 2007.

[36] L. Bertini and G. Giacomin. Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.*, 183(3):571–607, 1997.

[37] L. Bertini and B. Zegarlinski. Coercive inequalities for Kawasaki dynamics. The product case. *Markov Process. Related Fields*, 5(2):125–162, 1999.

[38] C. Bezuidenhout and G. Grimmett. The critical contact process dies out. *Ann. Probab.*, 18(4):1462–1482, 1990.

[39] M. Birkner, A. Greven, and F. den Hollander. Quenched large deviation principle for words in a letter sequence. *Probab. Theory Related Fields*, 148(3-4):403–456, 2010.

[40] M. Biskup. Reflection positivity and phase transitions in lattice spin models. In *Methods of contemporary mathematical statistical physics*, volume 1970 of *Lecture Notes in Math.*, pages 1–86. Springer, Berlin, 2009.

[41] M. Biskup and T. M. Prescott. Functional CLT for random walk among bounded random conductances. *Electron. J. Probab.*, 12:no. 49, 1323–1348, 2007.

[42] A. Borodin, I. Corwin, and T. Sasamoto. From duality to determinants for q-TASEP and ASEP. *Ann. Probab.*, 42(6):2314–2382, 2014.

[43] J.-P. Bouchaud. Weak ergodicity breaking and aging in disordered systems. *J. Phys. I (France)*, 2(9):1705–1713, 1992.

[44] J.-P. Bouchaud and D. Dean. Aging on Parisi's tree. *J. Phys. I (France)*, 5(3):265–286, 1995.

[45] A. Bourgeat and A. Piatnitski. Approximations of effective coefficients in stochastic homogenization. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(2):153–165, 2004.

[46] A. Bovier and V. Gayrard. Convergence of clock processes in random environments and ageing in the p-spin SK model. *Ann. Probab.*, 41(2):817–847, 2013.

[47] A. Bovier and I. Kurkova. Much ado about Derrida's GREM. In *Spin glasses*, volume 1900 of *Lecture Notes in Math.*, pages 81–115. Springer, Berlin, 2007.

[48] H. Brézis and I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282(17):Aii, A971–A974, 1976.

[49] D. C. Brydges, J. Fröhlich, and A. D. Sokal. A new proof of the existence and nontriviality of the continuum φ_2^4 and φ_3^4 quantum field theories. *Comm. Math. Phys.*, 91(2):141–186, 1983.

[50] F. Camia, C. Garban, and C. M. Newman. Planar Ising magnetization field I. Uniqueness of the critical scaling limit. *Ann. Probab.*, 43(2):528–571, 2015.

[51] S. Campanato. Proprietà di hölderianità di alcune classi di funzioni. *Ann. Scuola Norm. Sup. Pisa (3)*, 17:175–188, 1963.

[52] S. Campanato. Equazioni ellittiche del II° ordine e spazi $\mathfrak{L}^{(2,\lambda)}$. *Ann. Mat. Pura Appl. (4)*, 69:321–381, 1965.

[53] V. H. Can. Metastability for the contact process on the preferential attachment graph, preprint, arXiv:1502.05633.

[54] V. H. Can and B. Schapira. Metastability for the contact process on the configuration model with infinite mean degree, preprint, arXiv:1410.3061.

[55] F. Caravenna, R. Sun, and N. Zygouras. Universality in marginally relevant disordered systems, preprint, arXiv:1510.06287.

[56] M. Cassandro, R. Marra, and E. Presutti. Upper bounds on the critical temperature for Kac potentials. *Journal of statistical physics*, 88(3-4):537–566, 1997.

[57] R. Catellier and K. Chouk. Paracontrolled distributions and the 3-dimensional stochastic quantization equation, preprint, arXiv:1310.6869.

[58] J. Černý and T. Wassmer. Aging of the Metropolis dynamics on the random energy model. *Probab. Theory Related Fields*, in press, arXiv:1502.04535.

[59] S. Chatterjee and R. Durrett. Contact processes on random graphs with power law degree distributions have critical value 0. *Ann. Probab.*, 37(6):2332–2356, 2009.

[60] D. Cheliotis and F. den Hollander. Variational characterization of the critical curve for pinning of random polymers. *Ann. Probab.*, 41(3B):1767–1805, 2013.

[61] D. Chelkak, H. Duminil-Copin, C. Hongler, A. Kemppainen, and S. Smirnov. Convergence of Ising interfaces to Schramm's SLE curves. *C. R. Math. Acad. Sci. Paris*, 352(2):157–161, 2014.

[62] D. Chelkak, C. Hongler, and K. Izyurov. Conformal invariance of spin correlations in the planar Ising model. *Ann. of Math. (2)*, 181(3):1087–1138, 2015.

[63] J. W. Chen. The contact process on a finite system in higher dimensions. *Chinese J. Contemp. Math.*, 15(1):13–20, 1994.

[64] I. Corwin, J. Quastel, and D. Remenik. Renormalization fixed point of the KPZ universality class, preprint, arXiv:1103.3422.

[65] I. Corwin and L.-C. Tsai. KPZ equation limit of higher-spin exclusion processes, preprint, arXiv:1505.04158.

[66] J. T. Cox and R. Durrett. Some limit theorems for percolation processes with necessary and sufficient conditions. *Ann. Probab.*, 9(4):583–603, 1981.

[67] L. Cugliandolo and J. Kurchan. Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model. *Phys. Rev. Lett.*, 71(1):173, 1993.

[68] G. Da Prato and A. Debussche. Strong solutions to the stochastic quantization equations. *Ann. Probab.*, 31(4):1900–1916, 2003.

[69] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization. *Ann. Mat. Pura Appl. (4)*, 144:347–389, 1986.

[70] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization and ergodic theory. *J. Reine Angew. Math.*, 368:28–42, 1986.

[71] M. Damron, J. Hanson, and P. Sosoe. Subdiffusivity of random walk on the 2D invasion percolation cluster. *Stochastic Process. Appl.*, 123(9):3588–3621, 2013.

[72] E. De Giorgi. Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3)*, 3:25–43, 1957.

[73] A. De Masi, P. A. Ferrari, S. Goldstein, and W. D. Wick. Invariance principle for reversible Markov processes with application to diffusion in the percolation regime. In *Particle systems, random media and large deviations (Brunswick, Maine, 1984)*, volume 41 of *Contemp. Math.*, pages 71–85. Amer. Math. Soc., Providence, RI, 1985.

[74] A. De Masi, P. A. Ferrari, S. Goldstein, and W. D. Wick. An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Statist. Phys.*, 55(3-4):787–855, 1989.

[75] T. Delmotte and J.-D. Deuschel. On estimating the derivatives of symmetric diffusions in stationary random environment, with applications to $\nabla\phi$ interface model. *Probab. Theory Related Fields*, 133(3):358–390, 2005.

[76] A. Dembo, A. Guionnet, and C. Mazza. Limiting dynamics for spherical models of spin glasses at high temperature. *J. Stat. Phys.*, 126(4-5):781–815, 2007.

[77] A. Dembo and L.-C. Tsai. Weakly Asymmetric Non-Simple Exclusion Process and the Kardar–Parisi–Zhang Equation. *Comm. Math. Phys.*, 341(1):219–261, 2016.

[78] B. Derrida. Random-energy model: Limit of a family of disordered models. *Phys. Rev. Lett.*, 45(2):79, 1980.

[79] B. Derrida. Random-energy model: An exactly solvable model of disordered systems. *Phys. Rev. B*, 24(5):2613, 1981.

[80] B. Derrida and M. Retaux. The depinning transition in presence of disorder: A toy model. *J. Stat. Phys.*, in press, arXiv:1401.6919.

[81] M. Disertori, C. Sabot, and P. Tarrès. Transience of edge-reinforced random walk. *Comm. Math. Phys.*, 339(1):121–148, 2015.

[82] M. Disertori, T. Spencer, and M. R. Zirnbauer. Quasi-diffusion in a 3D supersymmetric hyperbolic sigma model. *Comm. Math. Phys.*, 300(2):435–486, 2010.

[83] M. Duerinckx, A. Gloria, and F. Otto. The structure of fluctuations in stochastic homogenization, preprint, arXiv:1602.01717.

[84] R. Durrett and X. F. Liu. The contact process on a finite set. *Ann. Probab.*, 16(3):1158–1173, 1988.

[85] R. Durrett and R. H. Schonmann. The contact process on a finite set. II. *Ann. Probab.*, 16(4):1570–1583, 1988.

[86] R. Durrett, R. H. Schonmann, and N. I. Tanaka. The contact process on a finite set. III. The critical case. *Ann. Probab.*, 17(4):1303–1321, 1989.

[87] J.-P. Eckmann and K. Osterwalder. On the uniqueness of the Hamiltonian and of the representation of the CCR for the quartic boson interaction in three dimensions. *Helv. Phys. Acta*, 44:884–909, 1971.

[88] A. Elgart. Lifshitz tails and localization in the three-dimensional Anderson model. *Duke Math. J.*, 146(2):331–360, 2009.

[89] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[90] A. Faggionato, F. Martinelli, C. Roberto, and C. Toninelli. The East model: recent results and new progresses. *Markov Process. Related Fields*, 19(3):407–452, 2013.

[91] J. Feldman. The $\lambda\varphi^4_3$ field theory in a finite volume. *Comm. Math. Phys.*, 37:93–120, 1974.

[92] J. S. Feldman and K. Osterwalder. The Wightman axioms and the mass gap for weakly coupled $(\Phi^4)_3$ quantum field theories. *Ann. Physics*, 97(1):80–135, 1976.

[93] J. Fischer and F. Otto. A higher-order large-scale regularity theory for random elliptic operators, preprint, arXiv:1503.07578.

[94] S. Fitzpatrick. Representing monotone operators by convex functions. In *Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988)*, volume 20 of *Proc. Centre Math. Anal. Austral. Nat. Univ.*, pages 59–65. Austral. Nat. Univ., Canberra, 1988.

[95] M. Flury. Large deviations and phase transition for random walks in random nonnegative potentials. *Stochastic Process. Appl.*, 117(5):596–612, 2007.

[96] J. Fröhlich, F. Martinelli, E. Scoppola, and T. Spencer. Constructive proof of localization in the Anderson tight binding model. *Comm. Math. Phys.*, 101(1):21–46, 1985.

[97] J. Fröhlich and T. Spencer. Absence of diffusion in the Anderson tight binding model for large disorder or low energy. *Comm. Math. Phys.*, 88(2):151–184, 1983.

[98] J. Gärtner. Convergence towards Burgers' equation and propagation of chaos for weakly asymmetric exclusion processes. *Stochastic Process. Appl.*, 27(2):233–260, 1988.

[99] V. Gayrard. Convergence of clock processes and aging in Metropolis dynamics of a truncated REM. *Ann. Henri Poincaré*, 17(3):537–614, 2016.

[100] V. Gayrard. Aging in Metropolis dynamics of the REM: a proof, preprint, arXiv:1602.06081.

[101] V. Gayrard and O. Gün. Aging in the GREM-like trap model, preprint, arXiv:1501.02986.

[102] N. Ghoussoub. *Self-dual partial differential systems and their variational principles*. Springer Monographs in Mathematics. Springer, New York, 2009.

[103] G. Giacomin. *Random polymer models*. Imperial College Press, London, 2007.

[104] G. Giacomin, J. L. Lebowitz, and E. Presutti. Deterministic and stochastic hydrodynamic equations arising from simple microscopic model systems. In *Stochastic partial differential equations: six perspectives*, volume 64 of *Math. Surveys Monogr.*, pages 107–152. Amer. Math. Soc., Providence, RI, 1999.

[105] G. Giacomin and F. L. Toninelli. Estimates on path delocalization for copolymers at selective interfaces. *Probab. Theory Related Fields*, 133(4):464–482, 2005.

[106] M. Giaquinta and L. Martinazzi. *An introduction to the regularity theory for elliptic systems, harmonic maps and minimal graphs*, volume 11 of *Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)]*. Edizioni della Normale, Pisa, second edition, 2012.

[107] J. Glimm. Boson fields with the $:\Phi^4:$ interaction in three dimensions. *Comm. Math. Phys.*, 10:1–47, 1968.

[108] J. Glimm and A. Jaffe. Positivity of the ϕ_3^4 Hamiltonian. *Fortschr. Physik*, 21:327–376, 1973.

[109] J. Glimm and A. Jaffe. *Quantum physics*. Springer-Verlag, New York, second edition, 1987.

[110] J. Glimm, A. Jaffe, and T. Spencer. Phase transitions for φ_2^4 quantum fields. *Comm. Math. Phys.*, 45(3):203–216, 1975.

[111] A. Gloria and D. Marahrens. Annealed estimates on the Green functions and uncertainty quantification. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, in press, arXiv:1409.0569.

[112] A. Gloria, S. Neukamm, and F. Otto. An optimal quantitative two-scale expansion in stochastic homogenization of discrete elliptic equations. *ESAIM Math. Model. Numer. Anal.*, 48(2):325–346, 2014.

[113] 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.

[114] A. Gloria, S. Neukamm, and F. Otto. A regularity theory for random elliptic operators, preprint, arXiv:1409.2678.

[115] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3):779–856, 2011.

[116] A. Gloria and F. Otto. An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, 22(1):1–28, 2012.

[117] A. Gloria and F. Otto. The corrector in stochastic homogenization: near-optimal rates with optimal stochastic integrability, preprint, arXiv:1510.08290.

[118] P. Gonçalves and M. Jara. Nonlinear fluctuations of weakly asymmetric interacting particle systems. *Arch. Ration. Mech. Anal.*, 212(2):597–644, 2014.

[119] Y. Gu. High order correctors and two-scale expansions in stochastic homogenization, preprint, arXiv:1601.07958.

[120] M. Gubinelli, P. Imkeller, and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi*, 3:e6, 75, 2015.

[121] M. Gubinelli and M. Jara. Regularization by noise and stochastic Burgers equations. *Stoch. Partial Differ. Equ. Anal. Comput.*, 1(2):325–350, 2013.

[122] M. Gubinelli and N. Perkowski. KPZ reloaded, preprint, arXiv:1508.03877.

[123] M. Gubinelli and N. Perkowski. Energy solutions of KPZ are unique, preprint, arXiv:1508.07764.

[124] M. Gubinelli and N. Perkowski. The Hairer-Quastel universality result in equilibrium, preprint, arXiv:1602.02428.

[125] M. Hairer. Solving the KPZ equation. *Ann. of Math. (2)*, 178:559–664, 2013.

[126] M. Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[127] M. Hairer and C. Labbé. A simple construction of the continuum parabolic Anderson model on $\mathbf{R}^2$. *Electron. Commun. Probab.*, 20:no. 43, 11, 2015.

[128] M. Hairer and C. Labbé. Multiplicative stochastic heat equations on the whole space, preprint, arXiv:1504.07162.

[129] M. Hairer and K. Matetski. Discretisations of rough stochastic PDEs, preprint, arXiv:1511.06937.

[130] M. Hairer and J. Quastel. A class of growth models rescaling to KPZ, preprint, arXiv:1512.07845.

[131] M. Hairer and H. Shen. A central limit theorem for the KPZ equation, preprint, arXiv:1507.01237.

[132] M. Hairer and W. Xu. Large scale behaviour of 3d phase coexistence models, preprint, arXiv:1601.05138.

[133] Q. Han and F. Lin. *Elliptic partial differential equations*, volume 1 of *Courant Lecture Notes in Mathematics*. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1997.

[134] B. Helffer and J. Sjöstrand. On the correlation for Kac-like models in the convex case. *J. Statist. Phys.*, 74(1-2):349–409, 1994.

[135] B. D. Hughes. Conduction and diffusion in percolating systems. In *Encyclopedia of complexity and systems science*, pages 1395–1424. Springer, 2009.

[136] A. Jaffe. Quantum fields, stochastic pde, and reflection positivity, preprint, arXiv:1411.2964.

[137] K. Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.*, 209(2):437–476, 2000.

[138] G. Jona-Lasinio and P. K. Mitter. On the stochastic quantization of field theory. *Comm. Math. Phys.*, 101(3):409–436, 1985.

[139] I. Junier and J. Kurchan. Microscopic realizations of the trap model. *J. Phys. A: Math. Gen.*, 37(13):3945, 2004.

[140] M. Kardar, G. Parisi, and Y.-C. Zhang. Dynamic scaling of growing interfaces. *Phys. Rev. Lett.*, 56(9):889, 1986.

[141] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(4):425–487, 1986.

[142] H. Kesten. Scaling relations for 2D-percolation. *Comm. Math. Phys.*, 109(1):109–156, 1987.

[143] H. Kesten. On the speed of convergence in first-passage percolation. *Ann. Appl. Probab.*, 3(2):296–338, 1993.

[144] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.

[145] F. Klopp. Weak disorder localization and Lifshitz tails. *Comm. Math. Phys.*, 232(1):125–155, 2002.

[146] E. Kosygina, T. S. Mountford, and M. P. W. Zerner. Lyapunov exponents of Green's functions for random potentials tending to zero. *Probab. Theory Related Fields*, 150(1-2):43–59, 2011.

[147] S. M. Kozlov. Averaging of differential operators with almost periodic rapidly oscillating coefficients. *Mat. Sb. (N.S.)*, 107(149)(2):199–217, 317, 1978.

[148] A. Kupiainen. Renormalization group and stochastic PDEs. *Ann. Henri Poincaré*, 17(3):497–535, 2016.

[149] T. M. Liggett. L_2 rates of convergence for attractive reversible nearest particle systems: the critical case. *Ann. Probab.*, 19(3):935–959, 1991.

[150] T. M. Liggett. Multiple transition points for the contact process on the binary tree. *Ann. Probab.*, 24(4):1675–1710, 1996.

[151] T. M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[152] L. Lundgren, P. Svedlindh, P. Nordblad, and O. Beckman. Dynamics of the relaxation-time spectrum in a cumn spin-glass. *Phys. Rev. Lett.*, 51(10):911, 1983.

[153] D. Marahrens and F. Otto. Annealed estimates on the Green function. *Probab. Theory Related Fields*, 163(3-4):527–573, 2015.

[154] P. Marcellini. Periodic solutions and homogenization of nonlinear variational problems. *Ann. Mat. Pura Appl. (4)*, 117:139–152, 1978.

[155] S. Martineau and V. Tassion. Locality of percolation for abelian cayley graphs, preprint, arXiv:1312.1946.

[156] P. Mathieu. Quenched invariance principles for random walks with random conductances. *J. Stat. Phys.*, 130(5):1025–1046, 2008.

[157] P. Mathieu and A. Piatnitski. Quenched invariance principles for random walks on percolation clusters. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 463(2085):2287–2307, 2007.

[158] P. Mathieu and E. Remy. Isoperimetry and heat kernel decay on percolation clusters. *Ann. Probab.*, 32(1A):100–128, 2004.

[159] F. Merkl and S. W. W. Rolles. Recurrence of edge-reinforced random walk on a two-dimensional graph. *Ann. Probab.*, 37(5):1679–1714, 2009.

[160] N. G. Meyers. An L^pe-estimate for the gradient of solutions of second order elliptic divergence equations. *Ann. Scuola Norm. Sup. Pisa (3)*, 17:189–206, 1963.

[161] R. Mikulevicius and B. L. Rozovskii. Martingale problems for stochastic PDE's. In *Stochastic partial differential equations: six perspectives*, volume 64 of *Math. Surveys Monogr.*, pages 243–325. Amer. Math. Soc., Providence, RI, 1999.

[162] T. S. Mountford. A metastable result for the finite multidimensional contact process. *Canad. Math. Bull.*, 36(2):216–226, 1993.

[163] T. S. Mountford. Existence of a constant for finite system extinction. *J. Statist. Phys.*, 96(5-6):1331–1341, 1999.

[164] C. Mueller. On the support of solutions to the heat equation with noise. *Stochastics Stochastics Rep.*, 37(4):225–245, 1991.

[165] A. Naddaf and T. Spencer. On homogenization and scaling limit of some gradient perturbations of a massless free field. *Comm. Math. Phys.*, 183(1):55–84, 1997.

[166] A. Naddaf and T. Spencer. Estimates on the variance of some homogenization problems, 1998, unpublished preprint.

[167] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80:931–954, 1958.

[168] E. Nelson. A quartic interaction in two dimensions. In *Mathematical Theory of Elementary Particles (Proc. Conf., Dedham, Mass., 1965)*, pages 69–73. M.I.T. Press, Cambridge, Mass., 1966.

[169] D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[170] H. Osada. Homogenization of diffusion processes with random stationary coefficients. In *Probability theory and mathematical statistics (Tbilisi, 1982)*, volume 1021 of *Lecture Notes in Math.*, pages 507–517. Springer, Berlin, 1983.

[171] K. Osterwalder and R. Schrader. Axioms for Euclidean Green's functions. *Comm. Math. Phys.*, 31:83–112, 1973.

[172] K. Osterwalder and R. Schrader. Axioms for Euclidean Green's functions. II. *Comm. Math. Phys.*, 42:281–305, 1975. With an appendix by Stephen Summers.

[173] G. C. Papanicolaou and S. R. S. Varadhan. Boundary value problems with rapidly oscillating random coefficients. In *Random fields, Vol. I, II (Esztergom, 1979)*, volume 27 of *Colloq. Math. Soc. János Bolyai*, pages 835–873. North-Holland, Amsterdam, 1981.

[174] G. Parisi and Y. S. Wu. Perturbation theory without gauge fixing. *Sci. Sinica*, 24(4):483–496, 1981.

[175] R. Pemantle. The contact process on trees. *Ann. Probab.*, 20(4):2089–2116, 1992.

[176] M. Prähofer and H. Spohn. Scale invariance of the PNG droplet and the Airy process. *J. Statist. Phys.*, 108(5-6):1071–1106, 2002.

[177] M. Röckner and T. S. Zhang. Uniqueness of generalized Schrödinger operators and applications. *J. Funct. Anal.*, 105(1):187–231, 1992.

[178] C. Sabot, P. Tarrès, and X. Zeng. The vertex reinforced jump process and a random Schrödinger operator on finite graphs, preprint, arXiv:1507.04660.

[179] C. Sabot and X. Zeng. A random Schrödinger operator associated with the vertex reinforced jump process on infinite graphs, preprint, arXiv:1507.07944.

[180] M. Sasaki and K. Nemoto. Analysis on aging in the generalized random energy model. *J. Phys. Soc. Japan*, 69(9):3045–3050, 2000.

[181] M. Sasaki and K. Nemoto. Numerical study of aging in the generalized random energy model. *J. Phys. Soc. Japan*, 70(4):1099–1104, 2001.

[182] T. Sasamoto and H. Spohn. One-dimensional kardar-parisi-zhang equation: an exact solution and its universality. *Phys. Rev. Let.*, 104(23):230602, 2010.

[183] B. Schapira and D. Valesin. Extinction time for the contact process on general graphs, preprint, arXiv:1509.04133.

[184] H. Shen and W. Xu. Weak universality of dynamical Φ^4_3: non-Gaussian noise, preprint, arXiv:1601.05724.

[185] V. Sidoravicius and A.-S. Sznitman. Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Related Fields*, 129(2):219–244, 2004.

[186] J. Sjöstrand. Correlation asymptotics and Witten Laplacians. *Algebra i Analiz*, 8(1):160–191, 1996.

[187] S. Smirnov. Conformal invariance in random cluster models. I. Holomorphic fermions in the Ising model. *Ann. of Math. (2)*, 172(2):1435–1467, 2010.

[188] H. Song, K.-N. Xiang, and S.-C.-H. Zhu. Locality of percolation critical probabilities: uniformly nonamenable case, preprint, arXiv:1410.2453.

[189] A. Stacey. The contact process on finite homogeneous trees. *Probab. Theory Related Fields*, 121(4):551–576, 2001.

[190] A. M. Stacey. The existence of an intermediate phase for the contact process on trees. *Ann. Probab.*, 24(4):1711–1726, 1996.

[191] L. Struik. Physical aging in plastics and other glassy materials. *Polymer Engineering & Science*, 17(3):165–173, 1977.

[192] J. W. Strutt, 3d Baron Rayleigh. On the influence of obstacles arranged in rectangular order upon the properties of a medium. *Philos. mag.*, 34(211):481–502, 1892.

[193] A.-S. Sznitman. Shape theorem, Lyapounov exponents, and large deviations for Brownian motion in a Poissonian potential. *Comm. Pure Appl. Math.*, 47(12):1655–1688, 1994.

[194] A.-S. Sznitman. *Brownian motion, obstacles and random media*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[195] C. A. Tracy and H. Widom. Integral formulas for the asymmetric simple exclusion process. *Comm. Math. Phys.*, 279(3):815–844, 2008.

[196] H. von Dreifus and A. Klein. A new proof of localization in the Anderson tight binding model. *Comm. Math. Phys.*, 124(2):285–299, 1989.

[197] W.-M. Wang. Localization and universality of Poisson statistics for the multidimensional Anderson model at weak disorder. *Invent. Math.*, 146(2):365–398, 2001.

[198] W.-M. Wang. Mean field bounds on Lyapunov exponents in $\mathbf{Z}^d$ at the critical energy. *Probab. Theory Related Fields*, 119(4):453–474, 2001.

[199] W.-M. Wang. Mean field upper and lower bounds on Lyapunov exponents. *Amer. J. Math.*, 124(5):851–878, 2002.

[200] V. V. Yurinskiĭ. On a Dirichlet problem with random coefficients. In *Stochastic differential systems (Proc. IFIP-WG 7/1 Working Conf., Vilnius, 1978)*, volume 25 of *Lecture Notes in Control and Information Sci.*, pages 344–353. Springer, Berlin-New York, 1980.

[201] V. V. Yurinskiĭ. Averaging of symmetric diffusion in a random medium. *Sibirsk. Mat. Zh.*, 27(4):167–180, 215, 1986.

[202] M. P. W. Zerner. Directional decay of the Green's function for a random nonnegative potential on $\mathbf{Z}^d$. *Ann. Appl. Probab.*, 8(1):246–280, 1998.

[203] R. Zhu and X. Zhu. Lattice approximation to the dynamical Φ^4_3 model, preprint, arXiv:1508.05613.