ABSTRACT. It is shown that in a large class of disordered systems with non-degenerate disorder, in presence of non-local interactions, the Integrated Density of States (IDS) is at least Hölder continuous in one dimension and universally infinitely differentiable in higher dimensions. This result applies also to the IDS in any finite volume subject to the random potential induced by an ambient, infinitely extended disordered media. Dimension one is critical: in the Bernoulli case, within the class of exponential interactions, the IDS measure undergoes continuity phase transitions, from absolutely continuous to singular continuous behaviour (the singularity in the latter case was known before). The continuity transitions do not occur for sub-exponential or slower decaying interactions, nor for $d \geq 2$. Technically, the case of polynomial decay is the simplest one.

The proposed approach provides a complement to the classical Wegner estimate which says, essentially, that the IDS in the short-range models is at least as regular as the marginal distribution of the disorder. In the models with non-local interaction the IDS is actually much more regular than the underlying disorder, which can even be discrete, due to the smoothing effect of multiple convolutions. In turn, smoothness of the IDS is responsible for a mechanism complementing the usual Lifshitz tails phenomenon.

It is also shown that the disorder can take various forms (e.g., substitution or random displacements) and need not be stochastically stationary (as in Delone–Anderson or trimmed/crooked Hamiltonians, for example); this does not affect the main phenomena observed already in the simplest setting.

Contrary to the situation with the usual lattice Bernoulli–Anderson Hamiltonians, the proof of Anderson localization in the models with infinite-range interaction follows in a fairly simple way from the main bounds on the finite-volume IDS. Another distinction from the approach developed by Bourgain and Kenig for the continuous Bernoulli–Anderson Hamiltonians, and later extended by Germinet and Klein to arbitrary (locally IID) disorder, is that all nontrivial marginal distributions are treated in a unified way, via harmonic analysis and without reduction to an embedded Bernoulli model, thus keeping potential benefits from less singular forms of underlying disorder.

Long-range models have an amazingly large number of connections to several classical problems of harmonic analysis, probability theory, dynamical systems and number theory.

CONTENTS

1. Introduction 3
2. Viète–Euler identity and smoothness of IDS: there and back again 8
2.1. Introductory remarks. Notation. Alloy transform 8
2.2. Viéte–Euler identity, Bernoulli alloys, and dynamical systems

3. Main results
3.1. Exponential and sub-exponential screening
3.2. Power-law screening
3.3. Localization under exponential screening
3.4. Localization under sub-exponential screening
3.5. Localization under power-law screening

4. Characteristic functions of sums of random variables

5. Fourier analysis I. Sub-exponential or exponential decay
5.1. Thermal bath estimates. Smoothness
5.2. Hölder continuity via integral estimates
5.3. "Frozen bath" estimate
5.4. Back to probability measures

6. Fourier analysis II. Summable polynomial decay
6.1. Thermal bath estimate
6.2. "Frozen bath" estimate
6.3. Back to probability measures

7. Wegner-type estimates

8. ILS estimates
8.1. Localization at the band edges
8.2. Large disorder
8.3. Dilute alloys

9. Localization: exponentially decaying potentials
9.1. Deterministic analysis of Green functions
9.2. Scaling of probabilities
9.3. Exponential scaling limit

10. Localization: polynomial potential
10.1. Deterministic analysis
10.2. Probabilistic analysis

11. Two-point correlation measures of the cumulative potential

12. Concluding remarks
12.1. On the proof of Theorem 3.4
12.2. Random dipoles and random displacements
12.3. Non-homogeneous models of disorder
12.4. Random magnetic fields
12.5. Correlated sources and statistical mechanics
12.6. On higher smoothness in the "frozen bath" approximations
12.7. Thin tails from neighbors

References
1. Introduction

Since the discovery of the phenomenon of quantum localization by Philip W. Anderson [1], a certain number of “simplifying” assumptions were made both in physical and mathematical models of disordered media. Probably, the most important of all are those concerning the nature of interactions between the quantum objects involved. Specifically, one has to distinguish two kinds of interactions:

- between the mobile objects (e.g., charged particles), and the “external” sources of disorder (e.g., heavy ions);
- between the above mentioned mobile objects themselves.

The latter interactions have been the main subject of recent physical and mathematical works over the last decade; cf. e.g. [66, 8, 27, 28, 29, 4, 24, 30, 50, 26, 22]. Importance of this new problematic initiated, on the mathematical level, in 2003 at the Isaac Newton Institute (Cambridge University, UK) has been recognized at the XVIth Congress on Mathematical Physics (2010) and reported by Aizenman and Warzel [5]. Speaking of rigorous aspects of the problem, a number of questions in this new area of spectral theory of random operators still remain wide-open and challenging; they will not be discussed in the present paper.

The main topic of this work is the impact of the non-local physical nature of interactions between the mobile quantum objects and an ambient disordered classical environment, on the qualitative characteristics of the Integrated Density of States (IDS) and, where applicable, the Density of States (DoS). The existence of the DoS (and more generally, regularity of the IDS) is usually derived in higher-dimensional models from local regularity of the marginal distribution of disorder, via Wegner’s estimate and its generalizations. Nothing can ever prevent mathematicians from assuming anything, yet it is quite natural to ask: “Where does a regular disorder come from?” A short answer to this question, which was at the origin of the present work, is that it is simply hard to avoid, for it emerges in a fairly universal way from virtually any kind of disorder more or less evenly distributed in the configuration space. The main mechanism is the smoothing effect of multiple convolutions, and the principal, very convenient mathematical tool for analyzing this effect is harmonic analysis of probability measures.

The one-dimensional systems are set apart in this respect, since the regularity of the IDS in the case of strongly singular local disorder (e.g. 1D Bernoulli) is known to follow by Hilbert transform from that of the Lyapunov exponents, but deep inside, one finds the same regularizing effect of multiple convolutions. The Riccati dynamics for the so-called Prüfer phase is of course nonlinear, but the Pastur–Figotin argument [104] shows that in the particular case of weak disorder the linearized dynamics alone leads to an asymptotically exact formula for the positive Lyapunov exponents, and even in the correlated case [25] the asymptotic behavior can be derived, in fact, from the linear harmonic analysis. All this can be done for the local models of disorder: the Lyapunov solutions, from which the Green functions (GFs) are built, provide in 1D convenient “test functions” accumulating the effect of multiple local fluctuations (be those linearized or not). Revealing a similar mechanism in higher dimension, with local (e.g., lattice IID) disorder seems much
harder a problem, but the situation changes radically as soon as we turn to physics and recall ourselves that fundamental interactions are NOT local. In particular, the Coulomb interaction between charged particles has infinite range.

Sometimes I refer to the mobile particles as electrons, but the detailed discussion of the real physical processes, especially on the level of second quantization, is most certainly beyond the scope of this paper, and occasional use of physical terminology is intended for terminological brevity only. The reality is of course much richer. Depending upon a specific physical model, the mobile objects may carry charge and spin either together or separately.

Speaking of electrically charged particles subject to electrostatic interactions (which can of course be complemented by magnetic fields), the fundamental Coulomb interaction is extremely slowly decaying, but in a large sample of a heterogeneous media, composed of a huge number of more or less mobile charges, it actually manifests itself only in a dampened, ”screened” form. The screening effects in solid state media have been since several decades an inalienable part of any physical work realistically describing quantum many-body systems. On the other hand, a vast majority of mathematical papers on Anderson localization operate with local models of disorder, starting with the pioneering papers on localization in one dimension (Goldsheid, Molchanov and Pastur [65] in \( \mathbb{R}^1 \); Kunz and Souillard [93] in \( \mathbb{Z}^1 \)) and in higher dimension (Fröhlich and Spencer [58]: exponential decay of Green functions; Fröhlich, Martinelli, Scoppola and Spencer [57]: pure point spectrum with exponentially decaying eigenfunctions).

In a few exceptions, the extended, non-local nature of the potential has been in prior works more of a nuisance, or perhaps an additional technical challenge, overcoming which would warrant certain sacrifices in the strength of the localization results one aimed to obtain. In fact, the first rather general result on correlated Gaussian potentials was obtained by von Dreifus and Klein [40] shortly after their reformulation of the energy-interval, or variable-energy, MSA (VEMSA) in the frequently cited paper [39]. Later on, Kirsch et al. [84] considered more general (non-Gaussian) marginal distributions. See also recent works [92, 81, 131, 132] and references therein. It seems only fair to explore the true role of non-local interactions (apparently, the only ones known in physics of solid state) in a broader context. This is precisely the main goal of the project the first part of which is presented in this paper. We shall see that the infinite range of interaction is indeed much deeper a subject than a technical mathematical nuisance.

We argue that a traditional reduction of the environment of a finite volume to the PDE-type boundary conditions hides a significant part of the story, and that the picture becomes substantially more complete in the traditional setting of statistical mechanics, where the environment acts as a thermal bath. It proves fairly instructive to decompose the integrated density of states into two components which, for the lack of better words and following a mechanical metaphor, we call respectively “tidal” and “ripple” components. In the daily movements of the ocean’s level on a sloping shore, local perturbations (wind, irregularities of the beach) determine a perceptible profile of the water surface, but the principal movements themselves are the result of incommensurably weaker fluctuations of gravitational forces from extremely remote sources; forces which would be unable to
move water, say, by a meter or two in vertical direction, in a strictly isolated container of the size comparable to the beach, e.g., in a lake, so they have to act through a much larger external volume. Imperceptible per se, those gravitational forces produce an easily perceptible by eye movement of water, back and forth, on an almost horizontal, yet sloping beach. Weak or not, it is the tidal mechanism which determines a considerable periodic evolution of the coastal area, that a local wind could not produce.

This is more than just a qualitative metaphor: the gravitational potential, like Coulomb, is a slowly decaying function in the sense that its gradient decays faster than the function itself, so distant sources produce "almost flat" (yet nowhere flat) fluctuations. We discuss this aspect briefly in Section 11 dedicated to the analysis of regularity of the two-point correlation measures, but it may have important implications for the many-body localization phenomena, in the light of some technical issues raised in [22].

This text is quite long and probably not easy to read; perhaps it is worthwhile to single out two techniques most useful from a pragmatic point of view. First, for a relatively simple proof of Hölder-type Wegner estimates, the integral bounds based on the techniques due to Wiener and Wintner [136] (closely related to [135]) are quite useful. In a broader context, these techniques have been used and further developed by Strichartz [127, 128]). Secondly, in the most relevant models with weak (polynomial) screening, Wintner’s technique [138] is both very simple and efficient for the proof of smoothness estimates.

Bourgain and Kenig [13] proved a remarkable eigenvalue concentration (EVC) estimate for Bernoulli–Anderson Hamiltonians in $\mathbb{R}^d$, which was later extended to arbitrary nontrivial disorder by joint efforts of Aizenman, Germinet, Klein and Warzel [2]. Their approach was based on a combinatorial argument (the Sperner lemma), but a reflex to saying "$\sqrt{N}$" among probabilists would certainly be even more Pavlovian than to saying "Jingle ..." in a preschool, on some 23rd December. In a way, the analysis given below ("thermal bath estimates") justifies that reflex. Observations made here evidence that the interactions of infinite range (the only physically relevant ones, anyway) actually provide a music easy to sing to, especially with the help of harmonic analysis. Singing the same lyrics but a capella remains an intriguing mathematical challenge, regardless of any physical applications.

We thus come to a more quantitative discussion of a model of the forces originating far away from a locus where their effects are to be studied.

The strongest form of screening occurs in 3D systems when charged particles are highly mobile, e.g., in plasma; the Debye–screened Coulomb potential originating from a given local source decays exponentially fast at large distances $r$, which have to be larger than some characteristic length, so as to enable several layers of induced waves of concentration of positive and negative charges to be created around the aforementioned remote source. The simplest approach relies on the classical statistical physics. Even in this case, screened Coulomb potential in dimension $d \leq 2$ is slowly decaying.
It was realized by physicists that the classical approximation results in an oversimplified and even qualitatively inaccurate picture of screening, particularly in solid state media. In more accurate models, the correction terms are no longer obtained by “commutative” probabilistic analysis but require a quantum description, the choice between Fermi–Dirac and Bose–Einstein quantum statistics, and complex diagrammatic techniques. Also, a quantum charged particle is not a point charge, and linear approximation to the Gibbs distribution is only an approximation. More importantly, one has to consider a full-fledged quantum many-body problem to achieve a good agreement with experiment. As a result, one has not a universal behaviour, but various forms of screening. The response of a large sample to a single source causes the so-called Friedel oscillations (cf. [55, 56, 94, 87]), observable experimentally, and the quantitative parameters, first of all the decay rate of the screened electrostatic potential from a given source, strongly depend upon the shape of the Fermi surface of the mobile particles responsible for the screening.

With these observations in mind, we shall explore various decay rates of the effective (screened) potential produced by heavy “ions” forming a spatial grid, periodic or not; these will range from the strongest (exponential) to the slowest power-law ones, just barely summable.

Below I am going to focus mainly on media of spatial dimension strictly higher than one, for two principal reasons.

- Firstly, from the perspective of applications to Anderson localization, the one-dimensional models are understood to a much greater extent than their higher-dimensional counterparts. The specifically 1D mechanisms, treated in terms of “back-scattering” in physical approaches, or with the help of products of random matrices, in the rigorous mathematical works, result in a much more complete and clear picture.

- Secondly, from the perspective of the continuity phase transition of the IDS which we are going to describe, it will become clear that for many intents and purposes, sufficiently “thick” quasi-one-dimensional media, i.e., those extended in one direction and having a finite cross-section, are much closer to higher-dimensional samples than to single-channel linear chains. More precisely, the dimensional threshold for the continuity phase transition – for a given exponential decay rate of the screened potential – is encountered already within the class of quasi-1D strips, for the cross-sections large enough. From this point of view, the macroscopic wires already have a cross-section very large in microscopic (atomic) units.

However, 1D systems are certainly worth a thorough investigation. Even a brief familiarization with physical literature, theoretical and experimental, is useful and can be recommended, to see that some mathematical issues, requiring in this paper a fair amount of space and efforts, are perhaps just that – mathematical ghosts from the land of might-have-been. Specifically, it seems logical to investigate the IDS continuity phase transition in low-dimensional media (viz. in 1D or in $1 < d \approx 1$), but our analysis evidences that the transition to singularity requires nothing less than exponential screening, and sufficiently strong one. The reader can see, e.g., in the works by Gabovich et al. [59] or by Petrashov et al. [105] how “strong” a 1D screening can be in physical reality...
At the same time, I would like to stress that the non-local tidal effects of disorder on the regularity of the DoS (or IDS) suggest that the dimensionality parameter has to be properly defined in models with heterogeneous structure, e.g., in the localization problem in a surface (or a specially designed internal) layer on a 3D substrate, or in a quasi-1D channel on the surface or inside a 2D/3D sample. While the quantum tunneling effects for the mobile agents may be limited to a linear sub-manifold of lower dimension (or a thin neighborhood thereof), the tidal DoS may or might be strongly influenced by the disorder in the ambient sample of higher dimensionality.

In models with a low concentration of “ions” creating a specific, gap-isolated energy band, this concentration itself may also become an important parameter near the critical point of the continuity phase transition: the decay rate of the screened potential is to be compared to the typical distance between nearest relevant loci.

More generally, the predictions concerning the continuity phase transition of the IDS are scale-dependent, as they result from a renormalization group (RG) type analysis, hence the effects become perceptible and sharp only beyond some minimal scale. In some mesoscopic systems, their size may or might be insufficient for the RG limit to give the right answer.

In physics, there is a number of characteristic lengths related to exponentially decaying functions: instead of $f(x) = e^{-a\|x\|}$ with $a > 0$, it is customary to write $f(x) = e^{-\|x\|/\xi}$, where $\xi$ describes the distance such that “for $\|x\| \gg \xi$, $f(x)$ is essentially nothing”. This may be true in many realistic situations, but the key equations of this work, (2.4) and (2.5), provide an instructive example of what can be the difference between an infinite series with uniformly bounded, exponentially decaying terms and any of its partial sums. Somehow, a measure supported by a finite number of atoms is “slightly” less regular than an a.c. measure; in higher dimensions, as we shall see, the latter even has a bad habit to become infinitely smooth.

**Disclaimer.** This paper focuses primarily on the fluctuations of the finite-volume IDS induced by a random media and on the regularity of their probability distribution, rather than on an exact form of the IDS. The main subject is therefore not the same as in many physical works.

The physical mechanisms of screening are not analyzed. The effects of a given interaction $u$ are studied regardless of whether or not it can occur in realistic models of a given dimensionality $d$, although the $d$-dependence of regularity properties of the cumulative potential and of the IDS is studied. The main goal is to find out how the most regular forms of disorder can emerge from the most singular ones, under the most difficult conditions.

The list of bibliographical references, although it is rather long, is quite possibly incomplete, despite my best efforts. In three words, an explanation but not an excuse, is: paid online access.
2. Viéte–Euler identity and smoothness of IDS: there and back again

2.1. Introductory remarks. Notation. Alloy transform. In presence of non-local interactions, one has to distinguish two kinds of potentials:

- the ”source” potential, described by the amplitudes at the origin points;
- the ”target”, cumulative potential registered at each point of the space.

We do not discuss the self-consistent, many-body models, so the basic disorder comes from the immobile sources, the configuration of which is to be determined in the framework of statistical physics; we assume the sources to be stochastically independent. One possible model is the so-called alloy potential (displacements models will be briefly discussed; they can be treated in essentially the same way),

\[ x \mapsto \sum_{y \in \mathcal{Z}} q_y u(x - y), \]

where \( \mathcal{Z} \) is a countable subset of the configuration space \( \mathcal{X} \), e.g., of \( \mathcal{X} = \mathbb{Z}^d, \mathbb{R}^d \). For definiteness, we will consider \( \mathcal{X} = \mathbb{Z}^d \). The registered cumulative potential, as a function on \( \mathcal{X} \), is defined through a linear mapping, which can be called alloy transform,

\[ U : \mathbb{Q} \mapsto U[\mathbb{Q}] = V, \quad V : \mathbb{Z}^d \to \mathbb{R}, \quad (2.1) \]

where

\[ V(x) = (U[\mathbb{Q}])(x) = \sum_{y \in \mathbb{Z}^d} u(y - x)q_y. \quad (2.2) \]

The interaction potential \( u(\cdot) \) will always be assumed absolutely summable,

\[ \sum_{x \in \mathbb{Z}^d} |u(x)| \leq C < \infty, \]

and nonnegative\(^1\), chosen from the class of power-law or (sub-)exponential functions, for we shall need lower bounds on the decay of \( u(\cdot) \), too. These notations will be used in the context of deterministic estimates and statements, to avoid confusion with probabilistic arguments where \( q \) are replaced by random variables, assumed IID in this paper, forming a random field on \( \mathbb{Z}^d \) relative to a probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \). To keep parallels with the deterministic setup, we denote \( \omega = \{\omega_x, x \in \mathbb{Z}^d\} \) samples of the random field of amplitudes \( \omega_x \); the potential registered at a site \( x \) has the same linear-algebraic form \( V(x; \omega) = (U[\omega])(x) \), with \( \omega \) replacing \( q \) in \( (2.2) \).

Fixing a point \( x \in \mathcal{X} \), we come to the analysis of regularity of the probability distribution \( \nu_x \) of \( V(x, \omega) \). Assuming that \( \omega_x \) are IID with common probability measure \( \mu \), \( \nu_x \) is the image of a transform of \( \mu \) parameterized by

1. a countable subset \( \mathcal{Z} \subset \mathcal{X} \),
2. a function \( u : \mathbb{R}_+ \to \mathbb{R} \),

\(^1\)In physical models, correlations can be sign-indefinite. We usually deal with absolute amplitudes, but even these can be somewhere closer to 0 than in average. It will be clear from our analysis that exclusion of some radii is harmless for the main phenomena. A more detailed analysis will be carried out in a forthcoming work.
which provides an interesting generalization, and not just an abstract one, of the theory of random series, closely related to the theory of self-similar measures. Considering all \( x \in \mathcal{X} \) at once, or in a bounded domain, we encounter an even more intriguing problem for the \((Z, u)\)-parameterized transform of a measure \( \mu \) into a random field \( V(x, \omega) \) on \( \mathcal{X} \). Discussion in Section 11 barely scratches the surface of the latter problematic.

As functions on the lattice, both \( q \) and \( \omega \) will be assumed uniformly bounded; in the case of \( q \) this is a non-ambiguous statement (in other words, \( q \) will be assumed elements of \( \ell^\infty(\mathbb{Z}^d) \), and even of a finite ball at the origin thereof), while \( \omega \) requires a bit of formalities: the uniform boundedness is to be assumed with probability one. Alternatively, we can simply define \( \Omega = [0, 1]^{\mathbb{Z}^d} \).

With these remarks, \( U \) is well-defined on all admissible \( q \) or \( \omega \), considered as elements of \( \ell^\infty(\mathbb{Z}^d) \). In Sections 4–6 we will have to control the dependence of the image \( U(q) \) (or, respectively, \( U(\omega) \)) upon the values \( q_x \) (resp., \( \omega_x \)) inside and outside some finite balls \( B \) in \( \mathbb{Z}^d \). To this end, we canonically inject \( \ell^\infty(B, \ell^\infty(B^c)) \rightarrow \ell^\infty(\mathbb{Z}^d) \) by zero-extensions, and note that for \( q = q_B + q_{B^c} \), where \( q_B \cdot q_{B^c} \equiv 0 \) as function on \( \mathbb{Z}^d \), one has \( U(q) = U(q_B) + U(q_{B^c}) \), but of course there is no reason in general for \( U(q_B) \cdot U(q_{B^c}) = 0 \). Indeed, assuming for example that \( u \) is strictly positive everywhere, one has \( U[\delta_x] \) also strictly positive everywhere, with \( \delta_x \) being the lattice delta-function at an arbitrary point \( x \).

The infinite range of the single-point potentials (scatterers) is certainly a double-edged sword, as the reader will see on a number of occasions. However, one thorny problem of rigorous Anderson localization – incomplete covering and an inevitable recurs to some form of the Unique Continuation Principle, alas, unavailable in general discrete models – simply has no raison d’etre in presence of realistic, long-range scatterers.

The reverse of the medal starts with the non-local dependencies between the events referring to localization (insufficient/no localization) in distant finite domains, or proximity of local spectra in distant domains, possibly leading to a long-distance tunneling.

The former issue had been addressed long ago by Kirsch et al. [84] who proposed one possible way around this problem; we adapt it to our problem in Sections 9–10. As to the latter, this is one of the instances where the infinite range of interaction proves salutary, and transforms even the most singular nontrivial disorder distribution into a highly regular one, thanks to multiple convolutions.

### 2.2. Viète–Euler identity, Bernoulli alloys, and dynamical systems.

In the sixteenth century\(^2\) François Viète\(^3\) discovered a remarkable identity

\[
\frac{2}{\pi} = \prod_{k=1}^{\infty} \cos \left( \frac{\pi}{2k} \right)
\]

(2.3)

which was generalized two centuries later by Leonhard Euler:

\[
\frac{\sin x}{x} = \prod_{k=1}^{\infty} \cos \left( \frac{x}{2k} \right);
\]

(2.4)

\(^2\)According to different sources, in 1579 or in 1593.

\(^3\)François Viète, or François Viette, or Franciscus Vieta (1540–1603). His last mathematical work “Opera mathematica ...” had remained unfinished, and was published only in 1646 by Frans van Schooten.
the latter follows by simple arguments from $\sin x = 2\sin \left(\frac{x}{2}\right) \cos \left(\frac{x}{2}\right)$. It is the opening topic of Mark Kac’ book [77]. The Viète–Euler identity provided long ago a bridge between two areas of mathematical analysis to emerge much later: harmonic analysis and fractal measures closely related to dynamical systems. Moreover, it makes unnecessary a delicate analysis of a critical model we are going to discuss a bit later.

Curiously, another remarkable elementary identity, figuring in Mark Kac’ book as Problem 5 (Chapter 1), is closely related to one of the cornerstones of a simple and very efficient smoothing technique used in an uncountable number of works on asymptotic formulae for the probability distribution functions and/or densities of normalized sums of independent random variables (identical or not); a topic we shall also come across in the discussion of “smoothness” of finite-volume IDS (see Section 12). The asymptotic expansions obtained in this way are widely used in statistics in general and in risk management, in particular.

As we shall see, answers to some tough questions appearing in spectral analysis of random Hamiltonians with long-range interactions can be found in the harmonic analysis of probability measures, and the specific form of some questions may bring new motivations to this classical area of mathematical analysis.

A prototypical form of the main mechanism we are going to exploit can be seen from the usual dyadic expansion of a real number

$$[0, 1] \ni \omega = \sum_{k=1}^{\infty} \frac{\omega_k}{2^k} \quad (2.5)$$

establishing an “almost” bijective isomorphism between $[0, 1]$ equipped with the Lebesgue measure and the set of infinite words $(\omega_1, \omega_2, \ldots) \in \Omega = \{0, 1\}^{\mathbb{N}}$ endowed with the structure of a measure space $(\Omega, \mathcal{B}\Omega)$ ($2^\Omega$ is the cylinder sigma-algebra rendering measurable all projections $\omega \mapsto \omega_i$), with the product measure $\mathbb{P}\Omega$ characterized by

$$\forall n \geq 1 \quad \mathbb{P}\Omega \{ \omega : \omega_n = 0 \} = \mathbb{P}\Omega \{ \omega : \omega_n = 1 \} = \frac{1}{2}.$$

The RHS of (2.5) can be interpreted as an alloy-type potential on $\mathbb{N}$ with exponentially decaying scatterer potential $u : r \mapsto 2^{-(r+1)}$, symmetric $(\frac{1}{2}, \frac{1}{2})$ Bernoulli distribution of the scatterers amplitudes, and evaluated at the origin. The above mentioned isomorphism transforms therefore the most singular nontrivial local disorder distribution into a perfect Lipschitz continuous one, with compactly supported density bounded by 1. The LHS of the Viète–Euler identity (2.4) is the characteristic function (= inverse Fourier transform) of the probability distribution of $\omega$ (i.e., of the Lebesgue measure on $[0, 1]$), while its RHS expresses it as the characteristic function of the sum of independent r.v. related to $2^{-k} \omega_k$ from (2.5) by a simple affine transformation: $\tilde{\omega}_k = 2\omega_k - 1 \in \{-1, +1\}$. A more symmetrical alloy model, on the entire lattice $\mathbb{Z}^1$, produces by independence a convolution of two uniform distributions, resulting in an even better – globally continuous – compactly supported density.

The relation between the admissible values of the individual amplitudes, 0 and 1 (or rather the distance between them) and the precise decay exponent of the function $u : k \mapsto 2^{-k}$, is crucial for the regularity of the induced single-site measure. Taking 0 and $a$ with
a > 1 results in a Cantor set supporting the infinite convolution measure, for there are obviously gaps in the set of values of the sums \( \sum_{n} \omega_n 2^{-n} \). For example,

\[
0 \cdot \frac{1}{2} + \sum_{n \geq 2} \frac{a}{2^n} < a \cdot \frac{1}{2} + \sum_{n \geq 2} \frac{0}{2^n}.
\]

However, even in such a case the resulting measure is (singular\(^4\)) continuous, even Hölder continuous. Moreover, a well-known example (cf., e.g., [51, v.2, Section V.4(d)]) shows that the convolution of two singular Cantor measures can be a.c. (Lebesgue measure on an interval).

The problem of decay (and where appropriate, decay rate) at infinity of the Fourier transform/coefficients\(^5\) of a probability measure on \( \mathbb{R} \) (or on \([0,1]\)) has a long and rich history. It all starts in 1854 with Riemann’s proof of decay at infinity of the Fourier coefficients of any periodic Riemann-integrable function on \( \mathbb{R} \); Lebesgue extended this result to Lebesgue-integrable functions. A systematic study of Borel measures on the torus \( \mathbb{T} = \mathbb{R}/\mathbb{Z} \cong [0,1) \) with decaying Fourier coefficients,

\[ \hat{\mu}(n) = \int_{\mathbb{T}} e^{-int} d\mu(t), \]

was carried out in 1920s by Rajchman [106, 107]; such measures have been called Rajchman measures; this class contains all a.c. measures.

Actually, Meneshov (”Menchoff” in the French-style transliteration used in many of his works) constructed in Ref. [100] an example of a singular Rajchman measure in 1916, precisely one century ago, although the term ”Rajchman measure” was not coined yet at that time.

Shortly after that (in 1918), Riesz introduced in [109] what is called today Riesz products,

\[ x \mapsto -x + \lim_{n \to \infty} \int_0^x \prod_{k=1}^n \left( 1 + \alpha_k \cos(m_k t) \right) dt, \]

with \( \alpha_k \in [-1,1] \). When \((m_k)_{k \geq 1}\) is a rapidly growing sequence of positive integers, the Fourier coefficients are not \( o \left( n^{-1} \right) \).

In 1920 Neder [101], answering a question raised by Riesz [109], proved that every Rajchman measure is continuous.

Ivašev–Musatov [74, 75] proved that the Fourier coefficients \( \hat{\mu}(n) \) of continuous measures \( \mu \) mutually singular with the Lebesgue measure are dominated by all functions of the form

\[ r(n) = \left( n \ln n \cdots \ln \cdots \ln n \right)^{-1/2}, \quad p \geq 1. \]

By Jessen–Wintner theorem [76, Theorem 11], an infinite product

\[
\varphi(t) = \prod_{k=1}^{\infty} \cos (r_k t), \quad (2.6)
\]

\(^4\)Of course, it is not the presence of gaps by itself which implies singularity; a Cantor set may have positive Lebesgue measure. Here the gaps are ”too big”, so the support has zero Lebesgue measure.

\(^5\)The term “coefficients” was actually introduced by François Viète.
giving the characteristic function of a random series \( S(\omega) = \sum_{k \geq 1} r_k X_k(\omega) \), with IID symmetric Bernoulli r.v. \( X_k(\omega) \in \{-1, +1\} \), is well-defined under the assumption \( \sum_k r_k^2 \equiv \sum_k E[(r_k X_k)^2] < \infty \) (cf. Kolmogorov’s three-series theorem [83], e.g. in [51, Section IX.9]), and in this case \( S(\cdot) \) has either purely s.c. or a.c. distribution.

Jessen and Wintner [76, Section 6] give an instructive set of examples of random series of scaled symmetric Bernoulli r.v. with the characteristic functions (2.6). In particular, Example 4 corresponds to the series \( \sum_{k \in \mathbb{Z}} 2^{-k} X_k \), with compactly supported bounded density \( \rho(x) = \left( \frac{1}{2} - \frac{1}{4} x^2 \right) 1_{[-2,2]}(x) \), and in Example 5 one has a series over \((\mathbb{N}^*)^2\),

\[
S^{(2)}(\omega) = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} 2^{-k-l} X_{k,l}(\omega),
\]

again with IID symmetric Bernoulli \( X_{k,l} \). The authors of [76] point out that

\[
\phi_{S^{(2)}}(t) = \prod_{k=1}^{\infty} \phi_{S^{(1)}}(2^{-k} t)
\]

with \( S^{(1)}(\omega) = \sum_{k=1}^{\infty} 2^{-l} X_l(\omega) \), thus the fast decay of \( \phi_{S^{(2)}}(t) \) at infinity implies that \( S^{(2)} \) has density \( \rho_{S^{(2)}} \in C^\infty(\mathbb{R}) \).

The case of polynomial decay did not escape their attention, either, although they consider (in Example 7) the situation where the series \( S(\omega) = \sum_{k=1}^{\infty} r_k X_k(\omega) \), \( r_k = k^{-4} \) converges in mean square but not absolutely, which gives rise to an unbounded r.v. with density \( \rho \in C^\infty(\mathbb{R}) \).

Wintner [137, 138] proposed a very natural and elementary upper bound for the characteristic functions, proving in the case of polynomially decaying \( k \mapsto r_k \) infinite derivability\(^6\) of the respective probability density. In fact, his technique from [138] alone suffices for a good half of main results of this paper, and applies to the most realistic physical models of disordered solid state media (with power-law screened interactions). On the other hand, for the reasons coming from the main application of this paper (to Anderson localization), I intentionally avoid below the discussion of the case where \( \sum_k r_k = +\infty \) (but \( \sum_k r_k^2 < +\infty \)) and the method from [138] is quite efficient.

Return to the characteristic function and write

\[
\ln |\phi(t)| = \sum_{k=1}^{\infty} \ln |\cos(b^{-k} t)|.
\]  

(2.7)

By parity, we can assume \( t > 0 \). Let

\[
K_t = m^{-1} \ln t, \quad m = \ln b,
\]

so that \( t b^{-K_t} = t e^{-m \ln t} = t^{1-mm^{-1}} = 1 \). For \( k > K_t \), \( \cos(b^{-k} t) \leq 1 - \frac{t^2}{4b^{2k}} \) thus

\[
\ln |\cos(b^{-k} t)|^{-1} \geq \ln \left( 1 - \frac{t^2}{4b^{2k}} \right)^{-1} \geq Cr^2 b^{-2k}.
\]

\(^6\)Jessen and Wintner [76, Section 6, Example 7] pointed out that for \( u(r) = 1/r \) the cumulative distribution has an analytic density in \( \mathbb{R} \) (actually, even in higher-dimensional convolution models), which is impossible whenever the series at hand converges to a bounded r.v., having necessarily a compactly supported probability distribution.
Decompose (2.7) into two sums:

\[
\ln |\varphi(t)|^{-1} = -\sum_{k=1}^{K_t} \ln |\cos (b^{-k} t)|^{-1} + \sum_{k>K_t} \ln |\cos (b^{-k} t)|^{-1}
\]

\[=: S_1(t) + S_2(t) \geq S_1(t) + C t^2 \sum_{k>K_t} b^{-2k} \geq S_1(t) + C t^2 b^{-2K_t}.
\]

Further,

\[S_2(t) \geq C t^2 b^{-2K_t} \geq C t^2 e^{-2m n^{-1} \ln t} = C',
\]

which provides no decay to $|\varphi(t)|$, so we turn to $S_1(t)$. Recalling $b^{-K_t} t = 1$,

\[S_1(t) = -\sum_{k=1}^{K_t} \ln |\cos (b^{-k} t)| = -\sum_{n=1}^{K_t} \ln |\cos (b^n \cdot b^{-K_t} t)|
\]

\[= -\sum_{n=1}^{K_t} \ln |\cos (b^n t)|. \tag{2.9}
\]

The last expression certainly calls for an ergodic theorem, namely the one for the fractional parts of $(\pi^{-1} b^n t)_{n \geq 1}$. Indeed, the threshold $K_t$ is chosen so as to ensure that for $k \geq K_t$ one has $b^{-k} \lesssim 1$, so in the reversed time scale, $K_t, K_t - 1, \ldots, 1$, we have a growing sequence of arguments $b^{-K_t} t \cdot b^k \sim C b^k$ of the periodic function $\cos$.

Unlike the "tidal" sum with a fixed $t$ and large $r$ leading to the Gaussian micro-scaling asymptotics, nothing precise can be said in general about any individual term in (2.9), but there are many of them; are they more or less evenly distributed or concentrated in the vicinity of $\pi Z$?

The equidistribution is often established with the help of Weyl’s criterion [134], applicable to a large variety of dynamical systems on tori. See also the works by Koksma [88], Dubickas [41], the monograph by Cornfeld, Fomin and Sinai [32] and further references provided therein.

Indeed, Kac, Salem and Zygmund [78] considered the equidistribution problem and noticed that (cf. [78, Section 5]) the expectation value

\[
\frac{1}{2\pi} \int_0^{2\pi} \ln |\cos \theta| \, d\theta = \ln 2
\]

suggests that for $b = 2$, by (2.8) with $m = \ln 2$,

\[|\varphi(t)| \lesssim e^{-K_t \ln 2} = e^{-\frac{\ln 2}{m} \ln |t|} = (\ln |t|)^{-\frac{\ln 2}{m}} = |t|.
\]

Of course, a simple integration does not suffice here (and neither was it used alone to infer rigorous consequences in [78]), for we deal with the logarithm of $|\varphi(t)|$, so the fluctuations cannot be taken lightly in the equidistribution argument. Quite fortunately, Viéte and Euler had solved for us the critical model a few centuries ago. Had they not,

\textsuperscript{7}In a forthcoming paper, the equidistribution mechanism will receive a proper treatment, based on a great wealth of results accumulated in this area. Cf. e.g. [68, 41], a more recent monograph [32] and references therein. The fractional parts of $(q^n t)_{n \geq 1}$ are equidistributed for a.e. $q > 1$ or a.e. $t \neq 0.$
the absolute continuity of the critical measure would follow immediately from the dyadic expansion (2.5), but in the non-critical cases the Fourier analysis proves more versatile.

In a more general context, the conditions for absolute continuity of infinite convolutions of Bernoulli measures (ICBM) have also been studied; cf., e.g., Erdös [47].

Kahane and Salem [79] proved the following nice result. For any $b > 1$, let

$$\mathbb{N} \ni q := \left\lceil \frac{\ln 2}{\ln b} \right\rceil \quad (\lceil \cdot \rceil \text{ stands for rounding up}),$$

then the measure $\mathcal{F} [\varphi]$ is Hölder continuous of order

$$\mathbb{R}_+ \ni \alpha := \frac{\ln 2}{\ln b}.$$ (2.10)

In particular, for any integer $q \geq 1$, the measure with scaling factor $b = 2^{1/q}$ is a.c. with density $\rho \in C^{q-1}(\mathbb{R})$. This can be considered as a generalization of [76, Section 6, Example 4].

Special values of the exponent give rise to interesting number-theoretic problems; see the papers by Hardy and Littlewood [69], Mahler [99], and more recent papers by Dubickas, e.g., [41].

Erdős [47] proved that if $b > 1$ is a so-called Pisot–Vijayaraghavan (PV) number, then $\varphi$ does not vanish at infinity, i.e., $\mathcal{F} [\varphi]$ is not a Rajchman measure.

Salem [116] proved the converse of the result by Erdős. Thus $\mathcal{F} [\varphi]$ is a Rajchman measure for Lebesgue-a.e. $b > 1$.

Levin [95] (the original Russian version published in 1979) proved that fractional parts of $(b^n)$ are completely equidistributed; this notion includes estimates for the deviations from equidistribution. (Cf. also Franklin [54].)

However, it is to be emphasized that most of these results apply to an exactly exponential decay rate of $a_{|x|} = u(|x|)$, and this is not the case in dimension $d > 1$ with Euclidean distance $|x - y|$, even for periodic lattices; models with integer-valued distances $|x - y|_1$ and $|x - y|_\infty$ are sometimes simpler. On the other hand, exponential screening in dimension $d \leq 2$ is physically questionable, and the case of screening weaker than exponential is technically easier.

Replacing the fractional parts of an exponential sequence with trajectories of a skew shift on the torus, one comes to the equidistribution problem for $\{n^2 \alpha \}$; see, e.g., a paper by Rudnick et al. [112]. In our problem, this corresponds to a polynomially decaying potential $u$. See also a review by Lyons [98], the works by Strichartz [128, 127, 129] and references therein.

**Remark 2.1.** Notice that for the measures satisfying the so-called Condition (C) introduced by Cramér, viz. $\limsup_{|t| \to \infty} |\varphi_\mu(t)| \leq \zeta < 1$, the analysis of equidistribution is unnecessary for the lower bounds on the sum $S_1(t)$ in (2.9), as $- \ln |\varphi_\mu(t)| \geq \ln \zeta^{-1} > 0$. This might seem like a very weak hypothesis, yet it rules out almost periodicity of the Fourier transform, since $\varphi_\mu(0) = 1$. The convolution analysis for these "poor man’s Rajchman measures" (or rather Cramér’s measures) is elementary and pleasant.
Naturally, the absolute integrability of $\phi(t)$ is neither required for non-singularity of the convolution measure $\mu = F[\phi]$ nor observed, e.g., in the critical case $b = 2$: the function $\sin t/t$ is not absolutely integrable, but square-integrable, and the density of the respective measure $\mu$, the indicator function $1_{[0,1]}$, is an exemplary element of $L^2(\mathbb{R})$, albeit neither smooth nor even continuous.

In this connection, recall that Wiener and Wintner [136], answering a question raised by Nina Bary [7, p. 113], proved that $\kappa^* = \frac{1}{2}$ is the critical decay exponent for the Fourier transforms of singular measures, in the following sense:

(i) no singular measure $\mu$ can have $|\hat{\mu}(t)| \leq C(1 + |t|)^{-\kappa}$ for $\kappa > \kappa^*$;
(ii) for any $\varepsilon > 0$ there are examples where

$$\hat{\mu}(t) = O \left( (1 + |t|)^{-(\kappa^* - \varepsilon)} \right).$$

Item (i) is due to the $L^2$-isometry of the Fourier transform, so [136] is essentially a construction of examples for (ii), complementing those by Menchoff [100] and Kershner [82].

Convolution products are always at least as "regular" as the best of the factors involved. In the case of a.c. measures with nice densities, one can forget about measures and deal directly with their densities viewed as functions. The classical example of B-splines (convolutions of interval indicators) shows that each additional factor, starting from $n = 3$ factors, brings one more derivative (Sobolev scale). The picture is however much more complex for singular measures. Again, the explanation is provided by harmonic analysis.

Convolution powers of a singular measure need not become absolutely continuous, as shows the example of integer-scaled Bernoulli measures [135]. Specifically, when $\mathbb{N} \ni M \geq 3$,

$$\phi(t) = \prod_{k=1}^{\infty} \cos \left( \frac{L}{M^k} \right) \quad (2.11)$$

obeys $\limsup_{|t| \to \infty} |\phi(t)| > 0$, and for large $R$,

$$\frac{1}{2R} \int_{-R}^{R} |\phi(t)|^2 = O \left( R^{\frac{\log 2}{3}} \right).$$

Hu and Lau [72] recently found that

$$\limsup_{|t| \to \infty} \phi(t) = \begin{cases} \phi(\pi), & M = 2n + 1, \\ \leq \phi(\pi), & M = 2n, \end{cases} \quad 3 \leq M \in \mathbb{N}.$$ 

The classical example\(^8\) [51, v.2, Section V.4(d)], on the other hand, shows that convolution products of non-identical singular measures can be more inclined to become more regular than either of the convolution factors. The explanation is simple: taking a product of two different infinite products of the form (2.11), one may sometimes overlap the unit (or nearly unit) factors from one product with very small factors from another product,

\(^8\)I had learned it long ago as a part of probabilistic folklore, but do not know who found it first.
which is impossible for identical products. A similar phenomenon is encountered in the theory of asymptotical expansions of sample distribution functions of sums of independent variables, where a number of results are proved differently (or available at all) with and without the so-called non-lattice distribution condition (cf. \([64, 51]\)).

However, there are many examples of singular measures on locally compact abelian groups with a.c. convolution powers; cf. Hewit and Zuckerman \([70, 71]\), Saeki \([115]\), Karanikas and Koumandos \([80]\). Again, an explanation is provided by harmonic analysis: if (and this is of course a big if) the Fourier transform \(\phi_\mu(t)\) of a measure \(\mu\) does actually have a power-law decay, 
\[
|\phi_\mu(t)| \leq C|t|^{-\varepsilon}, |t| \geq 1, \varepsilon > 0,
\]
then \(\mathcal{F}^{\pm 1}[\mu^\alpha](t) \leq C|t|^{-n\varepsilon}\), so it suffices to take \(n > 1/\varepsilon\). In view of the above mentioned Wiener–Wintner result \([136]\), there are s.c. measures with \(\tfrac{1}{2} < \varepsilon < 1\), so even squares of some s.c. measures are a.c.

The problem in general is that an s.c. can be not from Rajchman class, i.e. with no decay at all, let alone power-law rate.

An important particularity of the measures appearing in higher-dimensional non-local alloy models is that one has there infinite products of Fourier transforms, expressed themselves via infinite products coming from 1D chains filling a \(d\)-dimensional grid, like in \([76, \text{Section 6, Example 4}]\).

**Remark 2.2.** The above discussion is closely related to another topic which I only mention in passing here: improvement of regularity of the cumulative potential (hence, of the IDS/DoS) of multi-layer quasi-1D or quasi-2D media of finite cross-section \(W\), as \(W \to \infty\). Such models are sensitive to the geometric properties of the scatterers support (periodic/aperiodic grid of scatterers) and of the metric \(d(\cdot, \cdot)\) in the configuration space \(\mathcal{X} = \mathbb{Z}^d, \mathbb{R}^d\) figuring in the potential \(V(x; \omega) = \sum_y u(d(x, y))\omega_y\). For example, taking the distance \(|x - y|_\infty\), we would have in a strip of width \(W\) in \(\mathbb{Z}^2\) a convolution of \(W\) identical measures (cf. the paragraph preceding (2.11)). The Euclidean distance gives rise to an irrational and nonlinear scaling when we pass from one layer to another:
\[
r = \sqrt{x^2 + y^2} \sim \sqrt{x^2 + (y + 1)^2} = r\sqrt{1 + \frac{2y + 1}{r^2}},
\]

hence to a convolution of non-identical measures. Its quantitative analysis becomes, therefore, geometry-specific and less universal; it does not belong in this paper. In particular, the extension to Delone–Anderson Hamiltonians would not be automatic. What is clear, is that varying the width \(W\) of a strip, the grid spacing and the decay exponent \(c > 0\) of a potential \(u(r) = e^{-cr}\), one can rig the model so as to recover the classical example \([51, v.2, \text{Section V.4(d)}]\) with an a.c. convolution of s.c. Cantor measures. Therefore, the continuity phase transition is encountered already within the class of quasi-1D systems with exponential screening (be it possible or not in physical models). The distances \(|\cdot|_1\) and \(|\cdot|_\infty\) on \(\mathbb{Z}^d, d \geq 2\), on the other hand, give rise to interesting artificial models where some analytic aspects are simpler than for \(|\cdot|_2\).

**Remark 2.3.** I have consciously avoided using any ergodicity arguments regarding the spatial grid \(\mathcal{Z}\) of the scatterers. For definiteness, it is assumed, e.g., in the first paragraph of Section 5.1, that \(\mathcal{Z} = \mathbb{Z}^d\), but the actual calculations in Eqn. (5.2) evidence that the exact periodicity is unnecessary, provided the number of sites \(x \in \mathcal{Z}\) per sufficiently large
ball is bounded from below (even that can be slightly relaxed). Upper boundedness is only required for a uniform convergence and boundedness of cumulative potential, but not for the regularity as such. Pushing by force a spatially nonhomogeneous environment into the framework of ergodic systems proves quite useful in the analysis of Delone-Anderson Hamiltonians (cf. e.g. [110, 111, 63] and references therein; see also closely related works [85, 43] on ”crooked/trimmed” random operators). In the present context, however, any reference to spatial ergodicity would raise suspicions about a possible smuggling of an additional regularity in a disguised form in the first place; an old and efficient trick of some smart alchemists of the past centuries. Indeed, who says "ergodicity" says "with probability one" (except perhaps for the case of unique ergodicity), but it is obvious from our analysis that taking a "typical" grid \( \mathbb{Z} = \{ c_x, x \in \mathbb{Z}^d \} \subset \mathbb{R}^d \), where \( c_x = x + d_x \) and \( d_x \in \mathbb{R}^d \) are IID r.v. with a bounded density, would beat hands down any singularity of the amplitudes \( \omega_x \) (literally: you can even take \( \omega_x \equiv 1 \)) and produce a \( C^\infty \) PDF of the cumulative potential without breaking a sweat. Therefore, in order not to raise any doubts, when using the randomness of the amplitudes \( \omega_x \), I stick to a completely “quenched” spatial order/disorder of the grid \( \mathcal{Z} \).

On the other hand, in the pure displacements model (with \( \omega_x \equiv 1 \)), it suffices to have \( d_x \) taking two different values, as long as \( \mathbb{R}^d \ni x \mapsto u(|x-y|) \) is sufficiently non-flat.

I believe that a physically realistic modeling ought to take into account some additional continuity of the distribution of the "scatterers", first of all via the displacement degrees of freedom probably provided by the statistical-mechanical description. However, one has to be careful with the choice of the tools, since we deal here with very fine effects, and the classical, not quantum, statistical mechanics may be adequate or not. Lullabying ourselves with "Assume ..." is of course the easiest way. ◦

From a utilitarian, down-to-earth point of view, pointwise decay bounds on the Fourier transforms of (possibly) singular measures, quite handy when available, are far from being necessary for the proofs of Hölder continuity\(^9\) (of some positive order) of the infinite convolutions at hand, as we shall see in Section 5.2. Fortunately enough, integral estimates are both easier to establish and available for a large class of measures. A fairly explicit and constructive characterization of continuous measures whose Fourier coefficients decay in Cesaro sense follows from Wiener’s results [135] obtained in 1924: for a measure \( \mu \) on \([0,1]\) with the set of nonzero atoms denoted \( \Sigma_{p.p.}(\mu) \),

\[
\lim_{|n| \to \infty} \frac{1}{2n+1} \sum_{|k| \leq n} |\hat{\mu}(n)|^2 = \sum_{\lambda \in \Sigma_{p.p.}(\mu)} |\mu(\lambda)|^2,
\]

and for a measure on \( \mathbb{R} \) one has, as is well-known, by the same arguments (cf., e.g., [108, Theorem XI.114])

\[
\lim_{T \to +\infty} \frac{1}{2T+1} \int_{-T}^{T} |\hat{\mu}(t)|^2 dt = \sum_{\lambda \in \Sigma_{p.p.}(\mu)} |\mu(\lambda)|^2,
\]

\(^9\)Such EVC estimates allow in principle for the strongest localization results an MSA based method could provide today, just as the classical Wegner estimate for Lipschitz continuous marginal disorder.
with $\tilde{\mu}(t) := \int_{\mathbb{R}} e^{-itx} d\mu(x)$. As is equally well-known, this is a basis for the celebrated RAGE (Ruelle [113], Amrein and Georgescu [6], and Enss [45]) theorem.

Strichartz [127] established analogs of Wiener’s theorem for expansions in eigenfunctions of various Schrödinger operators, including the Hermite polynomials. Higher $L^p$: cf. [127, Corollary 5.4].

Perhaps, the reader might find unwarranted the amount of attention given above to the singular measures, since this paper focuses mainly on the smooth and multi-dimensional case. Indeed, just mentioning all these works, mostly related to 1D models, is like opening yet another Pandora’s box\(^{10}\), with all kinds of mathematical distractions from the physically most relevant situations.

It is worthwhile emphasizing the following points:

- **Large values of $b$ are not the only possible mechanism leading to singularity of the infinite convolutions of Bernoulli measures:** apart from creating “large” gaps in the Cantor-type support, “squeezing” the unit “mass” to a family of intervals of smaller and smaller total length (in the inductive construction), one can achieve a singular concentration by making Bernoulli measure asymmetric\(^{11}\). Specifically, any measure $\mu_p$ on the interval $[0,1]$ pulled-back by the isomorphism to the set of semi-infinite words $\{0,1\}^\mathbb{N}$ with $p = \mathbb{P}\{\omega_i = 0\} \not\in \{0,1/2,1\}$ has of course the full support $[0,1]$, but for any $p \neq p'$, the measures $\mu_p$ and $\mu_{p'}$ are mutually singular, as follows from the Law of Large Numbers for the limiting frequency of the digits $\omega_i = 0$. This shows that one should not expect “the” critical point for the continuity phase transition(s), but a number of critical parameter zones, where some important parameters are functional (PDF of a measure).

- **A sub-exponential or power-law decay of the potential $u$ results in a higher regularity than just continuity or mere absolute continuity with bounded density. A polynomial decay of any summable order gives rise to a compactly supported $C^\infty$-density.**

- **Whenever $u$ is a “slowly decaying” function (in the sense that its derivative $u'$ decays faster than $u$ itself), a local analysis of the induced single-site random potential reveals in any dimension $d \geq 1$ a Gaussian-like nature, due to a convolution of many independent contributions of comparable amplitudes, resulting in CLT (Central Limit Theorem) type approximations. It is shown below that some simple two-point Gaussian approximations are also possible to obtain. A full-fledged multivariate CLT in arbitrary finite domains is more difficult to establish. However, I conjecture that this is possible.**

---

\(^{10}\)This is one way to put it. A reader familiar with F. J. Dyson’s paper published in "Physics today” in 1967 knows another metaphor, which would sound today much less “politically correct”. Amazingly, the Soviet censorship authorities, otherwise paranoic, happened to overlook it in the well-known monograph by Lifshitz, Gredescul and Pastur [97], on the first page of Chapter II.

\(^{11}\)See e.g. a discussion in the work by Strichartz [128, Section 2] where it is shown that in a more general context of self-similar measures $\mu$, related to contractive maps, the dimension of $\text{supp } \mu$ is maximized by the so-called natural weights figuring in definition of $\mu$. In the Bernoulli case, with identical contraction exponents, the probabilities $p$ and $1 - p$ have to be equal to maximize the dimension of the support.

\(^{12}\)See also [76, Section 8, Example 2].
In any case, the last argument suggests that Gaussian models of disorder, understood within a suitably defined Gaussian “micro-scaling” limit, also should have a fairly universal value, at least as useful guides to more accurate models.

♦ The cumulative potential and the IDS. The relations between the regularity of the single-point marginal probability \( \mu \) of the potential registered at individual sites (cumulative potential, in the alloy models) and the IDS are not quite straightforward. Even in the simplest of the two directions (from regularity of the potential to that of the IDS), it took some time to extend the original Wegner’s result [133] to probability distributions (in the IID case) with an arbitrary continuity modulus. The turning point was the spectral averaging developed by Simon and Wolf [122] and the dimensional reduction via the Birman–Schwinger identity [11, 118] to a commutative, one-dimensional probabilistic analysis closely related to the Boole identity (1857) [12]. The one-dimensional models provide a good laboratory for studying these relations, particularly for deriving singularity of the IDS from that of the potential. Since dimension one is critical for the main phenomenon explored in the present paper, it is certainly worthwhile recalling some known key facts about the IID potentials in 1D.

Due to a result by Simon and Taylor [121], if \( \mu \) has a compactly supported density \( \rho_\mu \in L^\alpha(\mathbb{R}), \alpha > 0 \), then the DoS (in 1D !) exists and is in \( C^\infty(\mathbb{R}) \). As to arbitrary measures not supported by a single point (in the IID case), the IDS is always continuous, by Pastur’s result [103]. Craig and Simon [36, 35] established log-Hölder continuity of the IDS in any dimension (on a lattice), using the duality between the IDS and the Lyapunov exponents in one-dimensional or quasi-one-dimensional systems. On the other hand, in the 1D Bernoulli–Anderson model Simon and Taylor [121] conjectured that, in some parameter zones, IDS cannot be Hölder continuous of any order higher than some critical exponent \( \hat{\alpha} < 1 \). For a Bernoulli measure with values 0 and \( b \), Lipschitz continuity is ruled out for \( b \) large enough.

Carmona et al. [15] proved Anderson localization in 1D lattice models with arbitrary nontrivial disorder; they also proved the above mentioned Simon–Taylor conjecture, making rigorous the heuristic argument outlined in [121] and based upon an adaptation of Temple’s inequality [130] and on a result by Halperin [67]. Formally, it relies on the independence of the values of the potential. However, the key mechanism is the existence of EFs deterministically localized on single impurities embedded into the ambient constant potential, and this mechanism is robust enough to produce singular IDS at least for suitably chosen parameters of the long-range cumulative potential generated by the underlying Bernoulli disorder. Therefore, the dimension one harbors indeed transitions in the IDS measure, from absolute continuity to singular Hölder continuity.

Klein et al. [86] used the supersymmetry approach to prove Hölder continuity in 1D, under a relatively weak assumption on power-law decay of \( \hat{\mu}(t) \) at infinity. By comparison, the Fourier analysis used in our method easily proves \( C^\infty \)-regularity of the DoS under the same hypotheses [86, Eqs. (1.1)–(1.2)], for even a much weaker decay of \( \hat{\mu}(t) \) (viz. a mere fact that \( \mu \) is a Rajchman measure) is a dreams-come-true scenario in the regularity problem for infinite convolutions of (suitably scaled for convergence) singular measures. But recall that [86] deals with a harder, short-range disorder problem. The irony is that
the assumption of finite range of interaction had been initially made in Anderson-type models in order to "simplify" their analysis!

The authors of [86] conjectured that some hypotheses similar to [86, Eqns. (1.1)–(1.2)], i.e., relatively weak decay of the characteristic function \( \hat{\mu}(t) \), should be sufficient for (at least) Hölder continuity of the IDS in any dimension. In fact, their conditions refer to a nontrivial component \( \mu_1 \) in a mixture \( \mu = s\mu_1 + (1-s)\mu_2 \) with \( s > 0 \), regardless of \( \mu_2 \). The present paper only sheds some light on their general conjecture, since

- infinite range of interaction \( u \) is vital for our proofs;
- the case of a mixture is not considered here.

However, I conjecture that the infinite convolution mechanism is akin to the one observed in one-dimensional models, and that the lattice Bernoulli–Anderson Hamiltonian should have smooth DoS in dimension \( d > 1 \). (A mere log-Hölder continuity would suffice for Anderson localization.) The arguments in favor of this hypothesis are as follows. Regularity of the IDS in 1D is derived by Hilbert transform from that of the Lyapunov exponent(s); the latter come from the transfer-matrix analysis and ultimately from the Lyapunov solutions (those which grow exponentially), hence from the Green functions constructed from the Lyapunov solutions. Further, the Green functions traveling across a random media accumulate, like a sort of test functions, the random site-wise fluctuations, which results in multiple nonlinear convolutions. Moreover, an asymptotically sharp analysis by Pastur and Figotin [104] (IID case), and by Chulaevsky and Spencer [25] and Schulz-Baldes and Seidel [117] (correlated case) evidences that in the weak disorder, the linearized convolutions give rise to the ergodic theorem and CLT approximation providing the leading order of magnitude of the Lyapunov exponent, under weaker assumptions that IID (fast decay of correlations). The potential at a site \( y \) contributes to the value of a Green function at \( x \) with a weight which is at worst exponentially small in \( |x-y| \), but our analysis suggests that, were that contribution linear in \( V \), it would suffice for infinite derivability in any dimension \( d > 1 \).

To avoid any misunderstanding, let me stress: Anderson localization in a short-range Bernoulli disorder on a lattice \( \mathbb{Z}^d \) has been a natural and very tempting conjecture floating in the air ever since the publication of [13] in 2005. The conjecture I have mentioned above concerns

- (i) the suggested mechanism of regularity of the IDS, and
- (ii) infinite derivability of the DoS in higher dimension,

for compactly supported interactions \( u \).

◊ “Thin” tails. Another important aspect of infinite smoothness, combined with a.s. boundedness of the probability distribution, is that near every edge \( E^* \) (there may be gaps in its support) it features the decay \( O(|E-E^*|^{\infty}) \). This qualitative result follows from the infinite derivability without calculations or application of the large deviations theory. In this connection, recall that Exner, Helm and Stollmann [49] used earlier a very simple argument in the proof of the initial length scale (ILS) estimate for the MSA, based on the hypothesis of edge decay of the IDS of sufficiently high polynomial order and replacing the Lifshitz tails estimate. Such "ultra-thin" tails are therefore universal, above the critical point of the continuity phase transition for the exponential Bernoulli disorder. From the
perspective of Anderson localization, they are most valuable in the case of sign-definite interactions, but their "thin" nature is universal and does not require interaction to be sign-definite.

It is to be emphasized that the two kinds of "tails" have different nature and refer to different, albeit related, phenomena. The Lifshitz tails refer to upper, lower, or asymptotic bounds on the IDS, and as such do not presume any local regularity property of the IDS measure. The "thin" tails mentioned above are a direct consequence of $C^\infty$-smoothness of the compactly supported density of the cumulative potential. Further, Lifshitz tails asymptotics results from a collective behaviour of a sufficiently large sub-sample, hence manifesting itself in a sufficiently large finite volume, while the infinite derivability is a pointwise property, and as was demonstrated by Exner et al. [49], it can be used in a ball B without requiring it to be large. It is an individual, site-wise response to the collective behaviour of a large number of remote sources from the "thermal bath" surrounding the finite volume at hand. "Freezing" the bath outside some finite ball ("jacuzzi") may destroy continuity, let alone smoothness, but useful probabilistic upper bounds may be preserved (as are the Lifshitz tails unrelated to the thermal bath).

Summarizing, Lifshitz tails are a "ripple" phenomenon while the "thin" tails are tidal. Yet, both give rise to robust mechanisms of the onset of Anderson localization, without a physically questionable condition of strong disorder.

3. Main results

The word potential used alone refers below to the scatterer potential $u : \mathbb{Z}^d \to \mathbb{R}$, generated by the elements of the disordered media (sources), and the sum of potentials from all the sources registered at $x$ is called the cumulative potential.

3.1. Exponential and sub-exponential screening.

**Theorem 3.1.** Consider the potential $u(r) = e^{-mr^\delta}$, $m > 0$, $\delta \in (0, 1]$ and let $d \geq 1$. Then the characteristic functions of the random variables

$$V_x(\omega) = \sum_{y \in \mathbb{Z}^d} u(|y-x|)\omega_x$$

(3.1)

obey the upper bound

$$|\varphi_{V_x}(t)| = \left| \mathbb{E}\left[e^{itV_x(\omega)}\right] \right| \leq (1 + |t|)^{-Cm\ln \frac{d}{\delta} |t|}. $$

Consequently:

(A) for any $d > 1$ and $\delta \in (0, 1]$, as well as for $d = 1$ and $\delta \in (0, 1)$, the r.v. $V_x$ have probability densities $\rho_x \in C^\infty(\mathbb{R})$;

(B) for $d = \delta = 1$ and $m > 0$ small enough, $\rho_x$ have $C^m|m^{-1}|$ continuous derivatives;

(C) for $d = \delta = 1$ and any $m > 0$, the PDF of $V_x$ are Hölder continuous of some order $\alpha > 0$. 


3.2. Power-law screening.

**Theorem 3.2.** Consider the potential \( u(r) = r^{-A} \), \( A > d \) and let \( d \geq 1 \). Then the characteristic functions of the random variables of the form (3.1) obey the upper bound

\[
|\phi_{V_x}(t)| \leq \text{Const} e^{-c|t|^{d/A}}
\]

Consequently, for any \( d \geq 1 \) the r.v. \( V_x \) have probability densities \( \rho_x \in C^\infty(\mathbb{R}) \).

Regularity estimates for the partial sums of the random series (3.1), with \( y \) restricted to finite balls, are given in Sections 5 and 6. Evidently, allowing \( \omega_y \) to have singular distributions prevents one from having descent decay bounds on their characteristic functions. For example, in the case of a Bernoulli distribution a finite linear combination of random amplitudes \( \omega_y \) has a finite pure point support, hence a periodic or quasi-periodic characteristic function. Nevertheless, we shall see that one can still have satisfactory EV concentration bounds for the scaling analysis of localization.

3.3. Localization under exponential screening.

**Theorem 3.3.** Consider the potential \( u(r) = e^{-Mr} \), \( M > 0 \) and let \( d > 1 \). Fix some \( m \in (0,M) \). There exist \( L_0 \in \mathbb{N} \) and \( b = b(d) > 0 \) such that if for some interval \( I \subseteq \mathbb{R} \) and \( L_0 \geq L \) one has

\[
\sup_{E \in I} \sup_{x \in \mathbb{Z}^d} \mathbb{P}\left( \| \mathbf{1}_G \mathbb{G}_{B_{L_0}(x)}(E) \mathbf{1}_{\mathbb{G}_{B_{L_0}(x)}} \| > e^{-mL_0} \right) \leq L^{-b},
\]

then with probability one, \( H(\omega) \) has in \( I \) pure point spectrum with exponentially decaying eigenfunctions, and for any \( x, y \in \mathbb{Z}^d \) with \( |x - y| \geq 3L_k \) and any connected subgraph \( G \subseteq \mathbb{Z}^d \) one has

\[
\mathbb{E} \left[ \| \mathbf{1}_G \mathbb{P}_t (H(\omega) \mathbf{1}_y \| \right] \leq f(|x - y|)
\]

where \( f : (0, +\infty) \rightarrow \mathbb{R}_+ \) is asymptotically exponentially decaying:

\[
\lim_{L \rightarrow +\infty} \frac{\ln |\ln f(L)|}{\ln L} = 1,
\]

so that

\[
f(L) \leq e^{-L^{1-\kappa(L)}}, \quad \kappa(L) \downarrow 0 \text{ as } L \rightarrow +\infty.
\]

The hypothesis (3.2) can be substantially relaxed, in the spirit of the Germinet–Klein bootstrap MSA [60] (cf. also [18]). In this paper, we privilege a simultaneous derivation of (i) exponential decay of the Green functions (resulting in exponential decay of the EFs), and (ii) exponential scaling limit for the key probabilities, required for a similar behaviour of the EF correlators (cf. (3.3)–(3.5)).

A surprising feature of the exponentially screened infinite-range potentials is that, even for the weakest form of discrete marginal disorder (Bernoulli), they allow one to easily establish the strongest form of Anderson localization an MSA-type approach could prove today: exponential scaling limit of the EFs with a.s. exponentially decaying EFs.
3.4. **Localization under sub-exponential screening.**

**Theorem 3.4.** Consider the potential \( u(r) = e^{-r^\delta}, \delta \in (0, 1) \), and let \( d > 1 \). There exist \( L_* \in \mathbb{N} \) and \( b = b(d) > 0 \) such that if for some interval \( I \subset \mathbb{R} \) one has

\[
P\left\{ \| 1_x G_{B_{L_0}(x)}(E) 1_{\partial - B_{L_0}(x)} \| > e^{-mL_0} \right\} \leq L_0^{-b}.
\]  

(3.6)

then with probability one, \( H(\omega) \) has in \( I \) pure point spectrum with exponentially decaying eigenfunctions, and for any \( x, y \in \mathbb{Z}^d \) and any connected subgraph \( G \subseteq \mathbb{Z}^d \) one has, for some \( \kappa = \kappa(\delta) > 0 \),

\[
E\left[ \| 1_x P_I H_G(\omega) 1_y \| \right] \leq e^{-|x-y|^\kappa}.
\]  

(3.7)

3.5. **Localization under power-law screening.** Here the decay rate of the EFCs is the lowest one, among the three considered cases, but recall that one is allowed to have a barely summable polynomial decay of the screened interaction, i.e. a barely stable model from the statistical mechanics standpoint, hence with very slowly decaying correlations.

**Theorem 3.5.** Consider the potential \( u(r) = e^{-r^\delta}, \delta \in (0, 1) \), and let \( d > 1 \). There exist \( L_* \in \mathbb{N} \) and \( b = b(d) > 0 \) such that if for some interval \( I \subset \mathbb{R} \) and \( L_0 \geq L_* \) one has

\[
P\left\{ \| 1_x G_{B_{L_0}(x)}(E) 1_{\partial - B_{L_0}(x)} \| > e^{-mL_0} \right\} \leq L_0^{-b},
\]  

(3.8)

then with probability one, \( H(\omega) \) has in \( I \) pure point spectrum with exponentially decaying eigenfunctions, and for any \( x, y \in \mathbb{Z}^d \) and any connected subgraph \( G \subseteq \mathbb{Z}^d \) one has

\[
E\left[ \| 1_x P_I (H_G(\omega) 1_y \| \right] \leq \frac{C'}{(1 + |x-y|^C).}
\]  

(3.9)

4. **Characteristic functions of sums of random variables**

We restrict our analysis to probability distributions with compact support not reduced to a single point; the latter condition is obvious, while the former is motivated by physical applications (where no interaction between charged particles at distance can be infinite, except perhaps in specific artificial models), and also for a simpler presentation and more uniform set of results. Most notably, boundedness of random amplitudes puts their probability distributions in the universality (attraction) class of Gaussian distributions; of course, the finiteness of the second moment would suffice to that effect, but boundedness, being physically most adequate, results in a more clear picture.

Unlike the approach developed by Bourgain and Kenig [13] and later extended by Aizenman, Germinet, Klein and Warzel [2] to arbitrary non-constant IID local potentials via the Kolmogorov’s decomposition of arbitrary probability measures [90], our method does not use a reduction to an embedded Bernoulli measure.

The method of characteristic functions, i.e., (inverse) Fourier transforms of probability measures, is the most versatile tool for the asymptotic analysis of sums of independent r.v., successfully used both for identical and non-identical individual distributions. A number of general classical facts on the rate of convergence to the limiting distribution is available in the literature, but some of those results will require an adaptation to the problem at hand. The specificity of our analysis resides in the fact that we work not
in the framework of growing sums of r.v. with comparable amplitudes, hence not with the total variance tending to infinity, but with uniformly and absolutely convergent random series\(^{13}\), whose probability distribution cannot be Gaussian, being compactly supported (due to the assumed uniform convergence of the series). However, asymptotically Gaussian measures appear in the micro-scaling analysis, which is most important both for smoothness estimates of the (I)DS induced by an infinite ”thermal bath” and for the localization phenomena, through the Wegner-type estimates.

We consider the case of regularly (periodically) placed scatterers, and only briefly comment the displacement disorder and scatterers placed on Delone sets. We shall see that only crude upper and lower bounds on the number of sites in arbitrarily placed large balls (technically, annuli) are actually required.

Specifically, consider an integer lattice \(\mathbb{Z}^d\) with \(d > 1\). The scatterer potential is now assumed ”realistic”, viz. depending upon the Euclidean distance in \(\mathbb{Z}^d \rightarrow \mathbb{R}^d\), so the random potential values induced at \(\omega x u(|x|)\) with \(x\), say, near the max-norm spheres, \(S_r = \{x : |x|_\infty \approx r\}, r \geq 1\), are non-identically distributed or, better to say, non-identically scaled by the non-random factors \(u(|r|^2)\). However, these amplitudes remain comparable due to smoothness of \(u\); moreover, if \(u\) is ”slowly decaying”, i.e., its gradient decays faster than \(u\) itself, the amplitudes over \(S_r\) differ only by factors \(1 + o(1)\) for \(r \rightarrow \infty\). Naturally, the case of exponentially decaying potentials is somewhat different in this regard, but even in this case the amplitudes remain comparable to each other for all intents and purposes of our probabilistic analysis.

5. Fourier analysis I. Sub-exponential or exponential decay

Despite a great wealth of general results\(^{14}\) on Fourier transforms of self-similar measures on locally compact abelian groups, many of them are related directly only to the models with exponentially decaying potential \(u\) and only when the measure obtained by infinite convolutions is singular, with the support of a fractal dimension. We use below more basic tools applicable also to the models with a slower decay of \(u\).

5.1. Thermal bath estimates. Smoothness. Assume that \(u(r) = e^{-mr^\delta}, \delta \in (0, 1]\). If the random field \(\{\omega_x\}\) is translation invariant, which we assume below for simplicity, it suffices to study the probability distribution and the characteristic function for \(V(0; \omega)\). For notational brevity, we denote \(\mathcal{Z} = \mathbb{Z}^d\).

First, rearrange the series for \(V(0; \omega)\) as follows:

\[
V(0; \omega) = \sum_{x \in \mathcal{Z}} a_{|x|} \omega_x = \sum_{r=0}^{\infty} \sum_{x \in \mathcal{X}_r} a_{|x|} \omega_x
\]

where

\[
\mathcal{X}_r = \{x \in \mathcal{Z} : |x| \in [r, r+1)\}, \quad n_r = \text{card}(\mathcal{X}_r), \quad a_{|x|} = u(|x|).
\]

\(^{13}\)This topic also has been and continues to be popular in probability theory, particularly in the correlated case; cf. the papers by Khintchine and Kolmogorov [83], Kolmogorov [89], Lévy [96]. See also a paper [125] and a book [126] by Stout, and further references therein.

\(^{14}\)Cf. e.g. [73, 128].
Remark 5.1. For further use, notice that we will use the lower bound so for all absolute moment yielding for We have the r.v. following suitable definition of the threshold upon the ratio of the absolute (hence nonzero) moments example, so 

Next, we have the Taylor expansion for the logarithm, valid whenever \(|x| > 3\) valid if one decomposes the spherical layer with inner distance \(r\) into an arbitrarily large number \(B\) of sectors and keeps only one of them for the calculations. For \(r \geq r_0(B)\), each of these \(B\) sectors will have cardinality \(\geq C' r^{d-1} \gg 1\).

To assess the range of \(t\) for which \(|1 - \varphi_V(a_r t)| \sim \frac{1}{2} a_r^2 t^2\), use the finiteness of the third absolute moment \(\mu_3 = \mathbb{E} \left[ |\omega_x|^3 \right]\) and the general moment inequality

\[
(\mathbb{E} [ |X|^a ])^{1/a} \leq \left( \mathbb{E} [ |X|^b ] \right)^{1/b}, \quad 0 < a \leq b,
\]
yielding for \(a = 2\) and \(b = 3\)

\[
\sigma^6 \leq \mu_3^2.
\]

It follows from Taylor expansion for the exponential function that

\[
\forall |t| \leq \frac{\sigma^2}{\mu_3} \quad |1 - \varphi(t)| \leq \frac{1}{2} \sigma^2 t^2 < \frac{1}{2}.
\]

Next, we have the Taylor expansion for the logarithm, valid whenever \(|1 - \varphi(t)| < 1\),

\[
-(\ln \varphi(t) - \frac{1}{2} \sigma^2 t^2) = 1 - \varphi(t) - \frac{1}{2} \sigma^2 t^2 + \sum_{k \geq 2} \frac{1}{k} (1 - \varphi(t))^k
\]
so

\[
\left| \ln \varphi(t) + \frac{\sigma^2}{2} t^2 \right| \leq \frac{1}{6} \mu_3 |t|^3 + \frac{1}{4} \sigma^4 t^4 \leq \frac{5}{12} \mu_3 |t|^3.
\]

We have the r.v. \(a_r \omega_x\) with \(\mathbb{P}\{ |\omega_x| \leq 1 \} = 1\), so \(\mu_3 = \mathbb{E} \left[ a_x^3 \omega_x^3 \right] \leq a_3^3 = e^{-3m\delta}\). If, for example, \(|t| \leq \frac{3\sigma^2}{\mu_3}\), the last RHS \(\leq \sigma^2/4\). Therefore, by taking a suitable \(C\), depending upon the ratio of the absolute (hence nonzero) moments \(\mu_3\) and \(\sigma^2\), we come to the following suitable definition of the threshold \(R_t\) to be used in the sequel:

\[
C|t| = e^{m\delta} \Rightarrow R_t = C m \ln^{1/\delta} |t| = C/m \ln^{1/\delta} |t|,
\]
so for \(|t| > 0, r \geq R_t\)

\[
-\ln |\varphi_V(a_r t)| \geq \frac{\sigma^2}{4} t^2.
\]
Then we have
\[
\sum_{r \geq 1} \sum_{|x| \in [r,r+1)} \ln |\varphi_V(a_{|x|})|^{-1} \geq \sum_{r \geq R_i} \sum_{|x| \in [r,r+1)} \ln |\varphi_V(a_{|x|})|^{-1} \geq C^2 \sum_{r \geq R_i} r^{d-1} \frac{a_r^2}{\theta r}
\]
\[
\geq C^2 \sum_{r \geq R_i} r^{d-1} e^{-2r \delta} \geq C'' t^2 \int_{R_i}^{+\infty} r^{d-1} e^{-2r \delta} dr
\]
\[
= C'' t^2 \int_{2R_i^\delta}^{+\infty} s^{\frac{d-2}{\delta}} e^{-s} ds
\]
\[
= C'' t^2 \int_{2R_i^\delta}^{+\infty} s^{\frac{d-2}{\delta}} e^{-s} ds = C'' \delta^{-1} \Gamma \left( d \delta^{-1}, 2R_i^\delta \right)
\]
\[
\geq C'' t^2 R_i^\delta (d \delta^{-1} - 1) e^{-2R_i^\delta} = C'' t^2 R_i^{-\delta} e^{-2R_i^\delta}
\]
\[
= C''' m^{-1} t^2 (\ln |t|)^{\frac{d-\delta}{\delta}} e^{-2\ln \frac{1}{\delta} |t|}
\]
\[
\geq C''' m^{-1} \delta^{-1} (\ln |t|)^{\frac{d-\delta}{\delta}}.
\]

We used the well-known asymptotics for the incomplete upper Gamma function, showing that the contribution from a single term with \( r = R_i \) cannot be significantly improved by taking the entire tail sum \( \sum_{r \geq R_i} (\cdot) \). Therefore, for any \( R \geq R_i \), one has a similar bound
\[
\sum_{r = R}^{R} \sum_{|x| \in [r,r+1)} \ln |\varphi_V(a_{|x|})|^{-1} \geq C \delta^{-1} (\ln |t|)^{\frac{d-\delta}{\delta}}.
\]

so a suitable upper bound on \( \varphi_R(t) \) can be extended up to \( |t| \lesssim T_R \sim C e^{-R^\delta} \).

For \( d \geq 2 \) and any \( \delta \in (0, 1) \), one has \( \ln^\gamma |t| \) with \( \gamma = \frac{d}{\delta} - 1 \geq 1 \), hence \( |\varphi(t)| \leq C |t|^{-1} \), which may be insufficient for absolute integrability (hence, for smoothness) when \( d = 2, \delta = 1 \), but even in this case sufficient for square-integrability, hence for absolute continuity of the measure \( \mathcal{F}[\varphi] \).

In dimension \( d \geq 2 \), for any \( \delta \in (0, 1) \) one has an upper bound on the characteristic function
\[
|\varphi(t)| \leq e^{-c \ln |t| / |t|} = t^{-c}
\]
implying infinite derivability of the probability density, owing just to the tidal contribution to the IDS. The same is true for \( \delta = 1 \) (exponential decay) and dimension \( d = 2 + \theta \), \( \theta > 0 \). Although one may not be interested in non-integer dimensions, the above calculations evidence that the tidal contribution, in the case of exponential decay, guarantees by itself a power-law decay of any fixed order for a finite-width but sufficiently thick quasi-two-dimensional samples. In practice, this may well correspond to macroscopically very thin films, even without taking into account its 3D substrate. The specificity of the tidal potential is that it is "almost flat" and asymptotically Gaussian, contrary to the non-universal "ripple" component.

With \( d = 2, \delta = 1 \) (exponential decay), we have a bound
\[
|\varphi(t)| \leq e^{-c \ln |t|} = t^{-c}
\]
which may be insufficient for the absolute (or square) integrability of \( \varphi \), but it still is a decay bound! This situation is much better than the one for the characteristic functions of singular Cantor measures featuring no decay of \( t \mapsto \sup_{|s| \leq |t|} |\varphi(s)| \).

Now we shall use Remark 5.1, fix any \( N \ni B > 1 \), and decompose, for \( r \) large enough, the spherical layers \( \{ x : |x| \in [r, r + 1) \} \) into \( B \) sectors, each of cardinality \( \geq Cr^{d-1} \). Each sector provides the characteristic function of the sum of the corresponding r.v. \( a_{|x|} \omega_k \), decaying as \( |t|^{-c} \), \( c > 0 \). By independence, the entire characteristic function, being a product of \( B \) sectorial functions, decays as \( |t|^{-Bc} \), thus is absolutely integrable for \( B > c^{-1} \) and having at least \( c'B \) derivatives, \( c' > 0 \). Since \( B \) can be arbitrarily large, the infinite smoothness of the measure \( F[\varphi] \) follows.

\[ \square \]

Speaking more generally, the "ripple" component of the probability distribution, shaping the overall profile of the distribution of \( V(x; \omega) \), is essentially determined by the specific measure \( \mu \) of the scatterers, while the tidal one is more universal: it is determined by the universality (attraction) class of \( \mu \). For \( \mathbb{E} \left[ |\omega_{x}|^2 \right] < +\infty \), it is Gaussian, otherwise it depends upon the rate of decay of the tail probabilities \( \mathbb{P} \{ |\omega_{x}| > s \} \) as \( s \to \infty \). Therefore, the tidal contribution is universal in two different ways: probabilistically (as a r.v.) and spatially (being asymptotically flat). Flatness\(^{15}\) results in a very simple action of the tidal potential on the EVs of local Hamiltonians (and almost invariance of the associated EFs): without a recourse to the Birman–Schwinger argument, the probabilistic analysis becomes (almost) commutative and one-dimensional. Its universal (e.g., Gaussian) nature may also prove quite useful and give rise to some asymptotically exact calculations, unavailable for general probability measures. \[ \blacklozenge \]

For \( r < R_t \) we cannot benefit in the same way from the quadratic asymptotics. A particular situation with integer scaling factors, corresponding in our model to the two-dimensional lattice with the potential

\[ u(x) = 2^{-|x|_1} \equiv 2^{-|x_1| - |x_2|}, \]

exponentially decaying with respect to non-Euclidean distance \( |\cdot|_1 \), was considered by Jessen and Wintner [76, Section 6, Example 5]. In their example, the final distribution turns out to be an infinite convolution of scaled uniform distributions on finite intervals, hence with a \( C\infty \)-density.

For a pure point marginal measure of \( \omega_x \), we have an almost-periodic characteristic functions, so for Bernoulli or more general discrete distributions of \( \omega_x \), the equidistribution arguments, when applicable, give

\[
\sum_{r \geq 1} \sum_{|x| \in [r, r+1)} \ln |\varphi_V(a_{|x|}t)|^{-1} \geq C' \sum_{r=1}^{R_t} r^{d-1} \geq C'' R_t^d \]

\[ = C'' \ln^{d/\delta} |t| = C \ln^{1+\theta} |t|, \]

where \( \theta > 0 \) whenever at least one of the conditions is fulfilled:

(i) either \( d > 1 \), or

\(^{15}\)Recall that, physically speaking, Friedel oscillations are oscillations, and a more realistic quantitative analysis should take this into account.
(ii) decay of interaction is sub-exponential: $\delta < 1$.

Once again, I would like to emphasize that an exponential screening in dimension $d \leq 2$ may be perhaps nothing more than a mathematical abstraction, and for sub-exponentially decaying potentials $u$ the tidal phenomena (responsible for the Gaussian micro-scaling asymptotics) alone lead to infinite smoothness of the probability distribution of the cumulative potential.

For more general distributions of $\omega_x$, the easiest case is that of sub-exponential interactions, for they are "slowly" decaying functions: with $u(r) = e^{-r^\delta}$, $\delta \in (0, 1)$,

$$|u'(r)| = \frac{\delta}{r^{1-\delta}} u(r) \ll u(r), \quad r \gg 1.$$ 

Therefore, in longer and longer intervals $r \in [R', R'']$, $u(r) \asymp u(R')$, and the argument $u(r)t$ of $\varphi(\cdot)$ evolves asymptotically linearly. Notice that the slope over the long intervals $[R', R'']$, $[R'', R''']$, ..., also varies slowly from one interval to another. In the case of an almost-periodic function $\varphi$, we come again to equidistribution estimates, but in the general case one has to combine the above approach with the Wiener or Strichartz type techniques discussed in the next subsection.

Conclusions. (I). For any $\delta \in (0, 1]$ and $d = 1 + \theta \geq 1$,

$$|\varphi_{V(0; \cdot)}(t)| \leq e^{-C\ln |t|} \lesssim |t|^{-\frac{1+\theta}{\delta}-1}.$$ 

(II) For any $d \geq 2$ and any $\delta \in (0, 1]$, the tidal contribution alone guarantees infinite derivability of the PDF of the cumulative potential $V(x; \omega)$.

5.2. Hölder continuity via integral estimates. Integrable pointwise upper bounds on the characteristic functions $\varphi$ become rather technical and delicate as one approaches the critical point and restricts the mechanisms of regularity build-up to the tidal fluctuations. Integrability of $|\varphi(t)|$ is of course a fast track to absolute continuity of the respective single-site distribution of the cumulative potential. However, there is a simpler way to Hölder continuity of the latter, of some positive order $\beta > 0$. It goes back to another classics of the early 20th century, the Wiener theorem [135]. In the form useful for our purposes, the integral (not pointwise) decay estimates for the Fourier transforms of singular measures were obtained by Wiener and Wintner [136] and thoroughly investigated by Strichartz [127, 128] in a broader context. The bottom line is that for the infinite convolutions of scaled Cantor measures

$$\limsup_{T \to +\infty} \frac{1}{T^{1-\beta}} \int_{-T}^{T} |\varphi(t)| dt < +\infty. \quad (5.3)$$

Actually, [128, Theorem 4.4] provides also a positive lower bound for liminf of the integrals in the above RHS. A particular case of the canonical Cantor measure, where $\beta = \ln 2/\ln 3$, was already studied by Wiener and Wintner [136].

The paper [128], as many other works, focuses on the situation where the limiting convolution measure is singular and supported by a Cantor set of fractional dimension. Technically, this is implemented through the so-called "open set condition" introduced by
Hutchinson [73]. In the simplest case of the Cantor set of possible values of the random series

$$S(\omega) = \sum_{n=0}^{\infty} \frac{\omega_n}{b^n}, \quad b > 2$$

(5.4)

the open set condition basically says that there is a gap of positive length between the sets of values of the series

$$0 \left( \frac{b}{b^0} \right) + \sum_{n=1}^{\infty} \frac{\omega_n}{b^n} \quad \text{and} \quad b \left( \frac{b}{b^0} \right) + \sum_{n=2}^{\infty} \frac{\omega_n}{b^n},$$

so these two sets can be covered by non-overlapping open sets.

Strichartz proved in [128, Section 3] stronger variants of (5.3) for k-fold convolution powers of self-similar measures.

It is clear what (5.3) means for the measure \( \mu = \mathcal{F}[\varphi] \) (here \( \mathcal{F}[\cdot] \) stands for the Fourier transform): in view of the regularity estimates derived from the integral decay properties of \( |\varphi| \), one obtains an upper bound \( \mu([a, a + \varepsilon]) \leq CE\varepsilon^\gamma \) with \( \gamma > 0 \), hence Hölder continuity. On this path, one is still long way from infinite derivability of the DoS, yet much better off than with log-Hölder continuity crucial to the Bourgain–Kenig proof [13] of localization in \( \mathbb{R}^d \).

The approach based on Wiener-type, integral estimates has an advantage: here one gets exactly what is required for an integral estimate of the measure of an interval \([a, a + \varepsilon]\), \( \varepsilon \sim T^{-1} \ll 1 \).

Recall that an alternative is provided by the Kahane–Salem result (2.10), but the techniques from [127, 128] seem easier to adapt to more general scaling sequences (not exactly exponential).

As to the application to Anderson localisation, a reader familiar with the Germinet–Klein bootstrap method [60] knows that Hölder continuity of the IDS, of any positive order, suffices for the decay of the EF correlators with any rate \( f(L) = e^{-L^\zeta}, \zeta \in (0, 1) \). A refinement proposed in [23], used in a modified form in Section 9, allows one to achieve an exponential scaling limit, viz. to prove the decay of EF correlators with \( f \) obeying

$$\lim_{L \to \infty} \frac{\ln \ln f(L)}{\ln L} = 1,$$

once a Hölder continuity of the finite-volume EV concentration is established. For the moment, exponential decay of \( u \) is required for such a strong form of localization.

5.3. "Frozen bath" estimate. Now we assess only the partial sum

$$\sum_{r=1}^{R_t} \sum_{|x| \in [r, r+1]} \ln |\varphi_V(a_x; t)|^{-1} \geq C' \sum_{r=1}^{R_t} r^{d-1} \geq C'' R_t^d$$

(5.5)

$$= C''' R_t^d \ln^d |t| e^{-2\ln \delta \frac{1}{|t|}} = C'''' R_t^d |t|$$

$$= C'''' R_t \ln^{1+\theta} |t|, \quad \theta := \frac{d}{\delta} - 1,$$

where \( \theta > 0 \) whenever at least one of the conditions is fulfilled:

(i) either \( d = 1 + \sigma > 1 \), or
(ii) decay of interaction is sub-exponential ($\delta < 1$).

To make the notion of non-integer dimensionality non-abstract, one can define it via the growth rate of balls; for example, $\mathcal{X} = \{(x, y) \in \mathbb{N}^2 : 0 \leq y \leq Cx^\tau \}$, $\tau \in (0, +\infty)$.

In view of $m$-dependence of $C_{\sigma m}^m = m^{-1}C_\sigma$ (cf. (5.1)), for $m > 0$ small enough one has

$$|\varphi_V(t)| \leq (1 + |t|)^{-q}, \quad q > 1,$$

which is absolutely integrable, albeit this does not imply infinite derivability.

The equidistribution argument, improving the tidal estimate, may seem like a trick, but there is a deep probabilistic mechanism behind it. Saying that each term in the LHS sum in (5.5) contributes by a positive constant amounts to say that each r.v. $\omega_x, |x| \approx r$, contributes to the build-up of Lipschitz regularity of the measure of $V(0, \cdot)$, scale by scale. With $d = 1$ and $\delta = 1$, this may be insufficient and result in a singular Cantor measure: at each scale $e^{-mr}$, there are only so many contributing scatterers $\omega_x$. $\delta < 1$ means ”exponential decay with zero exponent”, which is much better than $2^{-r}$ leading to Lebesgue measure. $d > 1$ is essentially equivalent to a lower decay rate of the series with re-arranged terms (cf. [76, Section 6, Examples 4 and 6]). In other words, either of the conditions $d > 1$ or $\delta < 1$ provides a mechanism for the CLT to kick-in, so the entire series breaks into the sum of longer and longer interval sums, where the latter become more and more regular due to convergence in law to the Gaussian limit.

Quite naturally, when the probability distribution of scatterers’ amplitudes has a non-trivial continuous component, possibly is purely continuous, a better bound is to be expected. But for the moment, the only extension to non-atomic measures I have is more involved than it has to be. Another way around would be a Kolmogorov-type decomposition of $\omega_x$ into the sum of a discrete and continuous r.v., mutually independent; this would result in a loss of efficiency of the final bound. Neither option seems fitting, so I postpone this issue to a forthcoming work. After all, the keyword of this text is discrete disorder, and in addition, physical relevance of exponential screening in dimension $d < 3$ is questionable.

In the analysis of power-law screening in Section 6, most important for lower-dimensional media, we will be spared from this technicality: the tidal, Gaussian-like contribution alone will provide a satisfactory and quite robust regularity bound.

5.4. Back to probability measures. Fix $\varepsilon > 0$ and consider an interval $I_\varepsilon \subset \mathbb{R}$ of length $4\varepsilon$; some calculations are less cumbersome for $I_\varepsilon = [-2\varepsilon, 2\varepsilon]$, which we assume; a more general case $[c - 2\varepsilon, c + 2\varepsilon]$ is similar. It will be convenient to upper-bound its indicator function $\chi_\varepsilon$ as follows:

$$\chi_\varepsilon = 1_{[-2\varepsilon, 2\varepsilon]} \leq 1_{[-4\varepsilon, 4\varepsilon]} \ast \frac{1}{2\varepsilon} 1_{[-\varepsilon, \varepsilon]}$$

$$\leq 1_{[-4\varepsilon, 4\varepsilon]} \ast \frac{e^{-\frac{\varepsilon^2}{2\sigma^2}}}{\sqrt{2\pi \varepsilon}},$$
with $\sigma_\varepsilon = a\varepsilon$, $a \approx 1.2$. Using its Fourier transform

$$\hat{\chi}_\varepsilon(t) = \varepsilon e^{\frac{-\sigma_\varepsilon^2 t^2}{2}}$$

one can assess $\mu(I_\varepsilon)$, introducing some cut-off threshold $\zeta_\varepsilon > 0$:

$$\mu(I_\varepsilon) \leq \int_{|t| \leq \zeta_\varepsilon} \hat{\chi}_\varepsilon(t) \varphi(t) dt + \int_{|t| > \zeta_\varepsilon} \hat{\chi}_\varepsilon(t) \varphi(t) dt \leq J_1 + J_2.$$ 

With $\zeta_\varepsilon \leq T_R$,

$$J_1 \leq C \varepsilon \int_{|t| \leq \zeta_\varepsilon} e^{-C|t|^d/A} dt \leq C' \varepsilon,$$

$$J_2 \leq 2(1 - \Phi(\sigma_\varepsilon \zeta_\varepsilon)) \leq e^{-C \ln^2 \varepsilon^{-1}} \leq \varepsilon^2,$$

hence

$$\mu(I_\varepsilon) \leq J_1 + J_2 \leq C \varepsilon,$$

provided $\zeta_\varepsilon = \varepsilon^{-1} \ln^2 \varepsilon^{-1} \leq T_R \sim C e^{R^3}$. It suffices that, e.g.,

$$e^{R^3} \geq \varepsilon^{-\frac{1}{2}} \iff \varepsilon \geq e^{-\frac{1}{2}R^3}.$$ 

6. Fourier analysis II. Summable polynomial decay

Now we turn to the potentials $u(r) = r^{-A}$, $A > d$, in dimension $d$, aiming essentially at $d \geq 2$, as $d = 1$ calls for more efficient, specifically one-dimensional techniques. Allowing $A > d$ to be arbitrarily close to $d$ may look artificial, but note that for example in dimension $d = 2$ the decay exponent of a screened Coulomb potential can be in some models\(^\text{16}\) $A = \frac{5}{2} = d + \frac{1}{2}$. In any case, let us check that any summable power-law decay can be tolerated in the framework of our general approach.

6.1. Thermal bath estimate. One can repeat here the calculations from Section 5.1 with obvious adaptations to the polynomially decaying potential $u$. However, it will suffice for all purposes to restrict the analysis to the tidal component of the cumulative potential ("frozen bath" estimate) which provides alone a satisfactory bound on the Fourier transform.

6.2. "Frozen bath" estimate. Here a simple technique due to Wintner [138] suffices to achieve satisfactory regularity estimates.

Now we have to set

$$R_\varepsilon = C|t|^{1/A},$$

\(^\text{16}\)Cf. e.g. Gabovich et al.. [59].
so again for $C > 0$ determined by the form of $\varphi_V$, with $\mathbb{E} [\omega_k] = 0$,

$$
\sum_{r \geq R_t} \sum_{|x| \in [r,r+1]} \ln |\varphi_V(a|x|t)|^{-1} \geq C \sum_{r \geq R_t} r^{d-1} a_r^2 t^2 \\
\geq C' t^2 \sum_{r \geq R_t} r^{d-1} - 2A \\
\geq C'' t^2 R_t^{d-2A} = C'' |t|^{2 + \frac{d-2A}{2}} \\
\geq C'_d q_A |t|^{d/A}.
$$

One can see that, unlike the case of sub-exponential decay, a single term with $r = R_t$ is much smaller than the sum $\sum_{r \geq R_t}(\cdot)$. Still, taking for example $R \geq 2R_t$, one obtains a lower bound comparable with the infinite tail sum:

$$
\sum_{r = R_t+1}^{R} \sum_{|x| \in [r,r+1]} \ln |\varphi_V(a|x|t)|^{-1} \geq C't^2 \sum_{r = R_t+1}^{2R_t} r^{d-1} - 2A \geq C''' R_t^{d-2A} |t|^2.
$$

\ding{51} Conclusions: For the polynomially decaying potentials, the tidal contribution alone results in the upper bound

$$
|\varphi_{V(0;\cdot)}(t)| \leq e^{-C'|t|^{d/A}},
$$

thus $\rho_{V(0;\cdot)} \in \mathcal{C}^\infty(\mathbb{R})$.

For the finite-bath version, without the benefit of random fluctuations due to scatterers at all distances $r \geq R$, one can extend the above upper bound on $|\varphi_R(t)|$ up to $|t| \lesssim R^A$. This restriction will be very welcome in the MSA induction, requiring a certain degree of decorrelation at large distances as a surrogate of Independence At Distance.

Its another manifestation is a small but valuable possible improvement of regularity due to taking the finite annular bath radius $R$ slightly larger than the size of a ball to be "heated" by the annular bath at hand, i.e. where the IDS is to be assessed (cf. (7.1)).

Remark 6.1. There is a number of fine technical points and only briefly sketched arguments in this paper, so a reader may wonder, how much of it is really necessary in order to understand the basic regularity mechanisms in disordered models with long-range interaction. In the case of polynomially decaying interactions, apparently encompassing the largest number of physically relevant situations, the above elegant and elementary Wintner’s estimate is essentially all one needs to know.

6.3. Back to probability measures. We use again the smoothed indicator function $\chi_\varepsilon$ introduced in the previous section, with the Fourier transform

$$
\hat{\chi}(t) = \varepsilon \frac{\sin(4\varepsilon t)}{4\varepsilon t} e^{-\frac{\varepsilon^2 t^2}{2}}.
$$

As before, we obtain

$$
\mu(I_\varepsilon) \leq \int_{\mathbb{R}} \chi_\varepsilon(x) dF_R(x) \leq \int_{|t| \leq \mathcal{S}_\varepsilon} + \int_{|t| > \mathcal{S}_\varepsilon} = J_1 + J_2,
$$
with $\mathcal{I}_E \leq T_R$ and

$$J_1 \leq C\varepsilon \int_{|t| \leq \mathcal{I}_E} e^{-C|t|^d/\varepsilon} dt \leq C' \varepsilon,$$
$$J_2 \leq 2(1 - \Phi(\sigma \mathcal{I}_E)) \leq \varepsilon(-C\ln^2 \varepsilon^{-1}) \leq \varepsilon^2,$$

hence $\mu(I_{\varepsilon}) \leq J_1 + J_2 \leq C\varepsilon$, provided $\mathcal{I}_E = \varepsilon^{-1} \ln^2 \varepsilon^{-1} \leq T_R \sim CR^A$. It suffices that, with an arbitrary small $\beta > 0$,

$$R^A \geq \varepsilon^{-1+\beta} \Leftrightarrow \varepsilon \geq R^{-A} \varepsilon^\beta.$$

### 7. Wegner-Type Estimates

Wegner estimates are often obtained by spectral averaging; cf. the original work by Simon and Wolff [122], examples of applications in [31, Theorem 1.1], [124, Theorem 3.2], and some abstract functional-analytic presentation along with an extensive historical review in [114]. For a reader familiar with these techniques, note that the usual assumption of compact support of the scatterer potential (the source of usual problems related to the Unique Continuation Principle and particularly to the lack thereof in lattice models) is easily relaxed to a uniformly summable decay. This might be eclipsed in the work by Combes et al. [31] by the efforts required to overcome the main technical difficulty encountered there, the lack of complete covering of the configuration space by the compact supports of the scatterer potentials, but can be readily seen, e.g., from the proofs of slightly more abstract functional-analytic statements by Sabri [114, Proposition 2.1, Theorem 2.2].

**Proposition 7.1** (Cf. [124, Theorem 3.2]). Let $\mu$ be a probability measure on $\mathbb{R}$, $A$ a self-adjoint operator in a Hilbert space $\mathcal{H}$ and $B \geq 0$ a bounded operator on $\mathcal{H}$. For any bounded interval $I \subset \mathbb{R}$

$$\sup_{\|\Phi\|=1} \int_{\mathbb{R}} \left\langle B^{1/2}P_I(A + tB)B^{1/2}\Phi, \Phi \right\rangle d\mu(t) \leq 6\|B\|s_{\mu}(|I|).$$

A terminological remark: the measure $\mu$ need not be continuous, so $s_{\mu}$ is not exactly the continuity modulus of its PDF $F_\mu : \lambda \mapsto \mathbb{P}\{\mu((-\infty, \lambda])\}$, as the latter term is usually employed in the situation where $F_\mu$ is continuous. It still is what is called the Lévy concentration measure and can be used as such even for measures $\mu$ having atoms. This is not a "legal trickery"; it is shown in [20] with the help of standard probabilistic techniques, routinely used in stochastic (a.k.a. Monte-Carlo) modeling, that any probability measure $\mu$ on $\mathbb{R}$ admits stochastic regularizations of the following form: for any r.v. $X$ with PDF $F_\mu$ and any $\varepsilon > 0$ there is a stochastically equivalent r.v. $\tilde{X}$, i.e. a replica of $X$ on some other probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ but with the same PDF $F_{\tilde{\mu}}$, and another r.v. $\tilde{X}_{\varepsilon}$ on $\tilde{\Omega}$ such that

- $\tilde{\mathbb{P}}\{\left|\tilde{X}(\omega) - \tilde{X}_{\varepsilon}(\omega)\right| \leq \varepsilon\} = 1$;
- $\tilde{X}_{\varepsilon}$ has bounded probability density $\rho_{\varepsilon}$ with $\|\rho_{\varepsilon}\|_\infty \leq \varepsilon^{-1}s_{\mu}(\varepsilon).$
Therefore, one can \( \varepsilon \)-perturb \( \tilde{X} \) and replace it first by \( \tilde{X}_\varepsilon \), plug \( \rho_\varepsilon \) into any usual estimate requiring a bounded probability density, and then get back to the original r.v. \( X \) (equivalent to \( \tilde{X} \)) thus obtaining on the as-needed basis a slightly weaker concentration estimate:

\[
\forall a \in \mathbb{R} \ \forall \varepsilon > 0 \quad \mathbb{P}\left\{ X \in [a-\varepsilon, a+\varepsilon] \right\} \leq \mathbb{P}\left\{ \tilde{X}_\varepsilon \in [a-2\varepsilon, a+2\varepsilon] \right\}.
\]

The main outcome of the analysis of characteristic functions in Sections 5–6 given below is formulated in a way adapted to its applications to the MSA carried out in Sections 9–10. As we have seen, the bottleneck for the regularity estimates inside a given ball \( B_L(u) \) is the size of the area \( A = B_R(u) \supseteq B_L(u) \) where the randomness is not ”frozen” and can contribute by convolutions to the distribution of the cumulative potential \( V(x; \omega) \) at \( x \in B_L(u) \). In turn, the more regular is that distribution, the smaller \( \varepsilon \) one can afford in Wegner estimates, hence the smaller the probability of unwanted events one can achieve in the MSA. On the other hand, \( A \) should not be taken too large, due to the structure of the scaling analysis. The choice of \( R = R_L \) in (7.1) is dictated by the requirements of the MSA in Sections 9 and 10.

Needless to say, for the Bernoulli disorder the distribution of the potential \( V \) with a finite ”thermal bath” and ”frozen” infinite complement is purely atomic.

In this section, as well as in Sections 9–10, we focus for definiteness on lattice models. The main tool for proving EVC bounds in continuum media, spectral averaging, is well-known to work for various types of Hamiltonians. In fact, the Simon–Wolff technique has been applied by Kotani and Simon to the continuum Schrödinger operators \cite{122} shortly after the publication of the paper \cite{122}. An essential hypothesis was complete covering of the configuration space by the supports of the non-negative scatterer potentials; in the long-range models, the covering is not only complete but has infinite multiplicity. Considerable efforts were required to cope with the lack of complete covering in later works; cf., e.g., \cite{31, 124, 85}, a review and some abstract variants of Wegner’s estimate in \cite{114}. A detailed presentation of analytical aspects, inevitable in the case of unbounded self-adjoint operators, would make the present work, already too long, even longer.

\[\blacktriangle\] Note in this connection that the special role of the tidal component of the cumulative potential, particularly in the case of a polynomial (hence slowly decaying) potential \( u \) which I singled out at several places, becomes quite useful for the continuum Wegner estimates. Due to asymptotic flatness of the potentials originating far away from the target point, we encounter here a situation similar to the one considered by Fischer et al. \cite{52, 53} in the framework of Gaussian potentials \( V(x, \omega) \) with continuous argument \( x \in \mathbb{R}^d \). Such random fields also feature a surplus of complete covering: in any bounded domain \( \Lambda \subset \mathbb{R}^d \) one has the representation \( V(x, \omega) = \xi_\Lambda(\omega) \mathbf{1}_\Lambda(x) + \eta_\Lambda(x, \omega) \) where \( \xi_\Lambda(\cdot) \) has a regular (Gaussian) distribution conditional on the ”fluctuation” random field \( \eta_\Lambda(x, \cdot) \). In the simplest, lattice IID Gaussian case, \( \xi_\Lambda \) is merely the sample mean over \( \Lambda \) and \( \eta \) is generated by \( |\Lambda| - 1 \) linearly independent fluctuations chosen arbitrarily among \( \{ V(x, \omega) - \xi_\Lambda(\omega), x \in \Lambda \} \). In fact, flatness of the random background field is not necessary for the reduction employed in \cite{52, 53}. The tidal component in
our model is asymptotically Gaussian. In Section 11 we shall see that two-site correlations are non-degenerate, too, and in a sense, the entire approach presented here provides some justification for the use of genuine correlated Gaussian potentials (at least, as an approximation), but one is still a long way from a multivariate, functional CLT for the tidal potential field.

**Corollary 1** ("Frozen bath" Wegner estimates). Consider a ball $B = B_{L}(u)$ and let

$$R_{L} = \begin{cases} 4L, & \text{for } u(r) : r \mapsto r^{-A} \text{ or } e^{-r^{\delta}}, \delta \in (0,1), \\ L(1 + \frac{c}{\ln R}), & \text{for } u(r) : r \mapsto e^{-mr}, \end{cases}$$

and

$$\varepsilon_{L} = \begin{cases} R_{L}^{-1}, & \text{for } u(r) : r \mapsto r^{-A}, \\ e^{-\frac{1}{4}R_{L}^{\delta}}, & \text{for } u(r) : r \mapsto e^{-r^{\delta}}, \delta \in (0,1), \\ e^{-MRL}, & \text{for } u(r) : r \mapsto e^{-mr}; \end{cases}$$

(7.1) (7.2)

here $\beta > 0$ can be chosen arbitrarily small and $M \geq m$ arbitrarily large, for $L$ large enough. Next, consider the Hamiltonian $H_{B}$ and a larger set $A = B_{R_{L}}(x)$, and introduce the the product probability space $(\Omega_{A}, \mathcal{F}_{A}, P_{A})$ generated by the r.v. $\omega$, with $y \in \mathcal{A}$. Then for all $\varepsilon \geq \varepsilon_{RL}$

$$\mathbb{P}_{A}\left\{ \text{dist}(\Sigma_{B_{L}(x)}, E) \leq \varepsilon \right\} \leq C|B_{L}(x)|\varepsilon.$$  

(7.3)

All the values of $R_{L}$ and $\varepsilon_{L}$ can be easily improved.

As explained in Section 5.2, there is an alternative and easier way to prove Hölder-type bounds on the finite-volume EV concentration, by using integral bounds on the characteristic functions instead of their pointwise counterparts. This would make virtually no difference for exponential or sub-exponential potentials $u$, while making weaker the final results in the case of polynomial decay. I believe that the latter problem can be remedied by a more elaborate technique.

**8. ILS Estimates**

### 8.1. Localization at the band edges.

The first scenario leading to the onset of Anderson localization is more universal and robust than the one considered in the next subsection; here we do not make any assumption on the magnitude of the potential and do not attempt to achieve a global bound on the entire spectrum (which is usually possible in discrete systems and/or in one dimension). This will result in ILS estimates easily adapted to the continuous alloy models in $\mathbb{R}^{d}, d \geq 1$, as well as in a large class of quantum graphs, with tempered underlying combinatorial graphs of coupling vertices. Note that this is the scenario explored by Bourgain and Kenig [13] in the case of the Bernoulli potential, and later extended by Germinet and Klein [62] to general alloy potentials with arbitrary marginal distributions not concentrated on a single point.\(^{17}\)

To achieve a simple upper bound on the Green functions at the energies sufficiently close to the lower edge of the spectrum, we assume non-negativity of the scatterer potential. Therefore, by the standard Weyl-type argument, $\text{inf } \Sigma(H(\omega)) = 0$ with probability

\(^{17}\)Recall that the crucial EVC bound used in [62] was proved in [2], with the help of a result by Kolmogorov [90].
1. Furthermore, to have $\Sigma(B_L(x)) \cap [0, cL^{-2}]$ with sufficiently small $c$ (which can be chosen independently of $L$, provided $L \geq L_*$ with $L_*$ large enough but fixed), one needs that

$$|B_L(x)|^{-1} \sum_{y \in B_L(x)} V(y) \leq C(c), \quad \text{with } C(c) \searrow 0 \text{ as } c \searrow 0. \quad (8.1)$$

Due to the assumed non-negativity of the scatterer potential, (8.1) requires a more explicit condition on the basic random amplitudes $\omega_y$:

$$|B_L(x)|^{-1} \sum_{y \in B_L(x)} \omega_y \leq C(c), \quad \text{with } C(c) \searrow 0 \text{ as } c \searrow 0. \quad (8.2)$$

By standard large deviations estimates, the latter event has probability exponentially small in $|B_L(x)|$, whenever $C(c)$ is small enough.

It is to be emphasized that the above condition is necessary but may be insufficient, depending on the decay rate (and a precise profile, particularly in lower bounds) of the scatterer potential. Indeed, the positive tidal contribution from the ambient environment may significantly increase the mean value of the resulting potential in $B_L(x)$, hence the lowest EV in that ball.

To illustrate the above mentioned phenomenon, consider a one-dimensional lattice with an alloy potential of infinite range, with $u(x) = |x|^{-A}$ and Bernoulli amplitudes $\{\omega_x, x \in \mathbb{Z}\}$. Clearly, the condition of a.s. finiteness of the energy at an arbitrary given site $x$,

$$\sum_{0 \neq x \in \mathbb{Z}} \omega_x |x|^{-A} < +\infty \iff A > 1, \quad (8.3)$$

is weaker than the condition of a.s. finiteness of the energy of interaction of two disjoint half-axes,

$$\sum_{x \in \mathbb{N}^*} \sum_{y \in \mathbb{N}^*} \omega_x \omega_y |x-y|^{-A} < +\infty \iff A > 2. \quad (8.4)$$

Whenever (8.4) is violated but (8.3) holds true, the energy $U_L$ of interaction of a ball of a large radius $L$ with the ambient environment (a union of two half-axes) tends to $+\infty$ as $L \to \infty$:

$$U_L \sim \begin{cases} O(\ln L), & A = 2; \\ O(L^\alpha), & A \in (1, 2). \end{cases} \quad (8.5)$$

While even $O(L^\alpha)$ is relatively small compared to the total potential energy in $B_L(x)$ (of order $O(L)$), so that the elevation of the ground state energy in $B_L(x)$ induced by $U_L$ is vanishing as $L \to +\infty$, this elevation is still strong enough to modify the quantitative parameters of the zone of ”extreme” energies near the bottom of the spectrum, thus enhancing the conventional Lifshitz tails mechanism.

The impact of long-range correlations on the Lifshitz tails asymptotics was studied by Pastur [102].

A conclusion one can draw from this discussion is that the sign-definiteness of the screened interaction between a charged particle with remote ions (we ignore the interparticle interaction in the framework of single-particle Anderson models) is an important issue that cannot be discarded, if a realistic mathematical modelling of localization phenomena is intended. In the present paper, we do not pursue further this issue and rely
essentially on the lower and upper bounds on $|u(\cdot)|$, allowing in particular the decay exponent $A > d$ to be close to $d$. The motivation for this decision is dictated not by the desire to push analytic methods to their limits but first of all by the physical works where the so-called Friedel oscillations have been studied. Depending upon the specific properties of the crystalline media (viz. the asymmetry properties of the Fermi surface) various decay rates of the screened Coulomb interaction are believed to occur; for example, the power-law decay with exponent $A = 5/2$ is observed in some two-dimensional media, so in this case $A = d + \frac{1}{2} < d + 1$.

8.2. Large disorder. Increasing the size of the potential is well-known to increase the size of the energy zone where Lifshitz tails asymptotics or similar behaviour results in the onset of localization, even in the continuous systems; see, e.g., the analysis by Germinet and Klein [61]. We focus however on the discrete case, where the complete localization can occur, to see if it actually does occur in infinite-range alloys with structural disorder.

The ILS for the strongly disordered discrete systems usually do not require the initial scale length $L_0$ to be large (contrary to the extreme energy case where the LDE become efficient only for $L_0 \gg 1$). In fact, the smaller $L_0$, the better is usually for the strong-disorder variant of the ILS. In the model with purely discrete structural disorder the uniform upper bound on the probabilities of the events

\[
\{ \omega : B_{L_0}(u) \text{ is not } (E,m)\text{-NS} \}
\]

requires at least a surrogate Wegner-type estimate, hence a certain degree of continuity of the IDS in the ball $B_{L_0}(u)$. Were we interested only in the EVC estimate for an individual ball, it would suffice to set $L_0 = 0$, i.e., restrict the analysis to single-point ”balls”, apply the previously established universal continuity of the single-point effective potential (be it a.c. or s.c.), and take the coupling constant large enough. Then the usual, very robust argument going back to [39] would prove that

\[
\lim_{g \downarrow 0} \mathbb{P}\{ B_{L_0}(u) \text{ is not } (E,m)\text{-NS, for } H_g = H_0 + gV \} = 0.
\]

However, we need more than that: the EVC bound for the Hamiltonian in $B_{L_0}(u)$, in order to fit into the subsequent scale induction, must

- rely only on the disorder in a relatively small neighborhood of $B_{L_0}(u)$, and
- be stable under the fluctuations outside the above mentioned neighborhood of $B_{L_0}(u)$ carried inside $B_{L_0}(u)$ by the non-local scatterer potential.

Thus some additional technical analysis is in order.

Specifically, the analysis in Section 11 allows one to establish an even stronger EVC estimate, based on the ”regularity” of the two-point correlation measure for the pairs $(V(x; \omega), V(y; \omega))$, $x \neq y$, and establish an analog of the result easy to prove for the IID potentials with arbitrary continuous marginal distribution, no matter how singular:

\[
\forall \varepsilon > 0 \lim_{g \downarrow 0} \mathbb{P}\left\{ \min_{\substack{x,y \in B_{L_0}(u) \atop x \neq y}} |V(x; \omega) - V(y; \omega)| \leq \varepsilon, \text{ for } H_g = H_0 + gV \right\} = 0.
\]
8.3. **Dilute alloys.** Now I would like to briefly mention a third scenario: a substitution alloy of two (possibly more) kinds of atoms of which one (say, type I) creates lowest potential values \( U^{(I)} \) well-separated from all the others and has a sufficiently low concentration. Here the discrete nature of the disorder induced by the type I atoms plays an important role in the estimates and arguments, while in the other scenarios a discrete (e.g., Bernoulli) disorder was merely tolerated. Variational arguments imply in such a situation the existence of a spectral band (or better to say, zone, for its gapped or contiguous nature requires a more thorough analysis) emerging around \( U^{(I)} \) and separated by a gap from the rest of the spectrum. The size of the gap depends both on the magnitude of the hopping terms in the kinetic energy operator and on the rate of decay (perhaps, also on the precise profile) of the scatterer potential; the latter can be viewed as a perturbation of the local alloy potential with non-overlapping supports of the individual scatterers.

More generally, one could consider several types of atoms and have type I atoms, again of low concentration, with the potential values inside the spectrum but still well-separated from the remaining spectral bands.

9. **Localization: exponentially decaying potentials**

Now we turn to the proofs of Anderson localization and consider first the class of exponentially decaying potentials. Although this class might seem very close to the short-range interactions, curiously, it turns out to be fairly easy to achieve the strongest possible decay not only for the eigenfunctions (derived by standard methods from exponential decay of Green functions) but also for their correlators (due to strong bounds on probabilities of unwanted events). Specifically, as shows Theorem 3.3, one can establish here exponential scaling limit, in the spirit of the renormalization group approach.

Some arguments below are sketchy, and estimates are not optimal.

The technique used below is very close in spirit to the Bootstrap MSA developed by Germinet and Klein [60]. Technically, however, it is based on an observation made in [19, 21] that already the localization bounds from early papers by von Dreifus [38], Spencer [123], and von Dreifus and Klein [39] are actually self-improving in the course of scale induction, although the surplus of localization produced on scale \( L_k \) was not put in earlier works in a feedback loop for the scale \( L_{k+1} \). It is fair to say that the adaptive feedback scaling described in [23] is a perpetual bootstrap, not necessarily based on the scale-free bounds, and as such is a variant of renormalization group analysis. A small improvement made in this section, compared to [23], is that the exponential scaling limit for the EF correlators is proved simultaneously with the proof of genuine exponential decay of the EFs.

More precisely, the key ingredients for these two results are simultaneously obtained on the level of Green functions. The derivation of exponential spectral and asymptotically exponential dynamical localization is then a simple, well-understood exercise (cf. [23] with references to the original works, particularly [44, 60]).

Given an integer \( L_0 > 1 \), which will eventually assumed large enough, let

\[
\tau_k = \frac{1}{k^{1/4}}, \quad Y_{k+1} = L_k^{1+\tau_k} \quad L_{k+1} = Y_{k+1} L_k.
\]
We shall assume that $Y_1 \geq 40$, and set
\[ \sigma_k = \frac{Y_k}{20}\sigma_{k-1} = \sigma_0 \prod_{i=1}^{k} Y_i. \]

Next, let
\[ \gamma_k = 4 \sum_{x:|x|>4L_k} |u(|x|)|. \]

For further use, note that
\[ L^2_k \geq \left( \prod_{i=1}^{k} Y_i \right)^2 \geq Y^{k^2}_1 \geq (Y_1)^{k^{1/2}}, Y_1 > 1. \]

Also,
\[ \ln \ln L_k \sim \ln \prod_{i=1}^{k} (1 + \tau_k) \sim \int_{1}^{k} s^{-1/2} ds \sim k^{1/2}, \Rightarrow \ln L_k \sim e^{k^{1/2}} \ll e^{C_k}. \]

For $B > 1$ and any fixed $\sigma > 0$,
\[ L_{k+1}^{-2\sigma_k} \geq e^{-2\ln L_k + B^{-k}Y_{1-k}} \gg (e^{-mL_k})^{1-\sigma}. \]

A very natural approach to localization in long-range potential was proposed by Kirsch et al. [84]; their technique is adapted below to arbitrary nontrivial disorder.

9.1. Deterministic analysis of Green functions. We start with the definition of a standard notion of the MSA.

**Definition 9.1.** Fix a ball $B = B_{L_k}(x)$ with $k \geq 0$, $E \in \mathbb{R}$, and $\varepsilon > 0$. A function $V_B : B_L(x) \to \mathbb{R}$ is called

1. $(E, \varepsilon)$-NS (non-singular iff $E$ is not in the spectrum of $H_B = -\Delta_B + V_B$ and the resolvent obeys with
   \[ ||1_{|E-B|} G_B(E)1_{|E-B|}|| \leq (3L)^d \varepsilon; \]

2. $(E, \gamma)$-NR, with $k \geq 1$, iff
   \[ \text{dist} \left( \Sigma(H_{B_k}(u)), E \right) > \gamma \]
   \[ \text{and } (E, \gamma)$-CNR (completely non-resonant) iff for all $j = 1, \ldots, Y_k$
   \[ \text{dist} \left( \Sigma(H_{B_j}(u)), E \right) > \gamma. \]

The next definition is tailored to fit Lemma 9.1. Observe that the conditions in items (1) and (2) have different structure; this is because the non-resonance is going to be the first to check at each induction step; the verification will be straightforward and will not require induction. The second one will be derived inductively.

**Definition 9.2.** Let be given a ball $B = B_{L}(x)$. A function $q_{Bld} : B_{4L}(x) \to \mathbb{R}$ is called

1. $(E, \varepsilon)$-SNS (strongly non-singular or stable non-singular) iff for any configuration of amplitudes $q_{B} \in \Omega_B$ the function $V_B = U[q_{Bld} + q_{Bld}]_{B_{4L}}$ is $(E, \varepsilon)$-NS.
(2) \((E, \gamma)\)-SNR (strongly NR, or stable NR) iff for the configuration \(\Omega_{B^c} \ni q'_{B^c} \equiv 0\) the function \(V_B = U[q_{B^c} + q'_{B^c}]\) is \((E, \gamma + \gamma_k)\)-CNR.

The following deterministic statement is a simple adaptation of the well-known result of the MSA, in essence going back to the works [38, 123, 39], it is very close to [23, Lemma 2]. Its proof relies upon an important geometrical property of the Euclidean spaces and embedded periodic lattices, called in [23] \textit{uniform scalability}: given any integer \(Y > 1\), any ball of radius \(YL\) can be covered by \(M(Y) \in \mathbb{N}\) balls of radius \(L\), with \(M(Y)\) independent of \(L \in \mathbb{N}\). This property, first used by Spencer [123], significantly reduces combinatorial entropy in intermediate calculations and renders much more efficient scaling algorithms of the MSA. It is instrumental to the Germinet–Klein bootstrap MSA [60]. However, it was shown in [23] that it can be relaxed, so that the exponential scaling limit of the Green functions (and, as a result, of eigenfunctions and of their correlators) can be achieved in a larger class of graphs of tempered (polynomial) growth rate of balls. In the form given below\(^{18}\) it appears in [18], so its proof will be omitted for brevity.

Given \(L_{k+1} = Y_{k+1} L_k\), we cover redundantly the lattice with balls of radius \(L_k\), called admissible, so that a ball of size \(L_{k+1}\) is covered by \(C_d Y^d_{k+1}\) sub-balls of radius \(L_k/3\) of admissible balls; cf. [60, 23].

\textbf{Lemma 9.1} (Conditions for non-singularity; cf. [18, Lemma 4.1]). Suppose that \(B = B_{L_{k+1}}\) fulfills the following conditions:

(i) \(B\) is \(\epsilon^{1-c}\)-NR, \(c \in (0, 1)\);

(ii) \(B\) contains no collection of balls \(\{B_{L_k}(x_i), 1 \leq i \leq S + 1\}\), with pairwise \(4L_k\)-distant admissible centers, neither of which is \((\epsilon_k, E)\)-NS.

If

\[ N_{k+1} = Y_{k+1} - 24S_{k+1} + 3 \geq 1, \tag{9.6} \]

then \(B\) is \(\epsilon^{N_{k+1} + c}\)-NS.

The next important result stems easily from Lemma 9.1.

\textbf{Lemma 9.2.} Consider a ball \(B = B_{L_k}(u)\). Suppose:

(1) \(B\) contains no collection of balls \(\{B_{L_k}(x_i), 1 \leq i \leq S + 1\}\), with pairwise \(4L_k\)-distant admissible centers, neither of which is \((\epsilon_k, E)\)-NS;

(2) \(B\) is \(\epsilon^{1-c}\)-SNR. If

\[ N_{k+1} = Y_{k+1} - 24S_{k+1} + 3 \geq 3, \tag{9.7} \]

then \(B\) is \(\epsilon^{N_{k+1} + c}\)-SNS.

\textit{Proof.} One has to show that, with a fixed configuration \(q' \in \Omega_{B_{4L_{k+1}}(u)}\), the ball \(B\) is \((E, \epsilon)\)-NS for any cumulative potential \(V = U[q' + q'']\) for any choice of the complementary configuration \(q'' \in \Omega_{\mathbb{Z}^d \setminus B_{4L_{k+1}}(u)}\). It suffices to check that the assumptions (i) and (ii) of Lemma 9.1 remain fulfilled, no matter how \(q'' \in \Omega_{\mathbb{Z}^d \setminus B_{4L_{k+1}}(u)}\) is chosen, without changing \(q' \in \Omega_{B_{4L_{k+1}}(u)}\).

\(^{18}\)The factor 24 in (9.6) is 4 times larger than in [18, Lemma 4.1], due to the \(4L_k\)-belts around singular balls.
The stability of (i) follows from the min-max principle. As to (ii), it is readily seen that the support $q'' \in \mathcal{Q} \setminus B_{4L_k+1}(u)$ lies outside all the balls $B_{4L_k}(u_i) \subset B_{L_{k+1}}(u)$, hence cannot affect their strong non-resonance property.

Thus it follows from Lemma 9.1 that $B_{L_{k+1}}(u)$ is $(E, \varepsilon_{k+1})$-SNS with
\[
\ln \varepsilon_{k+1} = N_{k+1}m_kL_k - \sigma_k \ln L_{k+1}
\geq m_kL_kY_{k+1} \left( 1 - 24 \frac{S_{k+1}}{Y_{k+1}} \right) - \sigma_k \ln L_{k+1}
\geq L_{k+1}m_k(1 - \eta_k), \quad \eta_k := 24L_k^{-1} - \frac{\sigma_k \ln L_{k+1}}{m_kL_{k+1}}
\] (9.8)

Further,
\[
\sigma_k = \sigma_0B_1 \cdots B_k = \sigma_0L_0^{-1} 50^{-k}L_k
\] thus
\[
\frac{\sigma_k \ln L_{k+1}}{m_kL_{k+1}} \leq \frac{2\sigma_0 \ln L_{k+1}}{m_0L_0 50^kY_{k+1}}
\]

Recalling (9.2) and $\tau_k = k^{1/4}$, we conclude that (cf. (9.1) and (9.8))
\[
24L_k^{-\tau_k^2} + \frac{\sigma_k \ln L_{k+1}}{m_kL_{k+1}} \leq \text{Const} \left( Y_1^{-k^{1/2}} + (25)^{-k} \right),
\]
whence the convergence of the infinite product $\prod_{j \geq 1} (1 - \eta_j) \geq \frac{1}{2}$.

This shows that we can carry out induction in the length scales $L_k$ with the decay exponent ("mass") uniformly lower-bounded by $m_0/2 > 0$.

This ends the deterministic part of the scaling analysis. The main conclusion here is that, contingent upon a successful completion of the probabilistic analysis, one would eventually prove a genuine exponential decay of the Green functions at any scale, and then derive exponential decay of localized eigenfunctions by a well-known argument from [39].

Our next goal is exponential scaling limit for the probabilities of unwanted events, crucial to a similar behaviour of the EF correlators (cf. [60]).

### 9.2. Scaling of probabilities.

**Lemma 9.3** (Factorization of the probability of a bad cluster). Suppose that
\[
\mathbb{P}\{B_{L_k}(u) \text{ is not } (E, m)\text{-SNS}\} \leq p_k
\]
Let $S_{k+1}$ be the maximal cardinality of a collection of balls $B_{L_k}(u_i)$, $i = 1, 2, \ldots$, with pairwise $4L_k$-distant admissible centers, of which neither is $(E, m)$-SNS. Then
\[
\mathbb{P}\{S_{k+1} \geq S + 1\} \leq C_dY_{k+1}^{-S+1} p_k^{S+1}.
\]

**Proof.** Using induction, it suffices to prove that
\[
\mathbb{P}\left\{ B_{L_k}(u) \text{ is not } (E, m)\text{-SNS} \mid \tilde{\mathcal{Q}} \setminus B_{4L_k}(u) \right\} \leq p_k.
\]
By Lemma 9.1 the event \( \mathcal{J}_i := \{ B_{L_k}(u) \text{ is not } (E, m)\text{-SNS} \} \) is \( \mathfrak{F}_{dL_k(u)} \)-measurable, so the above bound holds and implies the claim by induction, since the number of all collections of \( S \) admissible centers \( u_i \in B_{L_{k+1}}(u) \) is bounded by \( Y_{k+1}^d \).

**Lemma 9.4** (Scaling of probabilities). Assume that

\[
\sup_{u \in \mathbb{Z}^d} \mathbb{P} \{ B_{L_k}(u) \text{ is not } (E, m)\text{-SNS} \} \leq p_k
\]

Then

\[
\sup_{u \in \mathbb{Z}^d} \mathbb{P} \{ B_{L_{k+1}}(u) \text{ is not } (E, m)\text{-SNS} \} \leq p_{k+1}.
\]

**Proof.** By Lemma 9.1, if \( B_{L_{k+1}}(u) \) is not \( (E, m)\text{-SNS} \), then either it is not \( (E, \mathcal{Y}_{k+1}\text{-SNR} \) or it contains a collection of at least \( S + 1 \) balls \( B_{L_k}(x_i) \) neither of which is \( (E, \varepsilon)\text{-SNS} \), with admissible and pairwise \( 4L_k \)-distant centers.

The probability of the former event is assessed with the help of the Wegner-type estimate, relying on the disorder in the balls \( B_{L_k}(u) \subseteq B \). Even the largest among them, \( B_{L_{k+1}}(u) \), is surrounded by a belt of width \( 4L_{k+1} \) where the random amplitudes are not fixed hence can contribute to the Wegner estimate with \( \varepsilon = L_k^{-2\sigma_k} \), hence the same is true for all of these balls: for any \( j \), we have

\[
\mathbb{P} \{ B_j(u) \text{ is not } (E)\text{-SNR} \} \leq Y_{k+1}^d L_k^{-2\sigma_k} = L_k^{-2\sigma_k + \tau_k} = L_k^{-\sigma_k + \frac{2\sigma_k - \tau_k}{1 + \tau_k}} \leq \frac{1}{2} L_k^{-\sigma_k}.
\]

The total number of such balls is \( L_{k+1} \), so with \( L_0 \) large enough,

\[
\mathbb{P} \{ B_{L_{k+1}}(u) \text{ is not } (E, \varepsilon_R)\text{-CNR} \} \leq L_k^{d+1} L_k^{-\sigma_k} = L_k^{-\sigma_k (1 - c_k)} , \quad c_k := \frac{d + 1}{\sigma_k}.
\]

Now assess the probability of \( (S + 1)\text{-fold singularity}:

\[
\mathbb{P} \{ (S + 1) \text{ distant singular balls} \} \leq \frac{1}{(S + 1)!} (C_d Y_{k+1}^d p_k)^{S+1}
\]

where \( C_d Y_{k+1}^d p_k \) can be made small using the ILS assumption (3.2), so

\[
-\frac{\ln p_{k+1}}{\ln L_{k+1}} \geq S_{k+1} \sigma_k (1 - c_k)
\]

\[
\geq \frac{1}{2} S_{k+1} \sigma_k \geq \frac{Y_{k+1}^{1-\tau_k}}{20} \sigma_k = B_{k+1} \sigma_k = \sigma_{k+1}.
\]

whence

\[
\mathbb{P} \{ B_{L_{k+1}}(u) \text{ is not } (E, \mathcal{Y}_{k+1}\text{-SNS}) \} \leq \frac{1}{2} L_k^{-\sigma_{k+1}} + \frac{1}{2} L_k^{-\sigma_{k+1}} = L_k^{-\sigma_{k+1}}.
\]

\[ \square \]
9.3. **Exponential scaling limit.** By construction of the sequence $\sigma_k$,

$$\ln \ln L_k^{\sigma_k} = \ln (\sigma_k \ln L_k) = \ln \ln L_k + \sum_{i=1}^k B_i$$

$$\geq \sum_{i=1}^k \ln B_i = \sum_{i=1}^k (1 - \tau_i) \ln Y_i.$$ 

Since $\tau_j \downarrow 0$ and $Y_j \uparrow +\infty$, it follows that

$$\lim_{k \to \infty} \frac{\ln \ln L_k^{\sigma_k}}{\ln L_k} \geq \lim_{k \to \infty} \frac{\sum_{i=1}^k (1 - \tau_i) \ln Y_i}{\sum_{i=1}^k \ln Y_i} = 1,$$

thus

$$L_k^{-\sigma_k} \leq e^{-L_k^{1-\kappa_k}}, \quad \kappa_k \downarrow 0.$$ 

10. **Localization: polynomial potential**

10.1. **Deterministic analysis.**

**Definition 10.1.** Let be given a ball $B = B_L(x)$. A function $q_{B_{4L}} : B_{4L}(x) \to \mathbb{R}$ is called

- (1) $(E, \varepsilon)$-SNS (strongly non-singular, or stable non-singular) iff for any configuration of amplitudes $q_{B^c} \in \Omega_{B^c}$ the function $V_B = U[q_{B_{4L}} + q_{B_{4L}}]_{B_{4L}}$ is $(E, \varepsilon)$-NS;

- (2) $(E, \gamma)$-SNR (strongly NR, or stable NR) iff for the configuration $\Omega_{B^c} \ni q_{B^c} \equiv 0$ the function $V_B = U[q_{B_{4L}} + q_{B_{4L}}]_{B_{4L}} = U[q_{B_{4L}}]_{B_{4L}}$ is $(E, \gamma + \gamma_k)$-CNR.

An immediate application of the above definition is that any event of the form

$$A(B_L(x), E, m) = \left\{ V_q(\cdot, \omega)\big|_{B_L(x)} \text{ is } (E, m)\text{-SNS} \right\}$$

is measurable w.r.t. the sigma-algebra $\mathcal{F}^{q}_{B_{4L}(x)}$. Indeed, the potential induced on $B_L(x)$ by the scatterers supported by $y \notin B_{4L}(x)$ has sup-norm within the stability limits of the projection on $B_L(x)$ of any sample $q' \in A(B_L(x))$.

From this point on and until the end of this section, we consider a scatterer potential $u(r) = r^{-\alpha}$, $A > d$. Further, fix an arbitrary number $b > d$, which will represent the polynomial decay rate of the key probabilities in the MSA induction, and let

$$\alpha > \tau > \frac{b}{A - d}, \quad \mathbb{N} \ni S > \frac{b\alpha}{b - ad}, \quad L_{k+1} = \lfloor L_k^{\alpha} \rfloor, \quad k \geq 0, \quad (10.1)$$

with $L_0$ large enough, to be specified on the as-needed basis. A direct analog of the deterministic Lemma 9.1 is the following statement.

**Lemma 10.1** (Conditions for strong non-singularity). *Consider a ball $B = B_{L_k+1}(u)$, $k \geq 0$, and suppose that*

(i) $B$ is $L_{k+1}^{-A\tau}$-SNR;

(ii) $B$ contains no collection of balls $\{B_{L_k}(x_i), 1 \leq i \leq S + 1\}$, with pairwise $2L_k^{\tau}$-distant centers, neither of which is $(E, m)$-SNS.

*Then $B$ is $(E, m)$-SNS.*
Proof. Derivation of the NS property can be done essentially in the same way as in [39], with minor adaptations.

To show that the strong (stable) non-singularity property also holds true, one can follow the same path as in the proof of Lemma 9.1: the quantitative adaptation to the case of a polynomial decay have been already made in Definition 10.1, replacing its counterpart for the exponentially decaying potentials, Definition 9.2. Stability of the non-resonance condition is immediate, owing to the min-max principle, and the SNS property is derived recursively. On the scale $L_0$ the non-singularity is derived from non-resonance, with a comfortable gap between an energy $E$ and the spectrum in the ball of radius $L_0$. □

10.2. Probabilistic analysis.

**Lemma 10.2** (Factorization of the probability of a bad cluster). Suppose that

$$\mathbb{P}\{B_{L_k}(u) \text{ is not } (E,m)\text{-SNS}\} \leq p_k.$$

Let $S_{k+1}$ be the maximal cardinality of a collection of balls $B_{L_k}(u_i)$, $i = 1, 2, \ldots$, with pairwise $2L_k^{\tau}$-distant admissible centers, of which neither is $(E,m)$-SNS. Then for any integer $S \geq 0$

$$\mathbb{P}\{S_{k+1} > S\} \leq C_dY_{k+1}^{(S+1)d}p_k^{S+1}.$$

Proof. Using induction on $j \in [1, S]$, it suffices to prove that, with

$$A_j = \bigcup_{i=1}^j B_{2L_k^{\tau}}(u_i), \quad j = 1, \ldots, S,$$

one has

$$\mathbb{P}\left\{B_{L_k}(u_{j+1}) \text{ is not } (E,m)\text{-SNS} \mid \tilde{\mathcal{F}}_{\mathcal{Z} \setminus A_j} \right\} \leq p_k. \quad (10.2)$$

By Lemma 10.1 the event $\{B_{L_k}(u) \text{ is not } (E,m)\text{-SNS}\}$ is $\tilde{\mathcal{F}}_{B_{L_k}(u)}$-measurable, so (10.2) holds true. Hence the claim follows by induction, since the number of all collections of $(S+1)$ admissible centers $u_i \in B_{L_{k+1}}(u)$, distant or not, is bounded by $C_dY_{k+1}^{(S+1)d}$. □

**Lemma 10.3** (Scaling of probabilities). Assume that

$$\sup_{u \in \mathbb{Z}^d} \mathbb{P}\{B_{L_k}(u) \text{ is not } (E,m)\text{-SNS}\} \leq p_k \leq L_k^{-b}.$$

Then

$$\sup_{u \in \mathbb{Z}^d} \mathbb{P}\{B_{L_{k+1}}(u) \text{ is not } (E,m)\text{-SNS}\} \leq p_{k+1} \leq L_{k+1}^{-b}.$$

Proof. By Lemma 10.1, if $B_{L_{k+1}}(u)$ is not $(E,m)$-SNS, then either it is not $(E, \gamma_{k+1})$-CNR or it contains a collection of at least $S+1$ balls $B_{L_k}(x_i)$ neither of which is $(E,m)$-SNS, with admissible and pairwise $2L_k^{\tau}$-distant centers.

The probability of the former event is assessed with the help of the Wegner-type estimate, relying on the disorder in the balls $B_{jL_k^{\tau}}(u) \subseteq B$. Even the largest among them, $B_{L_{k+1}}(u)$, is surrounded by a belt of width $L_{k+1}^{\tau}$ where the random amplitudes are not fixed hence can contribute to the Wegner estimate with $\varepsilon = \varepsilon_{R_{k+1}}$, $R_{k+1} = L_{k+1}^{\tau}$, hence the same is true for all of these balls: for any $j$, we have (cf. (7.2) and (7.3))

$$\mathbb{P}\{B_{jL_k^{\tau}}(u) \text{ is not } (E, \varepsilon_{R})\text{-SNR}\} \leq (L_{k+1}^{\tau})^{-A+d+\beta} \quad (10.3)$$
where we are free to choose $\beta > 0$ as small as we please (the actual correction to the power $A-d$ is logarithmic). Since $\tau > (b+1)/(A-d)$, we can pick $\beta$ so small that $(A-d-\beta)\tau > b+1$, hence the RHS of (10.3) is bounded by $\frac{1}{2}L^{-b'-1}_{k+1}$ with $b' > b$.

The total number of such balls is $Y_{k+1} = L^{\alpha-1}_{k+1} = L^{1-\alpha^{-1}}_{k+1}$, with $1 - \alpha^{-1} < 1$, therefore,

$$
\mathbb{P}\left\{ B_{L^{1+\tau}_{k+1}}(u) \text{ is not } (E, \epsilon_R)\text{-CNR} \right\} \leq \frac{1}{2}L^{-(b'+1)+1}_{k+1} < \frac{1}{2}L^{-b}_{k+1}.
$$

By Lemma 10.2

$$
\mathbb{P}\{ S_{k+1} > S \} \leq \frac{Y_{k+1}^{S+1}}{(S+1)!} p^{S+1}_{k} \leq \frac{1}{2} L^{-(S+1)b}_{k} \leq \frac{1}{2} L^{-(S+1)b}_{k+1},
$$

whence

$$
\mathbb{P}\{ B_{L^{k+1}_{k+1}}(u) \text{ is not } (E,m)\text{-SNS} \} \leq \frac{1}{2} L^{-b}_{k+1} + \frac{1}{2} L^{-b}_{k+1} = L^{-b}_{k+1}.
$$

By induction on $k$, we come to the conclusion of the fixed-energy MSA under a polynomially decaying interaction.

**Theorem 10.1.** Suppose that the ILS estimate

$$
\sup_{u \in \mathbb{Z}^d} \mathbb{P}\{ B_{L_{0}}(u) \text{ is not } (E,m)\text{-SNS} \} \leq L^{-b}_{0}
$$

holds for some $L_0$ large enough, uniformly in $E \in I \subset \mathbb{R}$. Then for all $k \geq 0$ and all $E \in I$

$$
\sup_{u \in \mathbb{Z}^d} \mathbb{P}\{ B_{L_{k}}(u) \text{ is not } (E,m)\text{-SNS} \} \leq L^{-b}_{k}.
$$

**Remark 10.1.** I believe that the EVC bounds used in this section are very far from optimal. The reason is that no effort has been made above to prove a higher smoothness of the asymptotic PDF for the "frozen bath" EV concentration. Yet, as is well-known from the classical results in this direction, starting from the Chebyshev’s work in 1887 (cf. [17, 42, 10, 48, 139, 37, 9, 119]), the Berry–Esseen theorem (cf., e.g., [51, Section XVI.5], [64, Chapter 8]) stating for arbitrary distributions that the normalized sample PDF is uniformly approximated by the standard Gaussian one, can be substantially improved under the condition (called the Condition (C) by Cramér)

$$
\limsup_{|t| \to \infty} |\varphi_{\mu}(t)| < 1.
$$

(10.4)

There are also adaptations to non-identically distributed r.v., e.g. Berry’s theorem [10]. But it is possible that such improvements require geometry-specific arguments.

11. TWO-POINT CORRELATION MEASURES OF THE CUMULATIVE POTENTIAL

The analysis carried out in this section is not directly applied to the MSA schemes used in previous sections, so a number of minor technical details will be omitted and only sketches of proofs will be given. Here we shall make only a small step towards a functional CLT.
Consider the joint probability distribution of the values of the potential \( V \) at two arbitrary distinct lattice points; by translation invariance, we can shift one of them to \( 0 \in \mathbb{Z}_d \); let the other have the form \( \rho x \) with \( \rho > 0 \) and \( \|x\| = 1 \).

Next, consider the potential values induced at \(-\rho x\) and at \(\rho x\) by two scatterers at two opposite lattice points

\[ \pm ru, \ r > 0, \|u\| = 1, \ \text{with} \ (u, x) =: \cos \theta_{u,x} \neq 0, \ \theta_{u,x} \in [0, \pi) \setminus \left\{ \frac{\pi}{2} \right\}, \ \varsigma := \rho/r > 0. \]

The condition \( \cos \theta_{u,x} \neq 0 \) assumes asymmetry of \( \pm ru \) w.r.t. the sites \( \pm \rho x \) in the sense that one of the scatterers at \( \pm ru \) is closer to \( 0 \) while the other is closer to \( \rho x \). This is necessary for the non-degeneracy of the two-point distribution.

The general strategy is as follows:

- The goal is to establish continuity (which will in fact be absolute continuity) of the probability distribution of the random vector \((V(x), V(y)) \in \mathbb{R}^2\).
- Varying the scatterer’s position \( ru \) in a large annulus

\[ \{ z : \|z - 0\| \sim \|z - \rho x\| \in [c_1 r, c_2 r], r \gg 1 \} \]

and fixing the potential induced by all remaining scatterers, we can obtain a large sample of values of the vector \((V(x), V(y))\); with \( r \gg 1 \), they will concentrate around the vector produced by the fixed (i.e., conditioned) scatterers; for simplicity we simply ignore this “background” potential, which results in the shift of the expectation but does not affect adversely regularity of the joint distribution: convolution with the independent “background” potential can only enhance regularity.

- Since the active, non-conditioned random scatterer potentials have comparable amplitudes (this is why the annulus is introduced), their convolution should obey a CLT, i.e. have an asymptotically Gaussian distribution, albeit of small variance. For the respective covariance matrix to be non-degenerate, one needs at least three non-aligned values in \( \mathbb{R}^2 \), and to this end we consider pairs of asymmetric scatterers, providing in the Bernoulli case four non-aligned points in \( \mathbb{R}^2 \).

We focus on the Bernoulli case for simplicity, but it will be clear from the calculations that an extension to arbitrary nontrivial distributions would not require such a radical modification of the general approach as in [2] compared to the combinatorial argument by Bourgain and Kenig [13] based on Sperner’s lemma.

Note that

\[ \cos \theta_{-u,x} = -\cos \theta_{u,x} \]

For brevity we omit the subscript \( x \) and write \( \theta_u \). Let

\[ \zeta_0 = \omega_{ru}(0) + \omega_{-ru}V_{-ru}(0), \]
\[ \zeta_{\rho x} = \omega_{ru}V_{ru}(\rho x) + \omega_{-ru}V_{-ru}(\rho x), \]
Denote $a = A/2$, $c = \cos \theta_u$. Then

$$
\frac{V_{ru}(ru)}{V_{ru}(0)} = \frac{r^{-2a}\|u - \varepsilon x\|^{-2a}}{r^{-2a}} = \|u - \varepsilon x\|^{-a} = (1 - 2\varepsilon \cos \theta_u + \varepsilon^2)^{-a}
$$

$$
= 1 - a(-2c\varepsilon + \varepsilon^2) + \frac{a(a+1)}{2}(-2c\varepsilon + \varepsilon^2)^2 + O(\varepsilon^3)
$$

$$
= 1 + 2ac\varepsilon + a(-1 + 2(a+1)c^2)\varepsilon^2 + O(\varepsilon^3)
$$

$$
= 1 + \alpha\varepsilon + \beta_+\varepsilon^2,
$$

with $\alpha := 2ac = Ac$, $\beta_+ := (-1 + 2(a+1)c^2)$.

We have

$$(u, x) \mapsto -(u, x) \Rightarrow c_u \mapsto c_u, \ \alpha \mapsto \alpha,$$

$$(u, x) \mapsto \pm(u, -x) \Rightarrow c_u \mapsto -c_u, \ \alpha \mapsto -\alpha,$$

Thus neglecting the terms $O(\varepsilon^3)$ and denoting

$$\alpha := 2ac = Ac,$$

$$\beta := -1 + 2(a+1)c^2,$$

and using the symmetry $V_{ru}(0) = V_{-ru}(0)$, we obtain

$$
\frac{V_{ru}(\rho x)}{V_{ru}(0)} = 1 + \alpha\kappa + \beta\varepsilon^2
$$

$$
\frac{V_{-ru}(-\rho x)}{V_{ru}(0)} \equiv \frac{V_{-ru}(-\rho x)}{V_{ru}(0)} = 1 + \alpha\kappa + \beta\varepsilon^2
$$

$$
\frac{V_{ru}(-\rho x)}{V_{ru}(0)} = 1 - \alpha\kappa + \beta\varepsilon^2
$$

$$
\frac{V_{-ru}(\rho x)}{V_{ru}(0)} \equiv \frac{V_{-ru}(\rho x)}{V_{ru}(0)} = 1 - \alpha\kappa + \beta\varepsilon^2
$$

The quantity $|V_{ru}(0)| = |a_r|$ defines therefore the common scale in which fluctuations of order $O(\varepsilon)$ or $O(\varepsilon^2)$ occur. Let

$$
\zeta_+ = \omega_{-ru}V_{-ru}(+\rho x) + \omega_{ru}V_{ru}(+\rho x)
$$

$$
= a_r \left[ \omega_{-ru}(1 - \alpha\kappa + \beta\varepsilon^2) + \omega_{ru}(1 + \alpha\kappa + \beta\varepsilon^2) \right],
$$

$$
\zeta_- = \omega_{-ru}V_{-ru}(-\rho x) + \omega_{ru}V_{ru}(-\rho x)
$$

$$
= a_r \left[ \omega_{-ru}(1 + \alpha\kappa + \beta\varepsilon^2) + \omega_{ru}(1 - \alpha\kappa + \beta\varepsilon^2) \right]
$$

Then the $a_r$-scaled values of $\zeta_\pm$ determined by $\omega_{\pm ru}$ are as follows:
For the characteristic functional, under the same restriction for $\cos$, now calculate the covariance matrix:

\[
\begin{array}{c|c|c}
\omega_{-\omega} & \omega_{\omega} & \zeta_+ / |a_r| \\
\hline
+1 & +1 & 2(1 + \beta \varepsilon^2) =: a \\
+1 & -1 & 2\alpha \varepsilon =: b \\
-1 & +1 & -b \\
-1 & -1 & -2a \\
\end{array}
\]

Now calculate the covariance matrix:

\[
\begin{align*}
\mathbb{E} [\zeta^2] &= \frac{1}{4}(a^2 + b^2 + b^2 + a^2) = \frac{1}{2}(a^2 + b^2) \\
\mathbb{E} [\zeta^2] &= \frac{1}{2}(a^2 + b^2) \\
\mathbb{E} [\zeta_+ ] &= \frac{1}{4}(a^2 - b^2 - b^2 + a^2) = \frac{1}{2}(a^2 - b^2) \\
\end{align*}
\]

\[
\begin{align*}
C_\zeta &= \frac{1}{2} \begin{pmatrix} a^2 + b^2 & a^2 - b^2 \\ a^2 - b^2 & a^2 + b^2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\
\det C_\zeta &= a^2b^2 - b^4 = 4\alpha^2 \varepsilon^2 \cdot 4(1 + (2(a + 1)\varepsilon^2 - 1)\varepsilon^2) \\
&= 64\alpha^2 \varepsilon^2 \cdot (1 + O(\varepsilon^2))^2 \\
&= 16\alpha^2 \varepsilon^2 \cos^2 \theta_u \cdot (1 + O(\varepsilon^2)) > 0
\end{align*}
\]

for $\cos \theta_u \neq 0$ and $\varepsilon$ small enough. For example, with $|\cos \theta_u| \geq 1/\sqrt{2}$ and small $\varepsilon$,

\[
\det C_\zeta \geq 4\alpha^2 \varepsilon^2 > 0.
\]

For the characteristic functional, under the same restriction $|\theta_u| \leq \pi/4$ and with $a_r \|t\| \leq \frac{1}{\pi}$, $\varepsilon = \rho/r$, we get

\[
\left| \mathbb{E} \left[ e^{i\omega \cdot (t, \zeta)} \right] \right| \approx 1 - \frac{a_r^2}{2} C_\zeta(t, t) \leq 1 - 2a_r^2 A^2 \varepsilon^2 \|t\|^2 = 1 - \frac{2\rho^2}{r^{2A+2}} \|t\|^2.
\]

Qualitatively, we thus have regularity properties of the joint two-point probability distribution similar to those for their single-point counterparts: on spatial scale relatively large compared to $r^{-(A+2)/2} = r^{-\alpha - 1}$, it is, approximately, at least as regular as a mixture of Gaussian measures of variance of order $O(r^{-A-2})$, while on much smaller scales its discrete, singular nature cannot be neglected.

The above asymptotic formula for the characteristic functional implies of course regularity bounds for any linear functional of the random vector $(V(\rho x, \omega), V(\rho x, \omega))$, but one can derive the one for the difference $\eta := V(\rho x, \omega) - V(\rho x, \omega)$ directly from the previous calculations: with precision quadratic in $\varepsilon$,

\[
\eta \approx \left[ \omega_{\omega} V_{\omega}(\rho x) + \omega_{-\omega} V_{-\omega}(\rho x) \right] - \left[ \omega_{\omega} V_{\omega}(-\rho x) + \omega_{-\omega} V_{-\omega}(-\rho x) \right]
= \omega_{\omega} \left[ V_{\omega}(\rho x) - V_{\omega}(-\rho x) \right] + \omega_{-\omega} \left[ V_{-\omega}(\rho x) - V_{-\omega}(-\rho x) \right]
= \omega_{\omega} [(1 + \alpha \varepsilon + \beta \varepsilon^2) - (1 - \alpha \varepsilon + \beta \varepsilon^2)] + \omega_{-\omega} [(1 - \alpha \varepsilon + \beta \varepsilon^2) - (1 + \alpha \varepsilon + \beta \varepsilon^2)]
= 2\alpha \varepsilon (\omega_{\omega} - \omega_{-\omega})
\]
whence, again with quadratic precision in \( \kappa \),

\[
\varphi_\eta(t) = E \left[ e^{i a_r \eta t} \right] \approx E \left[ 1 + i a_r \eta t - \frac{1}{2} a_r^2 \eta^2 t^2 \right] = 1 - 2 \cdot \frac{1}{4} \cdot \frac{1}{2} a_r^2 \kappa^2 \eta^2 t^2 = 1 - \frac{A^2 c^2 t^2}{r^{2A+2}}
\]

With \( |\theta_u| \leq \pi/4 \), hence \( c^2 = \cos^2 \theta_u \geq 2 \), \( \kappa = \rho / r \leq \frac{1}{2} \),

\[
- \ln |\varphi_\eta(t)| \geq - \ln \left| 1 - \frac{A^2 t^2}{2r^{2A+2}} \right| \geq \frac{A^2}{4} \left( \frac{t}{r^{A+1}} \right)^2
\]

This is still good enough for some satisfactory lower bounds in probability for the spectral spacings at an initial scale \( L_0 \), in a ball \( B_{L_0}(u) \), where in the strong disorder regime the EVs are essentially given by the values of the effective random potential \( V(\cdot, \omega) \).

12. Concluding remarks

12.1. On the proof of Theorem 3.4. In Theorem 3.4 we have a situation interpolating between exponential and polynomial decay of interaction \( u \). While the lack of a genuine exponential decay prevents one from using the same arguments as for Theorem 3.3 and establish exponential scaling limit, we are at the same time much better off with fractional-exponential decay than with polynomial one. Without the need for an elaborate scaling scheme, it is easy to adapt the one used for the polynomial decay and replace power-law probability bounds with their fractional-exponential analogs. The adaptation is fairly routine and omitted from this paper.

12.2. Random dipoles and random displacements. The simplest (binary) random displacement model, with two possible scatterer’s positions per cell, is close in spirit to a random dipole model: moving the source from position \( a_1 \) to position \( a_2 \) does not change the total charge but changes the orientation of a “dipole”. At a remote target point \( x \), not located on the median hyperplane for \( a_1 \) and \( a_2 \), the fluctuation of the registered potential is due to the non-flatness of the potential amplitude \( y \mapsto u(|y - x|) \), and essentially equivalent to the variation of the potential at either of the source points \( a_1, a_2 \), so we are still in the general framework of a Bernoulli disorder.

Naturally, any finite number of admissible source locations per cell gives rise to a similar situation, and any continuity of the probability distribution of the source points is very welcome for the regularity of the IDS and for localization.

12.3. Non-homogeneous models of disorder. Such models (Delone–Anderson Hamiltonians [63, 111], crooked/trimmed Hamiltonians [85, 43]) have been quite popular in the last few years. The lack of ergodicity is often the major technical problem in such models. Whenever the particle–media interaction potential has infinite range, particularly in the case of slowly decaying functions (in the same sense as above), tempered non-homogeneity has but a weak effect on the statistical properties of the mapping \( q \mapsto U[q](\cdot) \). The Gaussian micro-scaling remains valid (the attraction class remains the same) unless one is allowed to put arbitrarily large number of sources per unit volume (the same remark concerns the random displacements models).
12.4. Random magnetic fields. There is a vast literature on local random magnetic fields; a review can be found in [18], but here I would like to point out a paper by Erdős and Hasler [46] where an interesting technique was proposed to efficiently control the contribution of random fluctuations of magnetic fields to the local energy levels. Taking a full account of magnetic fluctuations from remote sources seems an interesting problem. Shielding from electromagnetic fields by thin metal films, and even by a tight grid (often preferred for its weight and flexibility) is a well-known and widely used technique (that’s why people buy these days insulated wallets for contactless cards), but, firstly, the penetration ”skin” still has a finite width; secondly, it has been demonstrated above that ”exponentially small” \( \neq ”zero”; \) and thirdly, the surface layer is exposed to external long-range fields. The final word here belongs to physicists.

12.5. Correlated sources and statistical mechanics. The random amplitudes \( \omega_x \) have been assumed independent to make the probabilistic analysis simpler. More realistically, they should be considered in the framework of a DLR (Dobrushin–Lanford–Ruelle) measure, hence correlated in general. Recall in this connexion that von Dreifus and Klein [40] pointed out this class of models and explained that a strong form of decay of correlation, emerging in spin models from the Dobrushin–Shlosman complete analyticity conditions, leads to Anderson localization in correlated potentials. However, it appears that in 1980s the complete analyticity could be established only in the models with a finite spin space [120], hence with a discrete local disorder, while the MSA on a lattice requires at least log-Hölder continuity of finite-volume EV distributions. It seems to be a natural further step to explore the regularity of the DoS for the Gibbsian random fields \( \{ \omega_x, x \in \mathbb{Z}^d \} \).

12.6. On higher smoothness in the ”frozen bath” approximations. The asymptotical smoothness of probability distributions of properly normalized sums of random variables, \( M_n = n^{-1} \sum_{i=1}^n X_i \), has been an area of active research in probability theory since very long time. Asymptotic expansions of the PDF and, where appropriate, probability density using the Chebyshev–Hermite polynomials go back to the 19th century (cf. Chebyshev [17]). In early 20th century this direction was further developed by Edgeworth [42], Bruns [14], Charlier [16]. Edgeworth expansions and their counterparts for the quantiles (so-called Cornish-Fisher expansions [33, 34]) are frequently used in statistics and in risk management. A good introduction can be found in the books by Cramér [37], Feller [51], and Gnedenko and Kolmogorov [64]. A classical condition allowing one to achieve a higher accuracy of approximation of a sample distribution by the Gaussian law, one of several bearing the name of H. Cramér [37] (usually called condition (C)), is that

\[
\limsup_{|t| \to \infty} |\hat{\varphi}_V(t)| < 1.
\]

The analysis in Sections 5–6 evidences that any uniform decay at infinity of the Fourier transform \( \hat{\varphi}_V \) of the marginal measure \( \mu_V \) would be even more welcome. However, it is known that such conditions require a fair amount of continuity on the part of \( \mu_V \). A milder condition is that \( \mu_V \) is not supported by any affine sub-lattice \( \mathcal{A} = T\mathbb{Z} + a, T > 0; \) the respective measures are called ”non-lattice” distributions, and many asymptotic results for the large sums of IID r.v. are proved either for the lattice or for non-lattice distributions.
Observe that the non-lattice condition is already fulfilled for a measure supported by three rationally incommensurate points \( \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R} \).

In any event, the wealth of knowledge accumulated in this area of probability and statistics may shed some light on accurate, asymptotically sharp results on regularity of the DoS in finite volumes subject to the non-local fluctuations occurring in the ambient thermal bath, which transgress the usual limits of the boundary conditions for local finite-difference or differential operators, or even on regularity of the DoS in a finite volume with the thermal bath temporarily "frozen" for some technical reasons (broken heater). Although we focused here on the results one can establish even for the most singular form of marginal disorder (Bernoulli), in a unified way and without reduction to the Bernoulli case, it is certainly worth investigating quantitatively, to what extent a discrete, e.g. substitution disorder is enhanced in realistic systems by additional, even relatively weak in amplitude, sources of very "shy", just barely continuous disorder.

12.7. **Thin tails from neighbors.** In an old, well-known story a boy asks his father:

- **Dad, a cubic meter of wood, is it much?**
- **To chop, quite a lot. To heat the house ... not really.**

Likewise, a question "Is a two-channel wire “quasi”-one-dimensional or pretty much two-dimensional?" is not philosophical but quantitative, to be asked for each particular problem and application. In this paper, I focused on the qualitative regularity properties of the cumulative potential induced by the most singular site-wise disorder, through multiple linear convolutions, as well as on its impact on the DoS measure. It has been demonstrated that there is a fairly universal and strong tendency to an extreme form of regularization – infinite derivability, and the main mechanism may only break down in *bona fide* 1D systems subject to the strongest – exponential – form of screening. From this particular point of view, even a nanotube, composed of several chains of atoms, may or may not qualify as "sufficiently one-dimensional". Add to this the requirement of exponential screening, and we probably end up with a purely mathematical curiosity; being not a physicist myself, I cannot be sure such systems exist at all. If they do, it would be very interesting indeed; if not, this would give to the regularization mechanism studied here an even greater universality. The Wegner estimate \[ 133 \], put in simple terms, says the regularity of a short-range IID disorder is preserved in the IDS measure. In the particular case of a bounded probability density \( \rho \) of the disorder, the DoS itself is bounded exactly by \( \| \rho \|_{\infty} \). Mathematical works extended this to arbitrary continuous measures \( \mu \) of the IID disorder: the IDS has (up to some constants) the same continuity modulus \( \mathfrak{s}_{\mu} \) as \( \mu \). We have seen that non-local interactions significantly improve the regularity of the underlying disorder, even when the latter is extremely singular. In turn, this gives rise to the "thin" tails at band edges, complementing the usual Lifshitz tails phenomenon.

At the very least, our analysis provides a possible answer to a fairly natural question: just **where does a regular disorder come from?** To a very efficient and elegant approach to Anderson localization developed by Aizenman and Molchanov \[ 3 \], local regularity after conditioning on the complement of a finite ball is simply vital. In this regard, the present paper, unfortunately, does not shed any light on the main hypothesis of the FMM
approach (not yet, anyway), while the MSA is much better off with the regularity coming from a discrete (e.g., substitution) disorder via non-local smoothing, as seen in Section 9.

However, dismissing regular underlying disorder is apparently unwarranted, since some additional regularity may come from the Gibbs distribution on the configurations of ”scattersers”. Note in this connection that in the case where the marginal distribution of the scatterer amplitudes is itself Lipschitz continuous, for whatever reason, and compactly supported, the convolution effect on the edge decay of probability density for the cumulative potential $V$ is much more immediate. For example, suppose $\omega \sim \text{Unif}[0, 1]$, and $u$ has range $\sqrt{d}$, so that only the nearest cubic ”sphere” surrounding a site $x$ affects $V(x; \omega)$. Apparently, one has to have a very fertile imagination even to look for a strong screening mechanism between nearest neighbors. We still have 8 neighbors in $\mathbb{Z}^2$, so the cumulative potential has $C^6$-density and edge decay $O(|E - E_*|^6)$. Respectively, in $\mathbb{Z}^3$ each site has 26 neighbors and $C^{24}$-density. Also, in a double-layer, quasi-2D sample each site has 17 neighbors, which shows that the dimensionality parameter figuring in various estimates is to be determined carefully. In practice, genuinely 2-dimensional, mono-layer samples are rare, as well as truly single-channel linear chains. Needless to say, the screening effects do not necessarily ”switch on” sharply beyond one or two atomic distances, so the role and universality of ”thin” tails cannot be just discarded in many applications, even if for some reasons one decided to neglect strongly (say, exponentially) screened potentials at large distances.

It seems interesting to explore possible local effects of weak screening at moderate distances undergoing a cross-over to a much stronger one beyond some typical radius. The paradigm of an infinite media brought to life a number of deep mathematical results and techniques, but the recent wake of interest in physics and technology to microscopical systems suggests one should not neglect such models either; chances are this preprint is visualized by the reader on a quantum dot based screen.

Acknowledgements

It is a pleasure to thank Misha Goldshtein, Günter Stolz and Ivan Veselić for stimulating discussions of localization mechanisms in correlated potentials, and David Khmelnitskii for fruitful discussions of physical mechanisms of screening phenomena.

Special thanks to the organizers of the semi-annual programs Mathematics and Physics of Anderson Localization: 50 Years After (2008) and Periodic and Ergodic Spectral Problems (PEP, 2015) held at the Isaac Newton Institute, Cambridge, UK, and to the INI team for the warm hospitality and opportunity to work in the unique, stimulating atmosphere of the Institute. The project outlined here has been initiated during the PEP program.

Resources of the Betty and Gordon Moore Library, of the Cambridge University, have been invaluable for the present project as well as for the one on multi-particle Anderson localization, initiated by Yuri Suhov and myself in 2003, also at the INI. I am greatly indebted to the founders and generous sponsors of the Library.
REFERENCES

[1] Anderson, P.: Absence of diffusion in certain random lattices. Phys. Rev. 109(5), 1492–1505 (1958)
[2] Aizenman, M., Germinet, F., Klein, A., Warzel, S.: On Bernoulli decompositions for random variables, concentration bounds, and spectral localization. Prob. Theory Rel. Fields 143(1), 219–238 (2009)
[3] Aizenman, M., Molchanov, S.A.: Localization at large disorder and at extreme energies: an elementary derivation. Commun. Math. Phys. 157(2), 245–278 (1993)
[4] Aizenman, M., Warzel, S.: Localization bounds for multiparticle systems. Commun. Math. Phys. 290, 903–934 (2009)
[5] Aizenman, M., Warzel, S.: Complete dynamical localization in disordered quantum multi-particle systems. In: XVIth International Congress on Mathematical Physics, pp. 556–565. World Sci. (2010)
[6] Amrein, W.O., Georgescu, V.: On the characterization of bound states and scattering states in quantum mechanics. Helv. Phys. Acta 46(10), 635–658 (1973)
[7] Bary, N.K.: Sur l’unicité du développement trigonométrique. Fund. Math. 9, 62–118 (1927)
[8] Basko, D.M., Aleiner, I.L., Altshuler, B.L.: Metal-insulator transition in a weakly interacting many-electron system with localized single-particle states. Ann. Phys. 321(5), 1126–1205 (2006)
[9] van Beek, P.: An application of Fourier methods to the problem of sharpening the Berry–Esseen inequality. Z. Wahrscheinlichkeitstheorie verw. Geb. 23, 187–196 (1972)
[10] Berry, A.C.: The accuracy of the Gaussian approximation to the sum of independent random variates. Trans. Amer. Math. Soc. 49, 122–136 (1941)
[11] Birman, M.S.: On the spectrum of singular boundary-value problems. Mat. Sbornik (N.S.) 55, 125–174 (1961)
[12] Boole, G.: On the comparison of transcendentals, with certain applications to the theory of definite integrals. Phil. Trans. of the Royal Society of London 147, 745–803 (1857)
[13] Bourgain, J., Kenig, W.: On localization in the continuous Anderson–Bernoulli model in higher dimension. Invent. Math. 161, 389–426 (2005)
[14] Bruns, H.: Wahrscheinlichkeitsrechnung und Kollektivmasslehre. Druck und Verlag von B. G. Teubner, Leipzig Berlin, Leipzig Berlin (1906)
[15] Carmona, R., Klein, A., Martinelli, F.: Anderson localization for Bernoulli and other singular potentials. Commun. Math. Phys. 108, 41–66 (1987)
[16] Charlery, C.V.L.: Über das Fehlergesetz. Ark. Math. Astr. och Phys. 2(8), 1–9 (1905)
[17] Chebyshev, P.L.: Sur le developpement des fonctions a une seule variable. Bull. de la Classe phys.-math. de l’Acad. Imp. des Sc. de St. Petersburg 1, 193–200 (1887)
[18] Chulaevsky, V.: Renormalization group limit of Anderson models. submitted
[19] Chulaevsky, V.: Direct scaling analysis of localization in single-particle quantum systems on graphs with diagonal disorder. Math. Phys. Anal. Geom. 15(4), 361–399 (2012)
[20] Chulaevsky, V.: Stochastic regularization and eigenvalue concentration bounds for singular ensembles of random operators. J. Stat. Mech. 2013, 931063 (2013)
[21] Chulaevsky, V.: From fixed-energy localization analysis to dynamical localization: An elementary path. J. Stat. Phys. 154, 1391–1429 (2014)
[22] Chulaevsky, V.: Efficient localization bounds in a continuous N-particle Anderson model with long-range interaction. Lett. Math. Phys. 106(4), 509–533 (2016)
[23] Chulaevsky, V.: Exponential scaling limit for single-particle Anderson models via adaptive feedback scaling. J. Stat. Phys. 162, 603–614 (2016)
[24] Chulaevsky, V., Boutet de Monvel, A., Suhov, Y.: Dynamical localization for a multi-particle model with an alloy-type external random potential. Nonlinearity 24(5), 1451–1472 (2011)
[25] Chulaevsky, V., Spencer, T.: Positive Lyapunov exponents for a class of deterministic potentials. Commun. Math. Phys. 168, 455–466 (1995)
[26] Chulaevsky, V., Suhov, Y.: Efficient Anderson localization bounds for large multi-particle systems. To appear in J. Spec. Theory
[27] Chulaevsky, V., Suhov, Y.: Wegner bounds for a two-particle tight binding model. Commun. Math. Phys. 283, 479–489 (2008)
[28] Chulaevsky, V., Suhov, Y.: Eigenfunctions in a two-particle Anderson tight binding model. Commun. Math. Phys. 289, 701–723 (2009)
[29] Chulaevsky, V., Suhov, Y.: Multi-particle Anderson localisation: Induction on the number of particles. Math. Phys. Anal. Geom. 12, 117–139 (2009)
[30] Chulaevsky, V., Suhov, Y.: Multi-scale analysis for random quantum systems with interaction, Progress in Mathematical Physics, vol. 65. Birkhäuser, Boston (2013)
[31] Combes, J.M., Hislop, P.D., Klopp, F.: An optimal Wegner estimate and its application to the global continuity of the integrated density of states for random Schrödinger operators. Duke Math. J. 140(3), 469–498 (2007)
[32] Cornfeld, I.P., Fomin, S.V., Sinai, Y.G.: Ergodic theory, Grundlehren der mathematischen Wissenschaften, vol. 245. Springer-Verlag, New York Heidelberg Berlin (1982)
[33] Cornish, E.A., Fisher, R.A.: Moments and cumulants in the specification of distributions. Rev. l’Inst. Nat. de Stat. 5(4), 307–320 (1938)
[34] Cornish, E.A., Fisher, R.A.: The percentile points of distributions having known cumulants. Technometrics 2(2), 209–225 (1960)
[35] Craig, W., Simon, B.: Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices. Commun. Math. Phys. 90, 207–218 (1983)
[36] Craig, W., Simon, B.: Subharmonicity of the Lyaponov index. Duke Math. J. 50, 551–560 (1983)
[37] Cramér, H.: Random variables and probability distributions, Cambridge tracts in mathematics and mathematical physics, vol. 36. Cambridge University Press (1970)
[38] von Dreifus, H.: On the effect of randomness in ferromagnetic models and Schrödinger operators. PhD thesis, New York University (1987)
[39] von Dreifus, H., Klein, A.: A new proof of localization in the Anderson tight binding model. Commun. Math. Phys. 124, 285–299 (1989)
[40] von Dreifus, H., Klein, A.: Localization for random Schrödinger operators with correlated potentials. Commun. Math. Phys. 140, 133–147 (1991)
[41] Dubickas, A.: On the fractional parts of the natural powers of a fixed number. Siber. Math. J. 47(5), 879–882 (2006)
[42] Edgeworth, F.Y.: The law of Error. Cambridge Philos. Soc. 20, 36–66, 113–141 (1905)
[43] Elgart, A., Klein, A.: Ground state energy of trimmed discrete Schrödinger operators and localization for trimmed Anderson models. J. Spec. Theory 4(2), 391–413 (2014)
[44] Elgart, A., Tautenhahn, M., Veselić, I.: Localization via fractional moments for models on $\mathbb{Z}$ with single-site potentials of finite support. J. Phys. A 43(8), 474,021 (2010)
[45] Enss, V.: Asymptotic completeness for quantum mechanical potential scattering in short range potentials. Commun. Math. Phys. 61, 285–291 (1978)
[46] Erdös, L., Hasler, D.: Wegner estimate and Anderson localization for random magnetic fields. Commun. Math. Phys. 309, 507–542 (2012)
[47] Erdös, P.: On a family of symmetric Bernoulli convolutions. Amer. J. Math. 61, 974–976 (1939)
[48] Esseen, C.G.: On the Liapounoff limit of error in the theory of probability. Ark. Mat. Astr. Fys. 28A, 1–19 (1942)
[49] Exner, P., Helm, M., Stollmann, P.: Localization on a quantum graph with a random potential on the edges. Rev. Math. Phys. 19(9), 923–939 (2007)
[50] Fauser, M., Warzel, S.: Multiparticle localization for disordered systems on continuous space via the fractional moment method. Rev. Math. Phys. 27(4), 1550010 (2015)
[51] Feller, W.: An introduction to probability theory and its applications, vol. 2. John Wiley and Sons, Inc., New York London Sidney (1966)
[52] Fischer, W., Hupfer, T., Leschke, H., Müller, P.: Existence of the density of states for multi-dimensional continuum Schrödinger operators with Gaussian random potentials. Commun. Math. Phys. 190, 133–141 (1997)
[53] Fischer, W., Leschke, H., Müller, P.: Spectral localization by Gaussian random potentials in multidimensional continuous space. J. Stat. Phys. 101(5), 935–985 (2000)

[54] Franklin, J.N.: Deterministic simulation of random processes. Math. Comput. 17, 28–29 (1963)

[55] Friedel, J.: The distribution of electrons round impurities in monovalent metals. Phil. Mag. 43(337), 153–189 (1952)

[56] Friedel, J.: Metallic alloys. Nuovo Cimento 7(2), 287–311 (1958)

[57] Fröhlich, J., Martinelli, F., Scoppola, E., Spencer, T.: Constructive proof of localization in the Anderson tight binding model. Commun. Math. Phys. 101, 21–46 (1985)

[58] Fröhlich, J., Spencer, T.: Absence of diffusion in the Anderson tight binding model for large disorder or low energy. Commun. Math. Phys. 88, 151–184 (1983)

[59] Gabovich, A.M., Il’chenko, L.G., Pashitskii, E.A., Romanov, Y.A.: Screening of charges and Friedel oscillations of the electron density in metals having differently shaped Fermi surfaces. JETP 48(1), 249–264 (1978)

[60] Germinet, F., Klein, A.: Bootstrap multi-scale analysis and localization in random media. Commun. Math. Phys. 222, 415–448 (2001)

[61] Germinet, F., Klein, A.: Explicit finite volume criteria for localization in continuous random media and applications. GAFA, Geom. Funct. Anal. 13, 1201–1238 (2003)

[62] Germinet, F., Klein, A.: A comprehensive proof of localization for continuous Anderson model with singular random potentials. J. Spec. Theory 15(1), 55–143 (2013)

[63] Germinet, F., Müller, P., Rojas-Molina, C.: Ergodicity and dynamical localization for Delone–Anderson operators. Rev. Math. Phys. 27(9), 1550020 (2015)

[64] Gnedenko, B.V., Kolmogorov, A.N.: Limit distributions for sums of independent random variables. Addison–Wesley, Cambridge, Cambridge (1954)

[65] Goldsheid, I., Molchanov, S.A., Pastur, L.A.: A pure point spectrum of the stochastic one-dimensional Schrödinger operator. Funct. Anal. Appl. 11, 1–8 (1977)

[66] Gornyi, I.V., Mirlin, A.D., Polyakov, D.G.: Interacting electrons in disordered wires: Anderson localization and low-T transport. Phys. Rev. Letters 95, 206603 (2005)

[67] Halperin, B.: Properties of a particle in a one-dimensional random potential. Adv. Chem. Phys. 31, 123–177 (1967)

[68] Hardy, G.H., Littlewood, J.E.: Some problems of Diophantine approximation. Acta Math. 37, 155–190, 193–238 (1914)

[69] Hardy, G.H., Littlewood, J.E.: Some problems of Diophantine approximation: A remarkable trigonometric series. Proc. Nat.Acad. Sci. USA 3(2), 583–586 (1916)

[70] Hewet, E., Zuckerman, H.S.: Singular measures with absolutely continuous convolution squares. Proc. Cambridge Philos. Soc. 62, 399–420 (1966)

[71] Hewet, E., Zuckerman, H.S.: Corrections to the paper ”singular measures with absolutely continuous convolution squares”. Proc. Cambridge Philos. Soc. 63, 367–368 (1967)

[72] Hu, T.Y., Lau, K.S.: Fourier asymptotics of Cantor type measures at infinity. Proc. Amer. Math. Soc. 130(9), 2711–2717 (2002)

[73] Hutchinson, J.E.: Fractals ans self-similarity. Indiana Univ. Math. J. 30(5), 713–747 (1981)

[74] Ivašev-Musatov: On the Fourier–Stiltjes coefficients of singular functions. Dokl. Akad. Nauk SSSR pp. 9–11 (1952)

[75] Ivašev-Musatov: On the coefficients of trigonometric null series. Izv. Akad. Nauk SSSR pp. 559–578 (1957)

[76] Jessen, B., Wintner, A.: Distribution function and the Riemann zeta function. Trans. Amer. Math. Soc. 38, 48–88 (1935)

[77] Kac, M.: Statistical independence in probability, analysis and number theory. The Mathematical Association of America, John Wiley and Sons (1959)

[78] Kac, M., Salem, R., Zygmund, A.: A gap theorem. Trans. Amer. Math. Soc. 63, 235–243 (1947)

[79] Kahane, J.P., Salem, R.: Sur la convolution d’une infinité de distributions de Bernoulli. Coll. Math. 6, 193–202 (1958)
[80] Karanikas, C., Koumandos, S.: Continuous singular measures with absolutely continuous convolution squares on locally compact groups. Illinois J. Math. 35(3), 490–495 (1991)
[81] Karsten, L., Peyerimhoff, N., Tautenhahn, M., Veselić, I.: Wegner estimate and localization for alloy-type models with sign-changing exponentially decaying single-site potentials. Rev. Math. Phys. 27, 1550007 (2015)
[82] Khintchine, A.Y., Kolmogorov, A.N.: Über Konvergenz von Reihen, deren Glieder durch den Zufall bestimmt werden. Mat. Sbornik 32, 668–677 (1925)
[83] Kirsch, W., Stollmann, P., Stolz, G.: Anderson localization for random Schrödinger operators with long range interactions. Commun. Math. Phys. 195, 495–507 (1998)
[84] Klein, A.: Unique continuation principle for spectral projections of Schrödinger operators and optimal Wegner estimates for non-ergodic random Schrödinger operators. Commun. Math. Phys. 323, 1229–1246 (2013)
[85] Klein, A., Martinelli, F., Perez, M.: A rigorous replica trick approach to Anderson localization in one dimension. Commun. Math. Phys. 106, 623–633–40 (1986)
[86] Kohn, W., Vosko, S.H.: Theory of nuclear resonance intensity in dilute alloys. Phys. Rev. 119, 912–936 (1960)
[87] Koksma, J.F.: Ein mengentheoretischer Satz über die Gleichverteilung modulo Eins. Compositio Mathematica 2, 250–258 (1935)
[88] Kolmogorov, A.: Über die Summen durch den Zufall bestimmt zufälliger Grössen. Math. Annalen 99, 309–319 (1928)
[89] Kolmogorov, A.N.: Sur les propriétés des fonctions de concentration de M. P. Lévy. Ann. Inst. H. Poincaré 16, 27–34 (1958)
[90] Kotani, S., Simon, B.: Localization in general one-dimensional random systems. II. Continuum Schrödinger operators. Commun. Math. Phys. 112, 103–119 (1987)
[91] Krüger, H.: Localization for random operators with non-monotone potentials with exponentially decaying correlations. Ann. Henri Poincaré 13(3), 543–598 (2012)
[92] Kunz, H., Souillard, B.: Sur le spectre des opérateurs aux différences finies aléatoires. Commun. Math. Phys. 78, 201–246 (1980)
[93] Langer, J.S., Vosko, S.H.: The shielding of a fixed charge in a high-density electron gas. J. Phys. Chem. Solids 12, 196–205 (1959)
[94] Levin, M.B.: Completely uniform distribution of fractional parts of the exponential function. J. Soviet Math. 31(5), 3247–3256 (1985)
[95] Lévy, P.: Sur les séries dont les termes sont des variables éventuelles indépendantes. Studia Mathematica 3, 119–155 (1931)
[96] Lifshitz, I.M., Gredescul, S.A., Pastur, L.A.: Introduction to the theory of disordered systems. Wiley, New York (1988)
[97] Lyons, R.: Seventy years of Rajchman measures. J. Fourier Anal. Appl. Kahane special issue, 363–377 (1995)
[98] Mahler, K.: On the fractional parts of the powers of a rational number. Math. Annalen 107, 691–730 (1932)
[99] Menchoff, D.E.: Sur l’unicité du développement trigonométrique. C. R. Acad. Sc. Paris, Sér. A-B 163, 433–436. (1916)
[100] Neder, L.: Über die Fourierkoeffizienten der Funktionen von beschränkter Schwankung. Math. Zeit. 6, 270–273 (1920)
[101] Pastur, L.A.: Behavior of some Wiener integrals as $t \to \infty$ and the density of states of Schrödinger equations with random potential. Theor. Math. Phys. 32(1) (1977)
[102] Pastur, L.A.: Spectral properties of disordered systems in one-body approximation. Commun. Math. Phys. 75, 179–196 (1980)
[103] Pastur, L.A., Figotin, A.L.: Spectra of random and almost-periodic operators. Springer, Berlin (1992)
[105] Petrashov, V.T., Antonov, V.N., Nilsson, B.: Electric field effects and screening in mesoscopic bismuth wires. J. Phys.: Condens. Matter 3, 9705–9711 (1991)
[106] Rajchman, A.: Sur l’unicité du développement trigonométrique. Fund. Math. 3, 287–302 (1922)
[107] Rajchman, A.: Sur une classe de fonctions à variation bornée. C. R. Acad. Sci. Paris 187, 1026–1028 (1928)
[108] Reed, M., Simon, B.: Methods of modern mathematical physics, vol. 3. Academic Press, Inc. (1979)
[109] Riesz, F.: Über die Fourierreihen und ihre harmonische Analyse. Math. Zeit. 2, 312–315 (1918)
[110] Rojas-Molina, C.: Characterization of the Anderson metal-insulator transition for non ergodic operators and application. Ann. H. Poincaré 13, 1575–1611 (2012)
[111] Rojas-Molina, C., Veselić, I.: Scale–free unique continuation estimates and applications to random Schrödinger operators. Commun. Math. Phys. 320, 245–274 (2013)
[112] Rudnick, Z., Sarnak, P., Zaharescu, A.: The distribution of spacings between the fractional parts of \( n/\alpha \). Invent. Math. 145, 37–57 (2001)
[113] Ruelle, D.: A remark on bound states in potential-scattering theory. Nuovo Cimento A 61, 655–662 (1969)
[114] Sabri, M.: Some abstract Wegner estimates with applications. Lett. Math. Phys. 104, 311–339 (2014)
[115] Sæki, S.: Singular measures with absolutely continuous convolution powers. Illinois J. Math. 21(2), 395–412 (1977)
[116] Salem, R.: Sets of uniqueness and sets of multiplicity. Trans. Amer. Math. Soc. 54(2), 218–228 (1943)
[117] Schulz-Baldes, B., Seidel, C.: Positive Lyapunov exponents and localization bounds for strongly mixing potentials. Adv. Theor. Math. Phys. 12, 1377–1399 (2008)
[118] Schwinger, J.: On the bound states of a given potential. Proc. Natl. Acad. Sci. USA 47, 122–129 (1961)
[119] Shevtsova, I.G.: On the absolute constants in the Berry–Esseen-type inequalities. Doklady Math. 89(3), 378–381 (2014)
[120] Shlosman, S.B.: private communication
[121] Simon, B., Taylor, M.: Harmonic analysis on SL(2,R) and smoothness of the density of states in the one-dimensional Anderson model. Commun. Math. Phys. 101, 1–19 (1985)
[122] Simon, B., Wolf, T.: Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians. Commun. Pure Appl. Math 39, 75–90 (1986)
[123] Spencer, T.: Localization for random and quasi-periodic potentials. J. Stat. Phys. 51, 1009–1019 (1988)
[124] Stollmann, P.: From uncertainty principles to Wegner estimates. Math. Phys. Anal. Geom. 13, 145–157 (2010)
[125] Stout, W.F.: Some results on the complete and almost sure convergence of linear combinations of independent random variables and martingale differences. Ann. Math. Stat. 39, 1549–1562 (1968)
[126] Stout, W.F.: Almost sure convergence. Academic Press, New York (1974)
[127] Strichartz, R.S.: Fourier asymptotics of fractal measures. J. Funct. Anal. 89, 154–187 (1990)
[128] Strichartz, R.S.: Self-similar measures and their Fourier transforms I. Indiana Univ. Math. J. 39(3), 797–817 (1990)
[129] Strichartz, R.S.: Mock Fourier series and transforms associated with certain Cantor measures. J. d’Analyse Math. 81, 209–238 (2000)
[130] Temple, G.: The theory of Rayleigh’s principle as applied to continuous systems. Proc. Roy. Soc. London Ser. A 119, 276–293 (1928)
[131] Veselić, I.: Wegner estimates for sign-changing single site potentials. Math. Phys. Anal. Geom. 13, 299–313 (2010)
[132] Veselić, I.: Minami’s estimate: Beyond rank one perturbation and monotonicity. Ann. Henri Poincaré 15, 737–754 (2014)
[133] Wegner, F.: Bounds on the density of states in disordered systems. Z. Phys. B. Condensed Matter 44, 9–15 (1981)
[134] Weyl, H.: Über die Gleichverteilung der Zahlen mod. Eins. Math. Ann. 77, 313–352 (1916)
[135] Wiener, N.: The quadratic variation of a function and its Fourier coefficients. J. of Math. and Phys. 3, 72–94 (1924)
[136] Wiener, N., Wintner, A.: Fourier–Stiltjes transforms and singular infinite convolutions. Amer. J. of Math. 60, 513–522 (1938)
[137] Wintner, A.: On analytic convolutions of Bernoulli distributions. Amer. J. Math. 56, 659–663 (1934)
[138] Wintner, A.: On symmetric Bernoulli convolutions. Bull. Amer. Math. Soc. 41, 137–138 (1935)
[139] Zolotarev, V.: An absolute estimate of the remainder term in the central limit theorem. Theor. Probab. Appl. 11(4), 95–105 (1966)

Université de Reims, Département de mathématiques, 51687 Reims Cedex, France
E-mail address: victor.tchoulaevski@univ-reims.fr