

# Optimized Regularity Estimates of Conditional Distribution of the Sample Mean

Victor Chulaevsky

Department of Mathematics, University of Reims Champagne-Ardenne, B.P. 1039 51687 Reims Cedex 2, France

Copyright ©2015 Horizon Research Publishing All rights reserved.

**Abstract** We prove an optimized estimate for the regularity of the conditional distribution of the empiric mean of a finite sample of IID random variables, conditional on the sample "fluctuations". Prior results, based on bounds in probability, provided a Hölder-type regularity of the conditional distribution. We establish a Lipschitz regularity, using bounds in expectation. The new estimate, extending a well-known property of Gaussian IID samples, is a crucial ingredient of the Multi-Scale Analysis of multi-particle Anderson-type random Hamiltonians in a Euclidean space. In particular, the Hölder regularity of the multi-particle eigenvalue distribution, sufficient for the localization analysis of N-particle lattice Hamiltonians, with  $N \geq 3$ , needs to be replaced by Lipschitz regularity for similar Hamiltonians in the Euclidean space.

**Keywords** Multi-particle Anderson Localization, Eigenvalue Concentration Estimates

## 1 Introduction

### 1.1 The general probabilistic framework

Consider a sample of  $n > 1$  IID (independent and identically distributed) random variables  $X_1, \dots, X_n$  with Gaussian distribution  $\mathcal{N}(0, 1)$ , and introduce the sample mean  $\xi = \xi_n$  and the "fluctuations"  $\eta_i$  relative to the mean:

$$\xi_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad \eta_i = X_i - \xi_n, \quad i = 1, \dots, n.$$

It is well-known from standard courses of the probability theory (cf., e.g., [11]) that  $\xi_n$  is independent from the sigma-algebra  $\mathfrak{F}_\eta$  generated by  $\{\eta_1, \dots, \eta_n\}$  (the latter are linearly dependent and have rank  $n - 1$ ). To see this, it suffices to note that all  $\eta_i$  are orthogonal to  $\xi_n$  with respect to the standard scalar product in the linear space formed by  $X_1, \dots, X_n$ , given by

$$\langle Y, Z \rangle := \mathbb{E}[Y Z],$$

where  $Y$  and  $Z$  are real linear combinations of  $X_1, \dots, X_n$  (recall:  $\mathbb{E}[X_i] = 0$ ).

Therefore, the conditional probability distribution of  $\xi_n$  given  $\mathfrak{F}_\eta$  coincides with the unconditional one, so  $\xi_n \sim \mathcal{N}(0, n^{-1})$ , thus  $\xi_n$  has bounded density

$$p_\xi(t) = \frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi n^{-1}}} \leq \frac{n^{1/2}}{\sqrt{2\pi}}.$$

Moreover, for any interval  $I \subset \mathbb{R}$  of length  $|I|$ , we have

$$\text{ess sup } \mathbb{P} \{ \xi_n(\omega) \in I \mid \mathfrak{F} \} = \mathbb{P} \{ \xi_n(\omega) \in I \} \leq \sqrt{\frac{n}{2\pi}} |I|.$$

In some applications to the eigenvalue analysis for random Hamiltonians, discussed below in Section 2 and in subsection 1.2, one has to estimate the probabilities of the form

$$\mathbb{P} \{ \xi_n(\omega) \in I(\eta) \} = \mathbb{P} \{ \xi_n(\omega) \in [\tilde{\mu}(\eta), \tilde{\mu}(\eta) + s] \}, \quad (1.1)$$

where  $\tilde{\mu}(\cdot)$  is some measurable function of  $\eta_1, \dots, \eta_n$ , so the interval  $I(\eta) = [\tilde{\mu}(\eta), \tilde{\mu}(\eta) + s]$  is determined only by the fluctuations  $\{\eta_i, i = 1, \dots, n\}$ . For example, with  $n = 2$ ,

$$\xi = \xi_2 = \frac{X_1 + X_2}{2}, \quad \eta = \eta_1 = \frac{X_1 - X_2}{2},$$

one may consider the probability

$$\begin{aligned} & \mathbb{P} \{ \xi \in [\eta^2, \eta^2 + s] \} \\ &= (2\pi)^{-1} \int_{\mathbb{R}^2} dX_1 dX_2 e^{-\frac{1}{2}(x_1^2 + x_2^2)} \mathbf{1}_A(x_1, x_2), \end{aligned}$$

where the set  $A := \{(x_1, x_2) \in \mathbb{R}^2\}$  is defined by the inequalities

$$\frac{(x_1 - x_2)^2}{4} \leq \frac{x_1 + x_2}{2} \leq \frac{(x_1 - x_2)^2}{4} + s, \quad s > 0.$$

In this particular case – for Gaussian samples – the conditional regularity of the sample mean  $\xi_n$  (given the fluctuations) is granted, but this is not always so, as shows the following elementary example where the common probability distribution of the sample  $X_1, X_2$  is very regular:  $X_i \sim \text{Unif}([0, 1])$ , so  $X_i$  admit a compactly supported probability density bounded by 1. In this simple example the random vector  $(X_1, X_2)$  is uniformly distributed in the unit square  $[0, 1]^2$ , and the condition  $\eta = c$  selects in the two-dimensional plane with coordinates  $(X_1, X_2)$  a straight line parallel to the main diagonal  $\{X_1 = X_2\}$ . The conditional

distribution of  $\xi$  given  $\{\eta = c\}$  is the uniform distribution on the segment

$$J_c := \{(x_1, x_2) : x_1 - x_2 = 2c, 0 \leq x_1, x_2 \leq 1\}$$

of length vanishing at  $2c = \pm 1$ . For  $|2c| = 1$ , the conditional distribution of  $\xi$  on  $J_c$  is concentrated on a single point, which is the ultimate form of singularity.

The above mentioned remarkable particularity of the Gaussian distributions enables one, among other things, to single out one component of a given sample and analyze its impact on one or another functional of the sample of size  $n > 1$ , while keeping fixed the remaining  $n - 1$  degrees of freedom. More to the point, in the spectral analysis of random operators (cf. Section 2), the sample mean modulates a scalar operator (a multiple of the identity operator), which shifts all eigenvalues in an explicit way, while the effect of the fluctuations is much harder to control quantitatively. The question of regularity of the probability distribution of the sample mean  $\xi_n$  with more general (in particular, singular continuous) marginal probability measures, conditional on all the "fluctuations", seems to be interesting in itself; for the author of these lines, however, the motivation came from the spectral theory of random operators.

For the moment, we can address only sufficiently regular marginal probability measures, which leaves the room for further research in this direction.

The main result of the present paper is given by Theorem 8.1, generalizing the estimates from Theorem 6.1 and Theorem 7.1.

## 1.2 Motivation: random $N$ -particle Hamiltonians

The celebrated Anderson localization theory, originated in the seminal paper by P. W. Anderson [1] in 1958 (Nobel Prize in 1977) studies the evolution of quantum particles in a disordered media. Mathematically speaking, one has to study the decay properties of the Green functions, eigenfunctions and eigenfunction correlators of random quantum Hamiltonians (see, e.g., the monographs [8, 9, 15] for a general introduction to the Anderson localization theory). It turns out that the free propagation of particles or waves can be strongly inhibited by disorder. The crucial role, both from mathematical and physical point of view, is played by the so-called "resonances", or "small denominators"; the latter appear in a number of problems in mathematical physics and beyond. In simple terms, one has to assess the regularity of the eigenvalue distribution of random self-adjoint operators (or finite-dimensional symmetric matrices). The first fairly general estimate of such kind was proved in the celebrated paper by Wegner [16]. Today there are various approaches leading to the Wegner-type EVC bounds (cf. [8, 15]) more general and more optimal than those relying on the results of the present paper. However, the situation is quite different in the area of multi-particle disordered quantum systems with a nontrivial inter-particle interaction. A considerable number of challenging problems here remain open due to the lack of adequate methods for assessing the eigenvalue concentration and eigenvalue comparison in interactive systems.

Theory of multi-particle Anderson Hamiltonians is a relatively new direction in this field (cf., e.g., [2, 4, 9]), and it presents greater challenges in the EVC estimation. The

main reason is that already for  $N = 2$  particles, the eigenvalues are "composite" quantities, and the resonances between them, in the form of "small denominators", are structural in nature and cannot be attributed to one or another specific locus in the physical, single-particle configuration space. As a result, correlations between composite EVs do not decay with distance; these strong correlations become particularly difficult to deal with starting from  $N = 3$ . Suffices to say that there still are open problems in the area of  $N$ -particle Anderson Hamiltonians in a Euclidean space, while the situation is somewhat simpler for the lattice Hamiltonians (cf. [10]).

A partial solution to this general problem, initially proposed in [6], used the specific property of Gaussian distributions which we discussed in subsection 1.1. In a modified form, it was extended to the IID random potentials with uniform marginal distributions and "smooth" perturbations thereof; see [5, 10]. However, the Hölder regularity (of any order  $\theta < 1$ ) of the  $N$ -particle EV distribution used in the context of lattice Hamiltonians turns out to be insufficient for the applications to the  $N$ -particle Anderson models in  $\mathbb{R}^d$ , where one needs the Lipschitz continuity of the EV distribution, and this is the main subject of the present paper.

We hope that the approach used in [6] and in this paper may prove useful in other areas of mathematical physics and applications of the probability theory.

A detailed discussion of the interactive Anderson models, especially in the Euclidean space, is certainly well beyond the scope of this short note; an introduction to this relatively recent direction of research in mathematical physics of random operators can be found in the monograph [9]. However, it is to be emphasized that the results of this paper are crucial for the rigorous proof of uniformly strong decay bounds on eigenfunction correlators with respect to the natural distance in the multi-particle configuration space and not in the so-called Hausdorff distance, used in the works [7, 13, 12] and some other papers. We also stress that the new EVC bounds are required for the proof of localization in a physically realistic geometrical setting (in finite-size disordered samples); see the discussion in [10].

## 2 An application to the Wegner-type bounds

Let  $\Lambda$  be a finite graph, with  $|\Lambda| = n \geq 1$ , and  $H(\omega) = H_\Lambda(\omega)$  be a discrete Schrödinger operator (cf., e.g., [8, 9]) acting in the finite-dimensional Hilbert space  $\mathcal{H} = \mathcal{H}_\Lambda = \ell^2(\Lambda)$ , with an IID random potential  $V : \Lambda \times \Omega \rightarrow \mathbb{R}$ , relative to some probability space  $(\Omega, \mathfrak{F}, \mathbb{P})$ . For example, one can take  $\Lambda = [-L, L] \cap \mathbb{Z}^1$  and define the random, second order discrete Schrödinger operator  $H(\omega) : \ell^2(\Lambda) \rightarrow \ell^2(\Lambda)$  by

$$\begin{aligned} (H(\omega)\psi)(x) &= ((H_0 + V(\omega))\psi)(x) \\ &:= -\psi(x - 1) - \psi(x + 1) + V(x, \omega)\psi(x), \end{aligned}$$

with Dirichlet boundary conditions (cf. [8, 9]) outside  $[-L, L]$ ; here  $V : \Lambda \times \Omega \rightarrow \mathbb{R}$  is a random field on  $\Lambda$  relative to some probability space  $(\Omega, \mathfrak{B}, \mathbb{P})$ . On an arbitrary connected finite graph  $\Lambda$ , one usually takes the kinetic energy operator  $H_0$  given by the negative graph Laplacian (cf., e.g., [3]),

$$((-\Delta_\Lambda)\psi)(x) = \sum_{y \in \Lambda: d(x,y)=1} (\psi(x) - \psi(y));$$

here  $d(x, y)$  is the canonical graph-distance on  $\Lambda$ , defined as the shortest path from  $x$  to  $y$  over the edges of the graph in question.

Decomposing the random field  $V$  on  $\Lambda$ ,

$$V(x; \omega) = \xi_n(\omega) + \eta_x(\omega),$$

we can represent  $H(\omega)$  as follows:

$$H(\omega) = \xi_n(\omega) \mathbf{1} + A(\omega),$$

where the self-adjoint operator  $A(\omega)$

$$A = -\Delta_\Lambda + \eta_x(\omega),$$

(here  $\eta_x$  is identified with the operator of multiplication by the "residual", "fluctuation" random potential  $x \mapsto \eta_x(\omega)$ ) is  $\mathfrak{F}_\eta$ -measurable, and so are its eigenvalues  $\mu_j(\omega)$ ,  $j = 1, \dots, n$ . Since  $A(\omega)$  commutes with the scalar operator  $\xi_n(\omega) \mathbf{1}$ , the eigenvalues  $\lambda_j(\omega)$  of  $H(\omega)$  have the form

$$\lambda_j(\omega) = \xi_n(\omega) + \mu_j(\omega). \quad (2.1)$$

The numeration of the eigenvalues  $\lambda_j(\omega)$ ,  $\mu_j(\omega)$  is of course not canonical, but they can be consistently defined as random variables on  $\Omega$ .

The representation (2.1) implies immediately the following EVC bound. Fix an interval  $I_s = [t, t + s]$  and let  $P_{I_s}(H(\omega))$  be the spectral projection of operator  $H(\omega)$  to the interval  $I_s$  (cf., e.g., [14]), then

$$\begin{aligned} \mathbb{P} \{ \text{tr } P_{I_s}(H(\omega)) \geq 1 \} &\leq \sum_{j=1}^n \mathbb{P} \{ \lambda_j(\omega) \in I_s \} \\ &= \sum_{j=1}^n \mathbb{E} [ \mathbb{P} \{ \xi_n + \mu_j \in I_s \mid \mathfrak{F}_\eta \} ] \\ &= \sum_{j=1}^n \mathbb{E} [ \mathbb{P} \{ \xi_n \in [-\mu_j + t, -\mu_j + t + s] \mid \mathfrak{F}_\eta \} ] \end{aligned} \quad (2.2)$$

Further, omitting the argument  $\omega$  for brevity and setting  $\tilde{\mu}_j(\omega) := -\mu_j(\omega) + t$ , we have

$$\begin{aligned} &\mathbb{P} \{ \xi_n + \mu_j \in I_s \mid \mathfrak{F}_\eta \} \\ &= \mathbb{P} \{ \xi_n \in [-\mu_j + t, -\mu_j + t + s] \mid \mathfrak{F}_\eta \} \\ &= \mathbb{P} \{ \xi_n \in [\tilde{\mu}_j, \tilde{\mu}_j + s] \mid \mathfrak{F}_\eta \} \end{aligned} \quad (2.3)$$

where  $\tilde{\mu}_j(\omega)$  are  $\mathfrak{F}_\eta$ -measurable, i.e., fixed under the conditioning. Therefore, the unconditional probability  $\mathbb{P} \{ \lambda_j \in I_s \}$  can be assessed by analyzing the regularity of the conditional probability distribution of  $\xi_n$  figuring in the RHS of (2.3).

### 3 Reduction to the local analysis in the sample space

Assume that the support  $\mathcal{S} \subset \mathbb{R}$  of the common continuous marginal probability measure of the IID random variables  $X_j$ ,  $1 \leq j \leq n$ , is covered by a finite or countable union of intervals:

$$\mathcal{S} \subset \cup_{k \in \mathcal{K}} J_k, \quad \mathcal{K} \subset \mathbb{Z}, \quad J_k = [a_k, b_k], \quad a_{k+1} \geq b_k.$$

Let  $\mathbf{K} = \mathcal{K}^n$ , and for each  $\mathbf{k} = (k_1, \dots, k_n) \in \mathbf{K}$ , denote

$$\mathbf{J}_{\mathbf{k}} = \times_{i=1}^n J_{k_i}.$$

Owing to the continuity of the marginal measure,  $J_k$  are essentially disjoint: two distinct intervals  $J_k, J_l$  can only have common endpoints. Respectively, the family of the parallelepipeds  $\{\mathbf{J}_{\mathbf{k}}, \mathbf{k} \in \mathbf{K}\}$  forms a partition  $\mathcal{K}$  of the sample space, which we will often identify with the probability space  $\Omega$ . Further, let  $\mathfrak{F}_{\mathcal{K}}$  be the sub-sigma-algebra of  $\mathfrak{F}$  generated by the partition  $\mathcal{K}$ . The probabilities of the general form (1.1) can be assessed as follows:

$$\begin{aligned} &\mathbb{P} \{ \xi_n \in [\tilde{\mu}, \tilde{\mu} + s] \} \\ &= \mathbb{E} [ \mathbb{P} \{ \xi_n \in [\tilde{\mu}, \tilde{\mu} + s] \mid \mathfrak{F}_{\mathcal{K}} \} ] \\ &= \sum_{\mathbf{k} \in \mathbf{K}} \mathbb{P} \{ \mathbf{J}_{\mathbf{k}} \} \mathbb{P} \{ \xi_n \in [\tilde{\mu}, \tilde{\mu} + s] \mid \mathbf{J}_{\mathbf{k}} \}. \end{aligned}$$

Let  $\mathbb{P}_{\mathbf{k}} \{ \cdot \}$  be the conditional probability measure, given the event  $\{X \in \mathbf{J}_{\mathbf{k}}\}$ ,  $\mathbb{E}_{\mathbf{k}} [ \cdot ]$  the respective expectation, and  $p_{\mathbf{k}} = \mathbb{P} \{ \mathbf{J}_{\mathbf{k}} \}$ . Then we have

$$\begin{aligned} &\mathbb{P} \{ \xi_n \in [\tilde{\mu}, \tilde{\mu} + s] \} \\ &= \sum_{\mathbf{k} \in \mathbf{K}} p_{\mathbf{k}} \mathbb{E}_{\mathbf{k}} [ \mathbb{P}_{\mathbf{k}} \{ \xi_n \in [\tilde{\mu}, \tilde{\mu} + s] \mid \mathfrak{F}_\eta \} ] \\ &\leq \sup_{\mathbf{k} \in \mathbf{K}} \mathbb{E}_{\mathbf{k}} [ \mathbb{P}_{\mathbf{k}} \{ \xi_n \in [\tilde{\mu}, \tilde{\mu} + s] \mid \mathfrak{F}_\eta \} ]. \end{aligned} \quad (3.1)$$

Therefore, one may seek a satisfactory bound on the LHS of (3.1) by assessing the "local" conditional probabilities  $\mathbb{P}_{\mathbf{k}} \{ \xi_n \in [\tilde{\mu}, \tilde{\mu} + s] \mid \mathfrak{F}_\eta \}$ , where each random variable  $X_j$  is restricted to a subinterval  $J_{k_j}$  of its global support, so the entire sample  $X = (X_1, \dots, X_n)$  is restricted to the parallelepiped  $\mathbf{J}_{\mathbf{k}} \subset \mathbb{R}^n$ .

In the next section, we perform such analysis in the case of a uniform distribution of the IID variables  $X_i$ .

### 4 Uniform marginal distributions. General setup

Let be given a real number  $\ell > 0$  and an integer  $n \geq 2$ . Consider a sample of  $n$  IID random variables with uniform distribution  $\text{Unif}([0, \ell])$ , and introduce again the sample mean  $\xi = \xi_n$  and the "fluctuations"  $\eta_i$  relative to the mean:

$$\xi_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad \eta_i = X_i - \xi_n.$$

Further, consider the  $n$ -dimensional Euclidean space of real linear combinations of the random variables  $X_i$ . Clearly, the variables  $\eta_i : \mathbb{R}^n \rightarrow \mathbb{R}$  are invariant under the group of translations

$$(X_1, \dots, X_n) \mapsto (X_1 + t, \dots, X_n + t), \quad t \in \mathbb{R},$$

and so are their differences  $\eta_i - \eta_j \equiv X_i - X_j$ ,  $1 \leq i < j \leq n$ . Introduce the variables

$$Y_i = \eta_i - \eta_n, \quad 1 \leq i \leq n-1. \quad (4.1)$$

Then the space  $\mathbb{R}^n$  is stratified into a union of affine lines of the form

$$\begin{aligned} \mathcal{L}(Y) &:= \{X \in \mathbb{R}^n : \eta_i - \eta_n = Y_i, i \leq n-1\} \\ &= \{X \in \mathbb{R}^n : X_i - X_n = Y_i, i \leq n-1\}, \end{aligned} \quad (4.2)$$

labeled by the elements  $Y = (Y_1, \dots, Y_{n-1})$  of the  $(n-1)$ -dimensional real vector space  $\mathbb{Y}^{n-1}$  orthogonal to the vector

$(1, \dots, 1)$ . Denote

$$\begin{aligned} \mathcal{X}(Y) &= \mathcal{L}(Y) \cap [0, \ell]^n \\ &= \{X \in [0, \ell]^n : X_i - X_n = Y_i, i \leq n-1\} \end{aligned}$$

and endow each nonempty interval  $\mathcal{X}(Y) \subset \mathbb{R}^n$  with the natural structure of a probability space inherited from  $\mathbb{R}^n$ :

- if  $|\mathcal{X}(Y)| = 0$  (an interval reduced to a single point), then we introduce the trivial sigma-algebra and the counting measure;
- if  $|\mathcal{X}(Y)| = r > 0$ , then we use the inherited structure of an interval of a one-dimensional affine line and the normalized measure with constant density  $r^{-1}$  with respect to the inherited Lebesgue measure on  $\mathcal{X}(Y)$ .

The transformation  $X \mapsto (\xi_n, \eta_1, \dots, \eta_{n-1})$  is non-degenerate, but not orthogonal. We will have to work with the metric on  $\mathcal{X}(Y)$ , induced by the standard Riemannian metric in the ambient space  $\mathbb{R}^n$ ; to this end, introduce an orthogonal coordinate transformation in  $\mathbb{R}^n$ ,  $X \mapsto (\tilde{\xi}_n, \tilde{\eta}_1, \dots, \tilde{\eta}_{n-1})$ , such that

$$\tilde{\xi}_n = n^{-1/2} \sum_{i=1}^n X_i = n^{1/2} \xi_n; \quad (4.3)$$

the exact form of  $\tilde{\eta}_j, j = 1, \dots, n-1$  is of no importance for our analysis, provided that the transformation is orthogonal.

*Remark 4.1.* For later use, note that, due to (4.3), each of the re-scaled variables  $n^{1/2} X_i$  can serve as the normalized length parameter on the elements  $\mathcal{X}(Y)$ . Along an element  $\mathcal{X}(Y)$ , one can simultaneously parameterize  $\tilde{\xi}$  and the variables  $X_i$ , by setting  $\tilde{\xi}(t) = c_0 + t, X_j(t) = c_j + n^{-1/2} t$ , with arbitrarily chosen constants  $c_j$ . Here,  $\xi_n$  is a natural length parameter on  $\mathcal{X}(Y)$ , since the transformation  $X \mapsto (\tilde{\xi}_n, \tilde{\eta}_1, \dots, \tilde{\eta}_{n-1})$  is orthogonal.

## 5 Short intervals are unlikely

**Lemma 5.1.** *Assume that the IID random variables  $X_1, \dots, X_n, n \geq 2$ , are uniformly distributed in an interval  $[0, \ell], \ell > 0$ . Then for all  $t \in (0, \ell/2]$  one has*

$$\mathbb{P}\{|\mathcal{X}(Y)| < t\} \leq t^2 n / \ell^2. \quad (5.1)$$

*Proof.* Let

$$\underline{X} = \underline{X}(X) = \min_i X_i, \quad \overline{X} = \overline{X}(X) = \max_i X_i. \quad (5.2)$$

While  $\overline{X}(X)$  and  $\underline{X}(X)$  vary along the elements  $\mathcal{X}(Y)$ , their difference  $\overline{X}(X) - \underline{X}(X)$  does not; it is uniquely determined by  $\mathcal{X}(Y)$ .

According to Remark 4.1, each  $n^{1/2} X_i, i = 1, \dots, n$ , restricted to  $\mathcal{X}(Y)$ , provides a normalized length parameter on  $\mathcal{X}(Y)$ ; thus the range of each  $n^{1/2} X_i|_{\mathcal{X}(Y)}$  is an interval of length  $|\mathcal{X}(Y)|$ . One can increase (resp., decrease), e.g., the value of  $X_1$ , as long as all  $\{X_i, 1 \leq i \leq n\}$  are strictly smaller than  $\ell$  (resp., strictly positive). Therefore, the maximum increment of  $X_1$  (indeed, of any  $X_i$ ) along  $\mathcal{X}(Y)$  is given by  $\ell - \overline{X}(X)$ , and its maximum decrement equals  $\underline{X}(X)$ , so the range of the normalized length parameter  $n^{1/2} X_1$  along  $\mathcal{X}(Y(X))$  is an interval of length  $n^{1/2}(\ell - \overline{X}(X) + \underline{X}(X))$ :

$$|\mathcal{X}(Y(X))| = n^{1/2}(\ell - \overline{X}(X) + \underline{X}(X)). \quad (5.3)$$

Since both  $\underline{X}(X)$  and  $\ell - \overline{X}(X)$  are non-negative, one has the implication

$$|\mathcal{X}(Y)| = \underline{X} + (\ell - \overline{X}) < t \implies \max\{\underline{X}, \ell - \overline{X}\} < t.$$

With  $0 \leq t \leq \ell/2$ ,  $(\ell - X_i < t)$  implies  $(X_i > t)$ , thus denoting

$$A_{ij}(t) := \{X_i < t\} \cap \{\ell - X_j < t\}, \quad (5.4)$$

we have, for any  $i$ ,

$$A_{ii}(t) = \{X_i < t\} \cap \{\ell - X_i < t\} = \emptyset. \quad (5.5)$$

Therefore,

$$\{\max\{\underline{X}, \ell - \overline{X}\} < t\} \subset \bigcup_{i \neq j} \{X_i < t, \ell - X_j < t\}.$$

Thus the union  $\cup_{i \neq j} A_{ij}(t)$  contains all the samples  $X$  with  $|\mathcal{X}(Y)| < t$ .

The sample  $\{X_k\}$  is IID, with common probability density bounded by  $\bar{\rho} = \ell^{-1}$ , so for any  $i \neq j$

$$\mathbb{P}\{A_{ij}(t)\} = \mathbb{P}\{X_i < t\} \cdot \mathbb{P}\{\ell - X_j < t\} = \bar{\rho}^2 t^2.$$

Therefore,

$$\begin{aligned} \mathbb{P}\{|\mathcal{X}(Y)| < t\} &= \mathbb{P}\left\{n^{1/2}((\ell - \overline{X}(X)) + \underline{X}(X)) < t\right\} \\ &= \mathbb{P}\left\{((\ell - \overline{X}(X)) + \underline{X}(X)) < rn^{-1/2}\right\} \\ &\leq \sum_{i \neq j} \mathbb{P}\{A_{ij}(rn^{-1/2})\} \leq n(n-1) (\bar{\rho} rn^{-1/2})^2 \\ &\leq r^2 n / \ell^2. \end{aligned} \quad (5.6)$$

This completes the proof.  $\square$

By a change of variable, one can extend the above result to  $n$  independent random variables uniformly distributed in their individual intervals  $J_i = [a_i, a_i + \ell]$ , for arbitrary  $a_1, \dots, a_n \in \mathbb{R}$ .

## 6 Regularity bound for the uniform distribution

**Theorem 6.1.** *Let be given IID random variables  $X_1, \dots, X_n$  and a measurable function  $\lambda : Y \mapsto \lambda(Y)$ . In each interval  $\mathcal{X}(Y) \subset \mathcal{L}(Y) \cong \mathbb{R}$ , introduce the sub-interval  $I_s(Y) = [\lambda(Y), \lambda(Y) + s] \cap \mathcal{L}(Y)$ . Then for any  $s \in (0, 1]$ , one has*

$$\mathbb{P}\{\xi(\omega) \in I_s(Y)\} \leq 3n^3 \ell^{-1} s. \quad (6.1)$$

*Proof.* Recall that the natural length parameter on the lines  $\mathcal{L}(Y)$  is given by  $\tilde{\xi} = \sqrt{n}\xi$  and not by  $\xi$  itself, since the gradient of  $\xi$  is given by the vector  $(n^{-1}, \dots, n^{-1})$  of norm  $1/\sqrt{n}$ . Therefore, when  $\tilde{\xi}$  runs through  $I_s(Y)$ ,  $\xi$  runs through the interval  $\tilde{I}_s = [\tilde{\lambda}(Y), \tilde{\lambda}(Y) + \sqrt{ns}]$ ,  $\tilde{\lambda}(Y) = \sqrt{n}\lambda(Y)$ , of length  $\sqrt{ns}$ .

Set  $l(\omega) := |\mathcal{X}(Y)|$ , and denote by  $\mathbf{1}_{l \geq s\sqrt{n}}$  the indicator function of the event  $\{l \geq s\sqrt{n}\}$ , then we have

$$\begin{aligned} \mathbb{P}\{\xi \in I_s(\eta)\} &= \mathbb{E}\left[\mathbb{P}\left\{\tilde{\xi} \in \tilde{I}_s(\eta) \mid \mathfrak{F}_\eta\right\}\right] \\ &= \mathbb{E}\left[\mathbf{1}_{l < s\sqrt{n}} \mathbb{P}\left\{\tilde{\xi} \in \tilde{I}_s(\eta) \mid \mathfrak{F}_\eta\right\}\right] \\ &\quad + \mathbb{E}\left[\mathbf{1}_{l \geq s\sqrt{n}} \mathbb{P}\left\{\tilde{\xi} \in \tilde{I}_s(\eta) \mid \mathfrak{F}_\eta\right\}\right] \\ &\leq \mathbb{P}\{l < s\sqrt{n}\} + \mathbb{E}\left[\mathbf{1}_{l \geq s\sqrt{n}} \mathbb{P}\left\{\tilde{\xi} \in \tilde{I}_s(\eta) \mid \mathfrak{F}_\eta\right\}\right], \end{aligned} \quad (6.2)$$

where, by virtue of (5.6),

$$\mathbb{P}\{l(\omega) < s\sqrt{n}\} \leq n(s\sqrt{n})^2/\ell^2 = n^2s^2/\ell^2, \quad (6.3)$$

yielding

$$\sup_{s>0} \frac{\mathbb{P}\{l(\omega) < s\sqrt{n}\}}{s^2} \leq n^2/\ell^2. \quad (6.4)$$

The second summand in the last RHS in (6.2) can be assessed as follows:

$$\begin{aligned} \mathbb{E}\left[\mathbf{1}_{l \geq s\sqrt{n}} \mathbb{P}\left\{\tilde{\xi} \in \tilde{I}_s(\eta) \mid \mathfrak{F}_\eta\right\}\right] &\leq \mathbb{E}\left[\mathbf{1}_{l \geq s\sqrt{n}} \frac{s}{l}\right] \\ &= s \mathbb{E}\left[\mathbf{1}_{l \geq s\sqrt{n}} (l(\omega))^{-1}\right] = s\sqrt{n} \int_{s\sqrt{n}}^{\ell\sqrt{n}} r^{-1} dF_l(r). \end{aligned} \quad (6.5)$$

Now we apply the well-known formula of integration by parts for the Stieltjes integral (cf., e.g., [11, Sect. V.6]) giving the moment of order  $\alpha$  of a probability distribution with probability distribution function (PDF)  $F(\cdot)$ :

$$\int_0^\infty r^\alpha dF(r) = \alpha \int_0^\infty r^{\alpha-1} F(r) dr,$$

valid under the assumption of convergence of the RHS integral. In our case, we only have to integrate over a finite interval  $[s\sqrt{n}, \ell\sqrt{n}]$ , thus avoiding singularity at  $r = 0$ . For the PDF  $F_l(\cdot)$ , using the upper bound (6.4), we obtain

$$\begin{aligned} \int_{s\sqrt{n}}^{\ell\sqrt{n}} \frac{dF_l(r)}{r} &= \frac{F_l(r)}{r} \Big|_{s\sqrt{n}}^{\ell\sqrt{n}} + \int_{s\sqrt{n}}^{\ell\sqrt{n}} r^{-2} F_l(r) dr \\ &\leq \frac{1}{\ell\sqrt{n}} + \ell\sqrt{n} \sup_{r>0} \frac{F_l(r)}{r^2} \leq \frac{1}{\ell\sqrt{n}} + \frac{\ell\sqrt{n} \cdot n^2}{\ell^2} \\ &\leq \frac{2n^{5/2}}{\ell}. \end{aligned} \quad (6.6)$$

Collecting (6.3), (6.5), (6.6), and using  $s/\ell \leq 1$ , the assertion follows:

$$\mathbb{P}\{\xi \in I_s(\eta)\} \leq n^2 \frac{s^2}{\ell^2} + \frac{2n^{5/2} \cdot s\sqrt{n}}{\ell} \leq \frac{3n^3}{\ell} s. \quad (6.7)$$

□

## 7 Smooth positive densities

Now we consider a richer class of probability distributions. While the conditions which we assume are certainly very restrictive, they are sufficient for the applications to some physically relevant Anderson models.

**Theorem 7.1.** *Assume that the common probability distribution of the IID random variables  $V_j$ ,  $j = 1, \dots, n$ , with the cumulative probability distribution function  $F_V$ , satisfies the following conditions:*

(i) *the probability distribution is absolutely continuous:*

$$dF_V(v) = \rho(v) dv, \quad \text{supp } \rho = [a, a + \ell]; \quad (7.1)$$

(ii) *the probability density  $\rho(\cdot)$  has bounded logarithmic derivative on  $(a, a + \ell)$ :*

$$\|(\ln \rho)' \mathbf{1}_{(a, a+\ell)}\|_\infty \leq C'_\rho < +\infty. \quad (7.2)$$

*Then there exists a constant  $C = C(F_V, \ell) < \infty$  such that for any  $s \in (0, \ell n^{-1})$  and any  $\mathfrak{F}_\eta$ -measurable random variable  $\lambda$ , setting  $I_s(\omega) := [\lambda(\omega), \lambda(\omega) + s]$ , one has the following bound:*

$$\mathbb{P}\{\xi_n(\omega) \in I_s(\omega)\} \leq Cn^2s. \quad (7.3)$$

*Proof.* Without loss of generality, it suffices to prove the claim for  $\text{supp } \rho = [0, \ell]$ , which we assume below.

As in Section 3, introduce a partition of the sample space into the cubes  $\mathbf{J}_\mathbf{k}$ , induced by the decomposition  $[0, \ell] = \sqcup_{\mathbf{k}} J_\mathbf{k}$ :

$$\mathbf{J}_\mathbf{k} = \prod_{i=1}^n J_{k_i}, \quad \mathbf{k} = (k_1, \dots, k_n),$$

where

$$J_k = [a_k, a_{k+1}], \quad a_k := \frac{k-1}{n}, \quad k = 1, \dots, n.$$

The hypothesis (7.2) implies that for any  $\mathbf{x} \in \mathbf{J}_\mathbf{k}$  the logarithm of  $\mathbf{p}(\mathbf{x})$  is well-defined and satisfies the upper bound

$$|\ln \mathbf{p}(\mathbf{x}) - \ln \mathbf{p}(\mathbf{a}_\mathbf{k})| \leq \sum_{i=1}^n |\ln \rho(x_i) - \ln \rho(a_{k_i})| \leq \frac{n C'_\rho \ell}{n},$$

thus

$$\forall \mathbf{x} \in \mathbf{J}_\mathbf{k} \quad \mathbf{p}(\mathbf{x}) = \mathbf{p}(\mathbf{a}_\mathbf{k}) e^{\alpha_n(\mathbf{x})},$$

where  $|\alpha_n(\mathbf{x})| \leq C\ell$ ; with  $\ell$  fixed in the condition (7.1),  $\mathbf{p}(\mathbf{x})$  is therefore uniformly bounded.

Now introduce in  $\mathbf{J}_\mathbf{k}$ :

- the uniform probability distribution  $\tilde{\mathbf{P}}_\mathbf{k}$ , i.e., the normalized measure with constant density  $\tilde{\mathbf{p}}_\mathbf{k}$  w.r.t. the Lebesgue measure;
- the probability distribution induced by  $\mathbf{P}$ , conditional on  $\{\mathbf{X} \in \mathbf{J}_\mathbf{k}\}$ , i.e., the normalized measure with density

$$\mathbf{p}_\mathbf{k}(\mathbf{x}) = Z_\mathbf{k}^{-1} \mathbf{p}(\mathbf{x}) = \frac{\mathbf{p}(\mathbf{x})}{\int_{\mathbf{J}_\mathbf{k}} \mathbf{P}(\mathbf{y}) d\mathbf{y}}.$$

Due to continuity of the density  $\mathbf{p}$ , we have  $\int_{\mathbf{J}_\mathbf{k}} \mathbf{P}(\mathbf{y}) d\mathbf{y} = c|\mathbf{J}_\mathbf{k}|$ , for some  $c \in [e^{-\alpha_n}, e^{+\alpha_n}]$ , so

$$\frac{\mathbf{p}_\mathbf{k}(\mathbf{x})}{\tilde{\mathbf{p}}(\mathbf{x})} = \frac{\mathbf{p}(\mathbf{x})}{c} \in [e^{-2\alpha_n}, e^{+2\alpha_n}]$$

Hence for any event  $\mathcal{A}$ , we have

$$e^{-2\alpha_n} \mathbb{P}\{\mathcal{A}\} \leq \mathbb{P}_\mathbf{k}\{\mathcal{A}\} \leq e^{2\alpha_n} \mathbb{P}\{\mathcal{A}\}. \quad (7.4)$$

Finally, it follows from (7.4) and (3.1) that

$$\mathbb{P}\{\xi \in I_s(\eta)\} \leq \sup_{\mathbf{k}} \mathbb{P}_\mathbf{k}\{\xi \in I_s(\eta)\} \leq C(F_V, \ell) n^2 s.$$

□

We will call the property of the common probability distribution of the IID random variables  $X_1, \dots, X_n$ , expressed by the inequality of the form (7.3), with the RHS polynomially bounded in  $n$ , the strong regularity of the conditional mean (SRCM).

## 8 Extension to convolution measures

The results of Section 7 can be easily adapted to a class of convolution probability measures  $\mathbb{P} = \mathbb{P}_1 * \mathbb{P}_2$  where at least one of the measures  $\mathbb{P}_1, \mathbb{P}_2$  satisfies the hypotheses of Theorem 7.1.

**Theorem 8.1.** *Assume that the common probability distribution  $\nu$  of the IID random variables  $V_j, j = 1, \dots, n$ , admits the representation*

$$\nu = \nu_1 * \nu_2, \tag{8.1}$$

where  $\nu_1$  satisfies the hypotheses of Theorem 7.1. Then there exists a constant  $C = C(F_V, \ell) < \infty$  such that for any  $s \in (0, \ell n^{-2})$  and any  $\mathfrak{F}_\eta$ -measurable random variable  $\lambda$ , setting  $I_s(\omega) := [\lambda(\omega), \lambda(\omega) + s]$ , one has the following bound:

$$\mathbb{P} \{ \xi_n(\omega) \in I_s(\omega) \} \leq C n^3 s. \tag{8.2}$$

*Proof.* It follows immediately from the assumption (8.1) that the random variables  $V_j$  can be represented as sums  $V_j(\omega) = A_j(\omega) + B_j(\omega)$ , where

- the family of random variables  $\{A_j, 1 \leq j \leq n\}$  is IID, with common probability distribution  $\nu_1$ ;
- the family of random variables  $\{B_j, 1 \leq j \leq n\}$  is also IID, with common probability distribution  $\nu_2$ ;
- the families  $\{A_j\}$  and  $\{B_j\}$  are mutually independent.

Let  $\mathbb{P}, \mathbb{P}_1$  and  $\mathbb{P}_2$  be the product probability measures in the sample space  $\cong \mathbb{R}^n$ , generated respectively by the marginal measures  $\nu, \nu_1$  and  $\nu_2$ . Further, let  $\mathfrak{F}_B$  be the sigma-algebra generated by the family  $\{B_j, j = 1, \dots, n\}$ . Then the conditional measure  $\mathbb{P} \{ \cdot | \mathfrak{F}_B \}$  is equivalent to  $\mathbb{P}_1$ , thus

$$\begin{aligned} \mathbb{P} \{ \xi_n(\omega) \in I_s(\omega) \} &= \mathbb{E}^{(\mathbb{P})} \left[ \mathbb{P} \{ \xi_n(\omega) \in I_s(\omega) | \mathfrak{F}_B \} \right] \\ &= \mathbb{E}^{(\mathbb{P})} \left[ \mathbb{P}_1 \{ \xi_n(\omega) \in I_s(\omega) \} \right] \\ &\leq C n^3 s. \end{aligned}$$

□

**Example 1.** The so-called triangular distribution  $\nu * \nu$ , where  $\nu$  is the uniform distribution in  $[0, 1]$ . More generally, one can take a convolution power  $\nu^{*n}, n \geq 2$ . This example shows that the property SRCM can hold in a class of probability measures with density vanishing at the edges of its support.

**Example 2.** The convolution  $\nu = \nu_1 * \nu_2$  of the uniform distribution  $\nu_1 = \text{Unif}([0, 1])$  with the exponential distribution  $\nu_2$  (with density  $\mathbf{1}_{[0, +\infty)} e^{-t}$ ). Here the property SRCM holds true for upper-unbounded random variables.

As pointed out in the Introduction (cf. also [6]), the Gaussian distribution features a particularly strong form of the property SRCM. Theorem 8.1 extends SRCM to a much larger class of unbounded random variables. In particular, note that the decay rate of the tail probabilities  $\mathbb{P} \{ |V(\omega)| > t \}$ , as  $t \rightarrow +\infty$ , can be quite slow. For example, one can take as  $\nu_2$  any stable law (cf., e.g., [11]) of index  $\alpha \in [1, 2]$ ; here  $\alpha = 2$  corresponds to the Gaussian distribution and  $\alpha = 1$  to the Cauchy distribution.

## 9 Conclusion

We have shown that the well-known and widely used property of the Gaussian IID samples of random variables, viz.

the independence of the sample mean of the "sample fluctuations", has a direct analog for a much larger class of marginal probability distributions. For the moment, our approach requires the marginal distribution to have a smooth probability density with respect to the Lebesgue measure. Compared to prior results (cf. [5, 10]), where the conditional distribution of the sample mean (given the sigma-algebra generated by all sample fluctuations) was proved to be Hölder-continuous of order  $\theta = 2/3$ , the present paper gives a sharp Hölder exponent  $\theta = 1$  (thus proving Lipschitz continuity of the conditional measure at hand).

The new result is an important component in the spectral analysis of random  $N$ -particle quantum Hamiltonians, describing the quantum transport of interacting particles (e.g., electrons) in a disordered environment. Quantum transport and/or localization in interacting disordered systems is a relatively new direction both in theoretical and mathematical physics, where there still are many challenging open problems, and our results shed a new light in this direction. They will constitute a major component in the localization analysis of  $N$ -particle Anderson Hamiltonians with nontrivial interaction in a Euclidean space, in a forthcoming paper extending the results of our recent work [10] to continuous systems.

We hope that the probabilistic problem investigated in this paper, as well as the method used here, can prove useful in a more general mathematical framework where analytic aspects intertwine with the probabilistic ones.

## REFERENCES

- [1] P. W. Anderson, P.W. *Absence of diffusion in certain random lattices.* Phys. Rev., **109** (1958), 1492–1505.
- [2] V. Chulaevsky and Y. Suhov, *Wegner bounds for a two-particle tight binding model.* Commun. Math. Phys. (2008) **283**, 479–489.
- [3] F. R. K. Chung (1997). Spectral graph theory, **92**. CBMS Regional Conference Series in Mathematics, Washington, DC .
- [4] V. Chulaevsky and Y. Suhov. *Eigenfunctions in a two-particle Anderson tight binding model.* Commun. Math. Phys., **289** (2009), 701–723.
- [5] Chulaevsky, V., A remark on charge transfer processes in multi-particle systems, Online available from <http://arxiv.org/pdf/1005.3387.pdf>.
- [6] Chulaevsky, V. (2011). On resonances in disordered multi-particle systems. *C. R. Acad. Sci. Paris, Ser. I*, **350**, 81–85.
- [7] Chulaevsky, V., Boutet de Monvel, A., & Suhov, Y. (2011). Dynamical localization for a multi-particle model with an alloy-type external random potential. *Nonlinearity*, **24**, 1451–1472.
- [8] Carmona, R., & Lacroix, J. (1990) *Spectral Theory of Random Schrödinger Operators*. Boston: Birkhäuser.
- [9] V. Chulaevsky, V., & Suhov, Y. (2013) *Multi-scale Analysis for Random Quantum Systems with Interaction*. Boston: Birkhäuser, Boston.
- [10] Chulaevsky, V., Suhov, Y., *Efficient Anderson localization bounds for large multi-particle systems.* J. Spec. Theory (to appear).

- [11] Feller, W. (1966). *An Introduction to Probability Theory and its Applications*. New York: Wiley.
- [12] Fauser, M. & Warzel, S., Multiparticle localization for disordered systems on continuous space via the fractional moment method, Online available from <http://arxiv.org/pdf/1402.5832.pdf>.
- [13] Klein, A., & Nguyen S. T. (2015). Bootstrap multiscale analysis and localization for multi-particle continuous Anderson Hamiltonians. *J. Spec. Theory* (to appear).
- [14] Reed, M. & Simon, B. (1980). *Methods of modern mathematical physics*. Academic Press, Sand Diego.
- [15] Stollmann, P. (2001). *Caught by disorder*. Boston: Birkhäuser, Boston.
- [16] Wegner, F. (1981). Bounds on the density of states in disordered systems. *Z. Phys. B. Cond. Matter*, 44, 9–15.