Centre de
Mathématiques
Laurent Schwartz

X ECOLE
POLYTECHNIQUE

François Germinet

On Bernoulli decomposition of random variables and recent various applications

ON BERNOULLI DECOMPOSITION OF RANDOM VARIABLES AND RECENT VARIOUS APPLICATIONS

FRANÇOIS GERMINET

ABSTRACT. In this review, we first recall a recent Bernoulli decomposition of any given non trivial real random variable. While our main motivation is a proof of universal occurence of Anderson localization in continuum random Schrödinger operators, we review other applications like Sperner theory of antichains, anticoncentration bounds of some functions of random variables, as well as singularity of random matrices.

Contents

1. Bernoulli decomposition

Let X be a real random variable that is non degenerate (i.e. non constant). Throughout this review, we shall make use of the following property (that clearly implies that X is non degenerate)

(H) *There exists $\rho \in]0, \frac{1}{2}[$ such that $\mathbb{P}(X < x^-) > \rho$ and $\mathbb{P}(X > x^+) > \rho$ for some real numbers $x^- < x^+$.*

To fix notations we say that ε is a Bernoulli random variable with parameter $p \in]0, 1[$, if $\mathbb{P}(\varepsilon = 0) = 1 - p$ and $\mathbb{P}(\varepsilon = 1) = p$. The first result asserts that from any non degenerate real random variable one can extract a Bernoulli part. Roughly this comes from the following observation: one can cut the support of the variable in two pieces of positive probabilities $1 - p$ and p, parametrize these two pieces, and use a Bernoulli variable to decide in which piece falls a given realization.

Definition 1.1. *Let X be a real random variable. Let f, δ be measurable functions, such that $f :]0, 1[\to \mathbb{R}$ is monotone increasing and $\delta :]0, 1[\to [0, +\infty[$. Let $p \in]0, 1[$. We say that (f, δ, p) is a Bernoulli decomposition of X if (in law)*

$$X = f(t) + \delta(t)\varepsilon, \tag{1.1}$$

---

where t and ε are two independent random variables, such that t has the uniform distribution in $]0,1[$; and ε is a Bernoulli with parameter p.

Theorem 1.2 ([AGKW]). *Let X be a real non degenerate random variable.*
1. For any $p \in]0,1[$ there exists a Bernoulli decomposition (f,δ,p) of X.
2. There exists $p \in]0,1[$ so that X admits a Bernoulli decomposition (f,δ,p) with $\inf_{t \in]0,1[} \delta(t) > 0$.
3. Assume Property (H). There exists $p \in]\rho, 1-\rho[$ so that X admits a Bernoulli decomposition (f,δ,p) with $\inf_{t \in]0,1[} \delta(t) > 0$.
4. Assume Property (H). Then the Bernoulli decomposition $(f,\delta,p=\frac{1}{2})$ satisfies $\mathbb{P}_t(\delta(t) > x_+ - x_-) \geq 2\rho$.

Remark 1.3. *Property (H) is convenient in order to get a control on the Bernoulli parameter p. Such a control turns out to be crucial when considering families of independent but not necessarily identical random variables.*

As mentioned in [AGKW], the presence of a Bernoulli component in any random variable was noted implicitly in the work of A. N. Kolmogorov [Ko] where it was put to use in an improvement of the earlier concentration bounds of W. Doeblin and P. Lévy [DoL, Do] on linear functions of independent random variables. Initially, Kolmogorov did not extract the maximal benefit from the method by not connecting it with Sperner theory, and in particular the concentration bound in [Ko] includes an unnecessary logarithmic factor; the corresponding improvement was made by B. A. Rogozin [R1]. The bounds were further improved in a series of works, in particular [Es, Ke, R2] where use was also made of other methods.

For random variables with values in $\mathbb{Z}$, a representation similar to the one of Theorem 1.2, but with δ taking values $0, 1$ (so not necessarily positive), has been derived by D. McDonald, and used in the analysis of local limit theorems for integer random variables ([McD]). An application of McDonald's decomposition to statistical mechanics can be found in [I].

For the reader's convenience, we sketch the proof of Theorem 1.2. Points 1 and 2 correspond to [AGKW, Theorem 2.1] and Point 4 comes from [AGKW, Theorem 2.2]. Point 3 can be found in [BrG].

Proof. 1. We denote by μ the law of X and by F its distribution function: $F(u) = \mu(]-\infty, u])$. We set, for any $t \in]0,1[$,

$$G(t) := \inf\{u, F(u) \geq t\}. \tag{1.2}$$

Note that

$$G(t) \leq u \quad \Longleftrightarrow \quad F(u) \geq t, \tag{1.3}$$

so that X and $G(t)$ have the same law ($G(t)$ can be seen as a parametrization of X). For $p \in]0,1[$ given, following [AGKW, Proof of Theorem 2.1], we set for $t \in]0,1[$:

$$
\begin{aligned}
Y_1(t) &:= G((1-p)t) \\
Y_2(t) &:= G(1-p+pt).
\end{aligned}
\tag{1.4}
$$

We then let

$$
\begin{aligned}
f(t) &:= Y_1(t), \tag{1.5}\\
\delta(t) &:= Y_2(t) - Y_1(t), \tag{1.6}
\end{aligned}
$$

so that, if ε is a Bernoulli variable with probabilities $(1-p,p)$ and t a random variable with uniform distribution in $]0,1[$, we have (in law)

$$X = f(t) + \delta(t)\varepsilon. \tag{1.7}$$

2. That $\inf_{]0,1[} \delta(t) > 0$ is immediate if $Y_2(0) - Y_1(1) = G(1-p+0) - G(1-p) > 0$, which turns out to be the case if X is a Bernoulli itself (choosing p to be its Bernoulli parameter). If X takes at least 3 values, then it is enough to note in full generality that there exists (at least one) $p \in]0,1[$ so that $T_1 > T_2$ where

$$
\begin{aligned}
T_1 &= \quad \inf\{t \in]0,1[: Y_1(t) = G(p_-)\} &\text{(arrival time of } Y_1),\\
T_2 &= \sup\{t \in]0,1[: Y_2(t) = G(1-p+0)\} &\text{(departure time of } Y_2).
\end{aligned}
$$

The latter implies that $\delta(t) > 0$ for all t.

3. Assume now that X satisfies the estimates of Property (H), with points $x_- < x_+$. We set $p_- = \mu(]-\infty, x_-[)$, $p_+ = \mu(]x_+, +\infty[)$. We show that $p = 1-p_-$ is a possible choice. Thanks to (H), $p \geq p_+ > \rho$, and $1-p = p_- > \rho$, so that $p \in]\rho, 1-\rho[$.

We always have $G(1-p) \leq x_- \leq G(1-p+0)$. If $G(1-p) < x_-$ then $\delta(t) \geq x_- - G(1-p) > 0$. Suppose $G(p_-) = x_-$. We claim that $T_1 = 1 > T_2$. It is easy to see that $T_2 \leq (p-p_+)/p < 1$. It remains to show that $T_1 = 1$. Suppose $T_1 < 1$. For any $t \in]T_1, 1[$ and for any $u < x_-$, one has $x_- = G(p_-t) > u$. Then (1.3) implies that $F(u) < p_-t$, and thus we get the following contradiction

$$p_- = \mu(]-\infty, x_-[) = \sup_{u < x_-} F(u) \leq p_-t < p_-. \tag{1.8}$$

4. We set $Y_1(t) = G(\frac{1}{2}t)$ and $Y_2(t) = G(1 - \frac{1}{2}t)$ and proceed as above. Because of Property (H), we have $Y_1(2\rho) \leq x_-$ while $Y_2(2\rho) > x_+$, so that for any $t \leq 2\rho$, $\delta(t) > x_+ - x_-$. $\qquad\square$

2. ANTICHAINS AND SPERNER THEORY

A possible motivation for looking into Sperner theory is the following quite natural question arising in arithmetics, see e.g. [An] and references therein. Consider distinct prime numbers $p_1, \cdots, p_N$ and the integer $M = p_1^{k_1} \cdots p_N^{k_N}$, with $k_i \geq 1$, $i = 1, \cdots, N$. Let $\mathcal{D} \simeq \bigotimes_{i=1}^{N} \{0, 1, \cdots, k_i\}$ be the set of divisors of M. We endow this set with a (discrete) probabilistic structure by considering $\mathbb{P} = \bigotimes_{i=1}^{N} \mu_i$ where for any i, μ_i is a discrete probability measure on $\{0, 1, \cdots, k_i\}$.

Let $\mathcal{A} \subset \mathcal{D}$ be so that for any r, r' in $\mathcal{A}$, neither $r|r'$ nor $r'|r$. Two such elements are said to be "non comparable", and a family of non comparable elements is called an antichain. The question is: what is the maximal size of such a set $\mathcal{A}$? Recasted in probabilistic terms, we would like to provide a bound on $\mathbb{P}(\mathcal{A})$. We first recall some basics from Sperner theory and then give an answer to the question.

We start with the simplest case, that is $k_i = 1$ for all i. The configuration space is $\{0,1\}^N$, and we consider a collection of Bernoulli random variables $\boldsymbol{\eta} = \{\eta_1, ..., \eta_N\}$. The set of configurations is partially ordered by the relation:

$$\boldsymbol{\eta} \prec \boldsymbol{\eta}' \quad \iff \quad \text{for all } i \in \{1, ..., N\}: \quad \eta_i \leq \eta_i'. \tag{2.1}$$

A set $\mathcal{A} \subset \{0,1\}^N$ is said to be an *antichain* if it does not contain any pair of configurations which are comparable in the sense of "$\prec$". The original Sperner's Lemma [Sp] states that for any such set: $|\mathcal{A}| \leq \binom{N}{[\frac{N}{2}]}$. An immediate computation using Stirling formula shows that the latter is bounded by $C2^N/\sqrt{N}$, with e.g.

$C = 2/\sqrt{\pi}$. As a consequence, if $\boldsymbol{\eta}$ is a collection of identical Bernoulli variables with even weigths $(\frac{1}{2}, \frac{1}{2})$, we get as an answer to the problem stated above: $\mathbb{P}(\mathcal{A}) \leq C/\sqrt{N}$ in this particular case. As we shall see, this $1/\sqrt{N}$ behaviour is fairly general.

The next step is to extend the previous bound to non even Bernoulli variables (but still identical). We need a more general result, called the LYM inequality, for antichains (e.g. [An]):

$$\sum_{\boldsymbol{\eta} \in \mathcal{A}} \frac{1}{\binom{N}{|\boldsymbol{\eta}|}} \leq 1 \,, \tag{2.2}$$

where $|\boldsymbol{\eta}| = \sum \eta_j$. The LYM inequality has the following probabilistic implication. If $\{\eta_j\}$ are independent copies of a Bernoulli random variable η with probabilities $(1 - p, p)$, then for any antichain $\mathcal{A} \subset \{0, 1\}^N$:

$$\mathbb{P}(\{\boldsymbol{\eta} \in \mathcal{A}\}) \leq \frac{2\sqrt{2}}{\sigma_\eta \sqrt{N}}, \tag{2.3}$$

where $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_N)$, $\sigma_\eta = \sqrt{pq}$ is the standard deviation of η. The same bound extends to antichains on larger alphabet: $\{0, 1, \cdots, k\}^N$ with $1 \leq k < \infty$ for equidistributed weights [An] as well as for general weights [En1, En2] (more than an upper bound, an asymptotics as N goes to ∞ is proven in those cases).

The following result extends those bounds to non identical measure with (possibly) infinite support. This seems to be a new result in Sperner Theory, which is mentioned in [AGKW, Remark 3.1]. In particular it provides an answer to the problem described as an introduction to this section.

Theorem 2.1. *Set $\mathcal{D} = \mathbb{Z}^N$ and let μ_i, $i = 1, \cdots, N$, be discrete probability measures on $\mathbb{Z}$. Set $\mathbb{P} = \bigotimes_{i=1}^N \mu_i$. Assume there is $\rho \in]0, \frac{1}{2}[$ such that for any $i = 1, \cdots, N$, there exists $m_i \in \mathbb{Z}$ s.t.*

$$\mu_i(] - \infty, m_i]) > \rho \text{ and } \mu_i([m_i + 1, \infty[) > \rho. \tag{2.4}$$

Then there exists $C < \infty$ (independent of N), such that for any antichain $\mathcal{A} \subset \mathcal{D}$,

$$\mathbb{P}(\mathcal{A}) \leq \frac{C}{\sqrt{\rho N}} \,. \tag{2.5}$$

Remark 2.2. *If the μ_i's are Bernoulli measures with support $\{0, 1\}$, then we recover the extension of (2.3) to independent Bernoulli variables, but not necessarily identically distributed, that is proven in [AGKW, Lemma 3.2]; except for the dependency in ρ of the constant, which behaves like ρ^{-1} in [AGKW, Lemma 3.2].*

As an illustration of the Bernoulli decomposition of Section 1, we sketch the proof of Theorem 2.1.

Proof. Let $(X_1, \cdots, X_n)$ be independent integer valued random variables with distribution $\mu_1, \cdots, \mu_N$, respectively. Hypothesis (H) of Section 1 is satisfied, and we can decompose X_i as $f_i(t_i) + \delta_i(t_i)\eta_i$ according to Theorem 1.2 Point 4. Since for all $i = 1, \cdots, N$, $\mathbb{P}_{t_i}(\delta(t_i) \geq 1) \geq 2\rho$, by a large deviation argument, it is enough to estimate $\mathbb{P}(\mathcal{A}; |J_t| \geq \rho N)$, where $J_t = \{i = 1, \cdots, N, s.t. \delta(t_i) \geq 1\}$. Now

$$\mathbb{P}(\mathcal{A}; |J_t| \geq \rho N) = \mathbb{P}_t \mathbb{P}_{(\eta_i)_{i \notin J_t}} \left(\mathbb{P}_{(\eta_i)_{i \in J_t}}(\mathcal{A}; |J_t| \geq \rho N) \right). \tag{2.6}$$

Let $\mathbf{t} \in \{|J_t| \geq \rho N\}$ and $(\eta_i)_{i \notin J_t}$ be given. We have $\delta_i(t_i) > 0$ for all $i \in J_t$, so that two elements $(f_i(t_i) + \delta_i(t_i)\eta_i)_{i \in J_t}$ and $(f_i(t_i) + \delta_i(t_i)\eta_i')_{i \in J_t}$ are comparable

if and only if $(\eta_i)_{i \in J_t}$ and $(\eta_i')_{i \in J_t}$ are comparable. Thus, $\mathbf{t} \in \{|J_t| \geq \rho N\}$ and $(\eta_i)_{i \notin J_t}$ being given, the set of $(\eta_i)_{i \in J_t}$ so that $(f_i(t_i) + \delta_i(t_i)\eta_i)_{i \in \{1, \cdots, N\}} \in \mathcal{A}$ is an antichain. By Sperner's bound

$$\mathbb{P}_{(\eta_i)_{i \in J_t}}(\mathcal{A}; |J_t| \geq \rho N) \leq \frac{2}{\sqrt{\pi |J_t|}} \leq \frac{2}{\sqrt{\pi \rho N}}. \tag{2.7}$$

$\square$

3. Singularity of random matrices

Let $M_n = (a_{ij})_{ij}$ be a random $n \times n$ matrix, where the a_{ij} are independent (non necessarily identically distributed) real random variables. We assume that the random variables a_{ij} satisfy the non-degeneracy property (H) of Section 1 with the same ρ, namely

(H') *There exists $\rho \in]0, \frac{1}{2}[$ such that for any $i, j = 1, \cdots, n$, $\mathbb{P}(a_{ij} > x_{ij}^+) > \rho$ and $\mathbb{P}(a_{ij} < x_{ij}^-) > \rho$ for some real numbers $x_{ij}^- < x_{ij}^+$.*

The note [BrG] provides an elementary proof of the following result.

Theorem 3.1. [BrG] *Let M_n be an $n \times n$ matrix whose coefficients are independent random variables satisfying* (H'). *Then $\mathbb{P}(M_n$ is singular$) \leq C_\rho/\sqrt{n}$, for some $C_\rho < \infty$.*

The study of the singularity of random matrices goes back, at least, to Komlós who showed in [Kom1] that $\mathbb{P}(M_n$ is singular$) = o(1)$ for independent and identically distributed (iid) Bernoulli entries, namely $a_{ij} = 0, 1$ with even probabilities $(\frac{1}{2}, \frac{1}{2})$. Using Sperner's Lemma, Komlós noticed that the probability was $O(n^{-1/2})$ [B], a result which has been further extended in [Sl] to the case of iid entries equally distributed over a finite set. For iid Bernoulli entries, the conjecture is that $\mathbb{P}(M_n$ is singular$) = (c + o(1))^n$ with $c = \frac{1}{2}$, which is the best possible since one clearly has $\mathbb{P}(M_n$ is singular$) \geq 2^{-n}$. Such an exponential behaviour has been successively obtained and improved in [KKoS, TV1, TV2] up to $c = \frac{3}{4}$. The value $c = \frac{1}{2}$ still seems to be out of reach.

If one turns to general entries, Komlós proved in [Kom2] that $\mathbb{P}(M_n$ is singular$) = o(1)$ for independent *and identically distributed* non degenerate random variables. Furthermore, as pointed out by Tao and Vu in [TV1, Section 8], it follows from their analysis that $\mathbb{P}(M_n$ is singular$) = o(1)$ for independent non degenerate entries, provided Property (H') holds. Under the same hypothesis Theorem 3.1 asserts that $\mathbb{P}(M_n$ is singular$) = O(n^{-1/2})$.

In [BrG], Komlós' argument as reproduced in [B] is extented to independent random variables satisfying the condition (H) taking advantage of the Bernoulli decomposition. The proof of Theorem 3.1 thus illustrates how the Bernoulli decomposition may be used in order to extend results known for Bernoulli to the general case of independent non degenerate random variables.

Sperner's bounds enter the proof of such a result through the use of the Littlewood-Offord Lemma [Er]: let $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_N)$ be a collection of independent Bernoulli variables, then for any reals $|\alpha_j| > 0$, $j = 1, \cdots, N$, $\beta \in \mathbb{R}$, $\mathbb{P}(\sum_{j=1}^{N} \alpha_j \eta_j = \beta) \leq C/\sqrt{N}$. Indeed different realizations of $\boldsymbol{\eta}$ can belong to the same hyperplane only if they are non comparable, composing hence an antichain. It is then one ingredient of the proof of Theorem 3.1 to notice that the Littlewood-Offord Lemma

easily extends to general non degenerate random variables thanks to a Bernoulli decomposition.

4. CONCENTRATION BOUNDS

Let X be a real random variable and μ its distribution. Its (Levy) concentration function $Q_X(s)$ (or equivalently the modulus of continuity of its measure μ), is defined as

$$Q_X(s) = \sup_{x \in \mathbb{R}} \mathbb{P}(X \in [x, x+s]) = \sup_{x \in \mathbb{R}} \mu([x, x+s]). \tag{4.1}$$

Bounds on probability of antichains as in Theorem 2.1 find their natural generalization in the following theorem, that deals with some functions of arbitrary non degenerate random variables.

Theorem 4.1. [AGKW] *Let $\boldsymbol{X} = (X_1, \ldots, X_N)$ be a collection of independent random variables whose distributions satisfy, for all $j \in \{1, ..., N\}$:*

$$\mathbb{P}(\{X_j < x_-\}) > \rho \quad and \quad \mathbb{P}(\{X_j > x_+\}) > \rho \tag{4.2}$$

for some $\rho > 0$ and $x_- < x_+$, and $\Phi : \mathbb{R}^N \mapsto \mathbb{R}$ be a function such that for some $\varepsilon > 0$

$$\Phi(\boldsymbol{u} + v\boldsymbol{e}_j) - \Phi(\boldsymbol{u}) > \varepsilon \tag{4.3}$$

for all $v > x_+ - x_-$, all $\boldsymbol{u} \in \mathbb{R}^N$, and $j = 1, \ldots, N$, with $\boldsymbol{e}_j$ the unit vector in the j-direction. Then, the random variable Z, defined by $Z = \Phi(\boldsymbol{X})$, obeys the concentration bound

$$Q_Z(\varepsilon) \leq \frac{C}{\sqrt{\rho N}}, \tag{4.4}$$

for some $C < \infty$ (independent of N).

Remark 4.2. *We ask for a uniform (H) property in the sense that $x_\pm$ are independent of the random variable. Compare to Theorem 2.1 and Theorem 3.1.*

If the random variables are identical Bernoulli variables then the link between Theorem 4.1 and Sperner's theory of antichains is quite obvious. Indeed, let $\boldsymbol{\eta}, \boldsymbol{\eta}'$ be two comparable realizations of $(X_1, \cdots, X_n)$, say $\eta_j \leq \eta'_j$ for all $j = 1, \cdots, N$. Then by (4.3), $\Phi(\boldsymbol{\eta})$ and $\Phi(\boldsymbol{\eta}')$ cannot both belong to a given interval of length ε. In other words, for any given $x \in \mathbb{R}$, realizations of $Z = \Phi(X_1, \cdots, X_n)$ that fall into $[x, x + \varepsilon[$ belong to an antichain; (2.3) above then yields (4.4).

It remains to extend such a reasoning to arbitrary non degenerate random variables. This is achieved by taking advantage of the Bernoulli decomposition, in the spirit of the proof of Theorem 2.1.

Remark 4.3. *1. The use of combinatorial estimates for concentration bounds first appeared in the context of Bernoulli variables in P. Erdös' variant of the Littlewood-Offord Lemma [Er]; that is for linear combinations of Bernoulli variables, improving on W. Doeblin, P. Lévy [DoL, Do]. Further developments for linear functions Φ can be found in A. N. Kolmogorov [Ko], B. A. Rogozin [R1], H. Kesten [Ke] and C. G. Esseen [Es] (see e.g. [R3]).*
2. We stress that in Theorem 4.1, the function Φ needs not to be linear. This is of importance for the application to random Schrödinger operators, where Φ will be an eigenvalue of finite reduction of the Hamiltonians depending on random variables.

5. Application to random Schrödinger operators

In this application, we consider random Schrödinger operators on $L^2(\mathbb{R}^d)$ of the type

$$H_{\boldsymbol{\omega}} = H_{\boldsymbol{\omega}} := -\Delta + V_{\boldsymbol{\omega}}, \qquad (5.1)$$

where Δ is the d-dimensional Laplacian operator, and $V_{\boldsymbol{\omega}}$ is an Anderson-type random potential,

$$V_{\boldsymbol{\omega}}(x) := \sum_{\zeta \in \mathbb{Z}^d} \omega_\zeta \, u(x - \zeta), \qquad (5.2)$$

where

(I) the single site potential u is a nonnegative bounded measurable function on $\mathbb{R}^d$ with compact support, uniformly bounded away from zero in a neighborhood of the origin, more precisely,

$$u_- \chi_{\Lambda_{\delta_-}(0)} \le u \le u_+ \chi_{\Lambda_{\delta_+}(0)} \quad \text{for some constants } u_\pm, \delta_\pm \in]0, \infty[; \qquad (5.3)$$

(II) $\boldsymbol{\omega} = \{\omega_\zeta\}_{\zeta \in \mathbb{Z}^d}$ is a family of independent identically distributed random variables, whose common probability distribution μ is non-degenerate with bounded support, and satisfies

$$\{0, 1\} \in \operatorname{supp} \mu \subset [0, 1]. \qquad (5.4)$$

To fix notations, the set of realizations of the random variables $\{\omega_\zeta\}_{\zeta \in \mathbb{Z}^d}$ is denoted by $\Omega = [0, 1]^{\mathbb{Z}^d}$; $\mathcal{F}$ denotes the σ-algebra generated by the coordinate functions, and $\mathbb{P} = \otimes_{\zeta \in \mathbb{Z}^d} \mu$ is the product measure of the common probability distribution μ of the random variables $\boldsymbol{\omega} = \{\omega_\zeta\}_{\zeta \in \mathbb{Z}^d}$. In other words, we work with the probability space $(\Omega, \mathcal{F}, \mathbb{P}) = \bigotimes_{\zeta \in \mathbb{Z}^d} ([0, 1], \mathcal{B}_{[0,1]}, \mu)$, where $\mathcal{B}_{[0,1]}$ is the Borel σ-algebra on $[0, 1]$. A set $\mathcal{E} \in \mathcal{F}$ will be called an event.

$H_{\boldsymbol{\omega}}$ is a $\mathbb{Z}^d$-ergodic family of random self-adjoint operators. It follows from standard results (cf. [KiMa, Sto2]) that there exists fixed subsets Σ, Σ_{pp}, Σ_{ac} and Σ_{sc} of $\mathbb{R}$ so that the spectrum $\sigma(H_{\boldsymbol{\omega}})$ of $H_{\boldsymbol{\omega}}$, as well as its pure point, absolutely continuous, and singular continuous components, are equal to these fixed sets with probability one.

For $x \in \mathbb{R}^d$, we set $\|x\|$ its sup norm, and $\Lambda_L(x) := \{y \in \mathbb{R}^d; \|y - x\| < \frac{L}{2}\}$ denotes the (open) box of side L centered at $x \in \mathbb{R}^d$. By a box Λ_L we mean a box $\Lambda_L(x)$ for some $x \in \mathbb{R}^d$. By χ_x we denote the characteristic function of the unit box centered at $x \in \mathbb{R}^d$, i.e., $\chi_x := \chi_{\Lambda_1(x)}$.

Localization is proved at the bottom of the spectrum for the Anderson Hamiltonian without any extra hypotheses. Spectral localization is proved in [AGKW] based on an extension of [BoK] given in [GK4]:

Theorem 5.1. *Let $H_{\boldsymbol{\omega}}$ be an Anderson Hamiltonian on $L^2(\mathbb{R}^d)$ as above with hypotheses* (I), (II). *Then there exists $E_0 = E_0(d, u_\pm, \delta_\pm, \mu) > 0$ such that $H_{\boldsymbol{\omega}}$ exhibits spectral localization in the energy interval $[0, E_0]$: with probability one, $\sigma_c(H_\omega) = \emptyset$.*

If one wants more detailed informations about the region of localization, the following result holds, based on the concentration bound given in Theorem 4.1.

Theorem 5.2. [GK4] *Let $H_{\boldsymbol{\omega}}$ be an Anderson Hamiltonian on $L^2(\mathbb{R}^d)$ as above with hypotheses* (I), (II). *Then there exists $E_0 = E_0(d, u_\pm, \delta_\pm, \mu) > 0$ such that*

$H_{\boldsymbol{\omega}}$ *exhibits Anderson localization as well as dynamical localization in the energy interval* $[0, E_0]$. *More precisely:*

- *(Anderson localization) There exists* $m = m(d, V_{\mathrm{per}}, u_\pm, \delta_\pm) > 0$ *such that the following holds with probability one:*
    - $H_{\boldsymbol{\omega}}$ *has pure point spectrum in* $[0, E_0]$.
    - *If* ϕ *is an eigenfunction of* $H_{\boldsymbol{\omega}}$ *with eigenvalue* $E \in [0, E_0]$, *then* ϕ *is exponentially localized with rate of decay* m, *more precisely,*

$$\|\chi_x \phi\| \leq C_{\boldsymbol{\omega},\phi}\, e^{-m|x|} \qquad \textit{for all} \quad x \in \mathbb{R}^d. \tag{5.5}$$

    - *The eigenvalues of* $H_{\boldsymbol{\omega}}$ *in* $[0, E_0]$ *have finite multiplicity.*
- *(Dynamical localization) For all* $s < \frac{3}{8}d$ *we have*

$$\mathbb{E}\left\{ \sup_{t \in \mathbb{R}} \left\| \langle x \rangle^{\frac{m}{2}} e^{-it H_\omega} \chi_{[0,E_0]}(H_\omega) \chi_0 \right\|_2^{\frac{2s}{m}} \right\} < \infty \quad \textit{for all} \ \ m \geq 1. \tag{5.6}$$

Remark 5.3. *For an extension to the non ergodic situation where impurities are located on a given Delone set (instead of a regular lattice), see* [G2].

Remark 5.4. *The same conclusions have been proved to hold for Schrödinger operators with a Poisson random potential* [GHK1, GHK2, GHK3].

The full proof of Theorem 5.2 is presented in [GK4]. In particular it combines the multiscale analysis of Bourgain and Kenig [BoK] together with the concentration bound of [AGKW] (Theorem 4.1 above). This yields Anderson localization (using [GK2] for finite multiplicity). To get dynamical localization, one builds on ideas that are by now standard and that come from [A, DeRJLSi, GDB, G1, DSto, GK1, GK2].

In the one-dimensional case the continuous Anderson Hamiltonian has been long known to exhibit spectral localization in the whole real line for any non-degenerate μ, i.e. when the random potential is not constant [GoMP, KotSi, DSiSt].

In the multidimensional case, localization at the bottom of the spectrum is already known at great, but nevertheless not all-inclusive, generality; cf. [Sto2, K, BoK] and references therein. First proofs of this result are due to Combes Hislop [CH1] and Klopp [Kl2], assuming that the single site probability distribution μ is absolutely continuous with bounded density. The result relies on a multiscale analysis argument "à la" Fröhlich Spencer [FrSp] and adapted from [DrK]'s discrete version; it took more time and quite some efforts to carry the Aizenman Molchanov approach of fractional moments [AM] over the continuum [AENSSt], still under the regularity assumption on μ.

The absolute continuity condition of μ can be relaxed to Hölder continuity of μ, both in the approach based on the multiscale analysis, and in the one based on the fractional moment method. The basis in the former case is an improved analysis of the Wegner estimate, which was first noticed by Stollmann in [Sto1]. Important improvements in Wegner estimates with (not too) singular continous measures μ have then been successively obtained in [CHNa, CHKl1, CHKlR, GKS, HuKiNaStoV] until the recent optimal form due to Combes, Hislop and Klopp [CHKl2]; all theses improved forms provide in particular some continuity property of the integrated density of states.

However, techniques relying on the regularity of μ seem to reach their limit with log-Hölder continuity. In particular, until recently the Bernoulli random potential had been beyond the reach of analysis in more than one dimension. For that

extreme case, i.e., of $H_{\boldsymbol{\omega}}$ with $\mu\{1\} = \mu\{0\} = \frac{1}{2}$, localization at the bottom of the spectrum was recently proven by Bourgain and Kenig [BoK].

In [BoK], the Wegner estimate is obtained along the lines of (an elaborated version of) the multiscale analysis, scale by scale, through a combination of a quantitative unique continuation principle together with a lemma due to Sperner [Sp]. Although it defenitely requires some technical care, it is quite clear from the analysis of [BoK] that the result extends to any measure for which a Sperner's type argument is valid. See for an illustration of this point the note [GK3] where μ is a uniform measure on some Cantor set (μ turns to be log log-Hölder continuous in this example).

Localization was thus proved for the two extreme cases: μ regular enough and μ Bernoulli, and with two different proofs, none of which applying directly to the other case. Our motivation was then to find a single proof for any non degenerated measure, and thus unifying these two extreme results. A key step, the concentration bound extending the Sperner's Lemma estimate used by Bourgain and Kenig, was obtained in [AGKW]. The full technical details of the extension of the multiscale analysis of [BoK] are provided in [GK4].

To fix notations, consider a scale L, $H_{L,\boldsymbol{\omega}}$ a suitable restriction of $H_{\boldsymbol{\omega}}$ to a cube Λ_L of side L with Dirichlet boundary condition, and $R_{L,\boldsymbol{\omega}}(z)$ its resolvent (that is now a compact operator). The spectrum of $H_{L,\boldsymbol{\omega}}$ is thus discrete and given $E \in \sigma(H_{\boldsymbol{\omega}}) = [0, +\infty[$ we want to investigate the size of $\|R_{L,\boldsymbol{\omega}}(E)\|$ and show it is $\leq e^{L^{1-\delta}}$, $\delta > 0$, with probability at least $1 - L^{-p}$, for some $p > 0$ (note that $\|R_{L,\boldsymbol{\omega}}(E)\|$ may be infinite, namely when $E \in \sigma(H_{L,\boldsymbol{\omega}})$, but typically, this should happen for a set of $\boldsymbol{\omega}$'s of small measure. This amounts to analyse the probability that $\mathrm{dist}(E, \sigma(H_{L,\boldsymbol{\omega}})) \geq e^{-L^{1-\delta}}$.

The strong form of the Wegner estimate reads as follows [CHKl2]: there exists $C_W < \infty$, such that for η small enough and L large enough, (recalling (4.1))

$$\mathbb{P}(\mathrm{dist}(E, \sigma(H_{L,\boldsymbol{\omega}})) < \eta) \leq C_W Q_{\omega_0}(2\eta) L^d. \tag{5.7}$$

It is worth pointing out that (5.7) is an a priori estimate that is independent of the existence of localized states. Applying (5.7) with $\eta = e^{-L^{1-\delta}}$ obviously leads to the needed estimate. A weaker version, corresponding to the approach of Bourgain Kenig, reads as follows. Let S be a subset of $D \cap \Lambda_L$, and $\omega_S = (\omega_\zeta)_{\zeta \in S}$. There exists $C_W < \infty$ and $\delta_0 > 0$ s.t., for suitable events $F_{L,\boldsymbol{\omega},S} \subset \mathcal{F}$ coming from the multiscale analysis, for L large enough, $\delta, \varepsilon > 0$ small enough,

$$\mathbb{P}_S(\mathrm{dist}(E, \sigma(H_{L,\boldsymbol{\omega}})) < e^{-L^{1-\delta}}; F_{L,\boldsymbol{\omega},S}) \leq L^\varepsilon Q_Z(2e^{-L^{1-\delta}}), \tag{5.8}$$

where $\mathbb{P}_S = \bigotimes_{\zeta \in S} \mu$ is the restriction of $\mathbb{P}$ to S, $Z = \Phi(\omega_S)$ is a random variable such that for any ω_S, for any $v_\zeta \geq \delta_0$,

$$\Phi(\omega_S + v_\zeta) - \Phi(\omega_S) > 2e^{-L^{1-\delta}}. \tag{5.9}$$

In practice, Φ is an eigenvalue of the finite volume operator, and property (5.9) follows from a quantitative unique continuation principle. Note that unlike what happens in the strong form, it is a only collective effect of the random variables ω_ζ, $\zeta \in S$, that provides some decay. The best universal bound is the following concentration bound coming from Theorem 4.1:

$$Q_Z(2e^{-L^{1-\delta}}) \leq C|S|^{-\frac{1}{2}}. \tag{5.10}$$

In practice, $|S| = L^{\frac{3}{4}d-}$, so that the probability in (5.8) amounts to $L^{-\frac{3}{8}d+}$.

Acknowledgements

It is the pleasure of the author to thank his co-authors, M. Aizenman, L. Bruneau, A. Klein, S. Warzel.

References

[A]   Aizenman, M.: Localization at weak disorder: some elementary bounds. Rev. Math. Phys. **6**, 1163-1182 (1994)

[AENSSt]   Aizenman, M., Elgart, A., Naboko, S., Schenker, J., Stolz, G.: Moment analysis for localization in random Schrödinger operators. Inv. Math. **163**, 343-413 (2006)

[AGKW]   Aizenman, M., Germinet, F., Klein, A., Warzel, S.: On Bernoulli decompositions for random variables, concentration bounds and spectral localization, to appear in Prob. Th. Rel. F.

[AM]   Aizenman, M., Molchanov, S.: Localization at large disorder and extreme energies: an elementary derivation. Commun. Math. Phys. **157**, 245-278 (1993)

[An]   Anderson, I.: *Combinatorics of finite sets*. Corrected reprint of the 1989 edition. Dover Publications Inc. Mineola, NY, 2002.

[B]   B. Bollobas, Random graphs, Academic Press 1985.

[BoK]   Bourgain, J., Kenig, C.: On localization in the continuous Anderson-Bernoulli model in higher dimension, Invent. Math. **161**, 389-426 (2005)

[BrG]   Bruneau, L., Germinet, F.: On the singularity of random matrices with independent entries. To appear in Proc. AMS.

[CH1]   Combes, J.M., Hislop, P.D.: Localization for some continuous, random Hamiltonians in d-dimension. J. Funct. Anal. **124**, 149-180 (1994)

[CHKl1]   Combes, J.M., Hislop, P.D., Klopp, F.: Hölder continuity of the integrated density of states for some random operators at all energies. IMRN **4**, 179-209 (2003)

[CHKl2]   Combes, J.M., Hislop, P.D., Klopp, F.: Optimal Wegner estimate. Duke Math. J. **140**, 469-498 (2007)

[CHKlNa]   Combes, J.M., Hislop, P.D., Klopp, F. Nakamura, S.: The Wegner estimate and the integrated density of states for some random operators. Spectral and inverse spectral theory (Goa, 2000), Proc. Indian Acad. Sci. Math. Sci. **112** , 31-53 (2002)

[CHKlR]   Combes, J.M., Hislop, P.D., Klopp, F, Raikov, G..: Global continuity of the integrated density of states for random Landau Hamiltonians. Comm. Partial Differential Equations. **29**, 1187-1213 (2004)

[CHNa]   Combes, J.M., Hislop, P.D., Nakamura, S.: The L^p-theory of the spectral shift function, the Wegner estimate and the integrated density of states for some random operators. Commun. Math. Phys. **218**, 113-130 (2001)

[DSiSt]   Damanik, D., Sims, R., Stolz, G.: Localization for one dimensional, continuum, Bernoulli-Anderson models. Duke Math. J. **114**, 59-100 (2002)

[DSto]   Damanik, D., Stollmann, P.: Multi-scale analysis implies strong dynamical localization. Geom. Funct. Anal. **11**, 11-29 (2001)

[DeRJLSi] Del Rio, R., Jitomirskaya, S., Last, Y., Simon, B.: Operators with singular continuous spectrum IV: Hausdorff dimensions, rank one pertubations and localization. J. d'Analyse Math. **69**, 153-200 (1996)

[Do]   Doeblin, W.: *Sur les sommes d'un grand nombre de variables aléatoires indépendantes*, Bull. Sci. Math. **63** 23–64 (1939).

[DoL]   Doeblin, W, Lévy, P.: *Calcul des probabilités. Sur les sommes de variables aléatoires indépendantes à dispersions bornées inférieurement*, C.R. Acad. Sci. **202** 2027-2029 (1936).

[DrK]   von Dreifus, H., Klein, A.: A new proof of localization in the Anderson tight binding model. Commun. Math. Phys. **124**, 285-299 (1989)

[En1]   Engel, K.: Optimal representations of partially ordered sets and a limit Sperner theorem. European J. Combin. **7**, 287-302 (1986)

[En2]   Engel, K.: *Sperner theory*. Encyclopedia of Mathematics and its Applications **65**. Cambridge University Press, Cambridge, 1997.

[Er]   Erdös, P.: *On a lemma of Littlewood and Offord*, Bull. Amer. Math. Soc. **51**, 898–902 (1945).

[Es]   Esseen, C. G.: *On the concentration function of a sum of independent random variables* Z. Wahrscheinlichkeitstheorie verw. Geb. **9**, 290-308 (1968).

[FrMaScSp] 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)

[FrSp]   Fröhlich, J., Spencer, T.: Absence of diffusion with Anderson tight binding model for large disorder or low energy. Commun. Math. Phys. **88**, 151-184 (1983)

[G1]   Germinet, F.: Dynamical localization II with an application to the almost Mathieu operator. J. Stat Phys. **95**, 273-286 (1999)

[G2]   Germinet, F.: Recent advances about localization in continuum random Schrödinger operators with an extension to underlying Delone sets. Proceedings of Qmath10 conference, Moeciu 2007. To appear.

[GDB]   Germinet, F., De Bièvre, S.: Dynamical localization for discrete and continuous random Schrödinger operators. Commun. Math. Phys. **194**, 323-341 (1998)

[GHK1]   Germinet, F., Hislop, P., Klein, A.: On localization for the Schrödinger operator with a Poisson random potential. C.R. Acad. Sci. Paris Ser. I **341**, 525-528 (2005)

[GHK2]   Germinet, F., Hislop, P., Klein, A.: Localization for the Schrödinger operator with a Poisson random potential. J. Europ. Math. Soc. 9, 577-607 (2007)

[GHK3]   Germinet, F., Hislop, P., Klein, A.: Localization at low energies for attractive Poisson random Schrödinger operators. Probability and Mathematical Physics: A Volume in Honor of Stanislav Molchanov, CRM Proceedings and Lecture Notes 42, 2007

[GK1]   Germinet, F., Klein, A.: Bootstrap Multiscale Analysis and Localization in Random Media. Commun. Math. Phys. **222**, 415-448 (2001)

[GK2]   Germinet, F., Klein, A.: New characterizations of the region of complete localization for random Schrödinger operators. J. Stat. Phys. **122**, 73-94 (2006)

[GK3]   Germinet, F., Klein, A.: Localization for some Cantor-Anderson Schrödinger operators. *Adventures in Mathematical Physics*, Contemp. Math., vol. 447, 2007

[GK4]   Germinet, F., Klein, A.: Universal occurence of localization for some continuum random Schrödinger operators. In preparation.

[GKS]   Germinet, F, Klein, A., Schenker, J.: Dynamical delocalization in random Landau Hamiltonians. Annals of Math. **166**, 215-244 (2007).

[GoMP]   Gol'dsheid, Ya., Molchanov, S., Pastur, L.: *Pure point spectrum of stochastic one dimensional Schrödinger operators*. Funct. Anal. Appl. **11**, 1–10 (1977).

[HuKiNaStoV] Hundertmark, D., Killip, R., Nakamura, S., Stollmann, P., Veselić, I.: Bounds on the spectral shift function and the density of states. Comm. Math. Phys. 262, 489-503 (2006).

[I]   Ignatiouk-Robert, I.: Martin boundary of a killed random walk on a half-space, J. th. prob. **21**, (2008).

[KKoS]   J. Kahn, J. Komlós, E. Szemerédi, *On the probability that a random ±1 matrix is singular*, J. Amer. Math. Soc. **8**, 223-240 (1995).

[Ke]   Kesten, H.: *A sharper form of the Doeblin-Lévy-Kolmogorov-Rogozin inequality for concentration functions*. Math. Scand. **25**, 133-144 (1969).

[KiMa]   Kirsch, W., Martinelli, F.: On the ergodic properties of the spectrum of general random operators. J. Reine Angew. Math. **334**, 141-156 (1982)

[KiStoSt]   Kirsch, W., Stollman, P., Stolz, G.: Localization for random perturbations of periodic Schrödinger operators. Random Oper. Stochastic Equations **6**, 241-268 (1998)

[K]   Klein, A.: Multiscale analysis and localization of random operators. In *Random Schrodinger operators: methods, results, and perspectives*. Panorama & Synthèse, Société Mathématique de France. To appear.

[Kl2]   Klopp, F.: Localization for continuous random Schrödinger operators. Commun. Math. Phys. **167**, 553-569 (1995)

[Kl3]   Klopp, F.: Internal Lifshits tails for random perturbations of periodic Schrödinger operators. Duke Math. J. **98**, 335-396 (1999)

[Kl4]   Klopp F.: Weak disorder localization and Lifshitz tails: continuous Hamiltonians. Ann. Henri Poincaré **3**, 711-737 (2002)

FRANÇOIS GERMINET

[Ko]   Kolmogorov, A. N.: *Sur les propriétés des fonctions de concentration de M. P. Lévy.* Ann. Inst. H. Poincaré **16**, 27-34 (1958-60).

[Kom1]   J. Komlós, *On the determinant of* $(0,1)$ *matrices*, Studia. Sci. Math. Hungar. **2**, 7-22 (1967).

[Kom2]   J. Komlós, *On the determinant of random matrices*, Studia. Sci. Math. Hungar. **3**, 387-399 (1968).

[KotSi]   Kotani, S., Simon, B.: Localization in general one-dimensional random systems. II. Continuum Schrdinger operators. Comm. Math. Phys. **112**, 103-119 (1987)

[LeMuW]   Leschke, H., Müller, P., Warzel, S.: A survey of rigorous results on random Schrödinger operators for amorphous solids. Markov Process. Related Fields **9**, 729-760 (2003)

[MaSc]   Martinelli, F., Scoppola, E.: Introduction to the mathematical theory of Anderson localization. Riv. Nuovo Cimento **10**, 1-90 (1987)

[McD]   McDonald, D.: *A local limit theorem for large deviations of sums of independent, nonidentically distributed random variables.* Ann. Probab. **7**, 526–531 (1979).

[R1]   Rogozin, B. A.: *An Estimate for Concentration Functions*, Theory Probab. Appl., **6**, 94-97 (1961) Russian original: Teoriya Veroyatnost. i Prilozhen. **6** 103–105 (1961).

[R2]   Rogozin, B. A.: *An integral-type estimate for concentration functions of sums of independent random variables* Dokl. Akad. Nauk SSSR **211**, 1067-1070 (1973) [in Russian].

[R3]   Rogozin, B. A.: Concentration function of a random variable. In *Encyclopaedia of Mathematics*, M. Hazewinkel (ed.), Springer 2002.

[Sl]   A. Slinko, *A generalization of Komlós's theorem on random matrices*, New Zealand J. Math. **30** no. 1, 81-86 (2001).

[Sp]   Sperner, E.: *Ein Satz über Untermengen einer endlichen Menge.* Math. Zeit. **27**, 544-548 (1928).

[Sto1]   Stollmann, P.: Wegner estimates and localization for continuum Anderson models with some singular distributions. Arch. Math. (Basel) **75**, 307-311 (2000)

[Sto2]   Stollmann, P.: *Caught by disorder: bound states in random media.* Boston: Birkhäusser, 2001.

[TV1]   T. Tao, V. Vu, *On Random* ± 1 *matrices: Singularity and Determinant*, Random Structures and Algorithms **28**, 1-23 (2006).

[TV2]   T. Tao, V. Vu, *On the singularity of random Bernoulli matrices*, J. Amer. Math. Soc. **20**, 603-628 (2007).

UNIV CERGY-PONTOISE, CNRS, IUF, DÉPARTEMENT DE MATHÉMATIQUES, F-95000 CERGY-PONTOISE, FRANCE

E-mail address: `francois.germinet@u-cergy.fr`