Ann. Henri Poincaré 26 (2025), 3189–3231

1424-0637/25/093189-43
published online December 12, 2024
https://doi.org/10.1007/s00023-024-01521-3

Annales Henri Poincaré

Typical Macroscopic Long-Time Behavior for Random Hamiltonians

Stefan Teufel, Roderich Tumulka and Cornelia Vogel

Abstract. We consider a closed macroscopic quantum system in a pure state ψ_t evolving unitarily and take for granted that different macro states correspond to mutually orthogonal subspaces $\mathcal{H}_\nu$ (macro spaces) of Hilbert space, each of which has large dimension. We extend previous work on the question what the evolution of ψ_t looks like macroscopically, specifically on how much of ψ_t lies in each $\mathcal{H}_\nu$. Previous bounds concerned the *absolute* error for typical ψ_0 and/or t and are valid for arbitrary Hamiltonians H; now, we provide bounds on the *relative* error, which means much tighter bounds, with probability close to 1 by modeling H as a random matrix, more precisely as a random band matrix (i.e., where only entries near the main diagonal are significantly nonzero) in a basis aligned with the macro spaces. We exploit particularly that the eigenvectors of H are delocalized in this basis. Our main mathematical results confirm the two phenomena of generalized normal typicality (a type of long-time behavior) and dynamical typicality (a type of similarity within the ensemble of ψ_0 from an initial macro space). They are based on an extension we prove of a no-gaps delocalization result for random matrices by Rudelson and Vershynin (Geom Funct Anal 26:1716–1776, 2016).

1. Introduction

We are concerned with an isolated quantum system in a pure state ψ comprising a macroscopically large number of particles. Studying such systems in order to study macroscopic behavior is an approach that is particularly used in connection with the *eigenstate thermalization hypothesis* (ETH) which has been investigated by physicists as well as mathematicians [6,8,9,40]. In particular, the phenomena of equilibration and thermalization in closed quantum

Birkhäuser

systems have attracted increasing attention in recent years [3,4,16–19,21,30–34,38,39,42,43].

We assume that the Hilbert space $\mathcal{H}$ of the system has very large but finite dimension.[1] Following Wigner [46], we model the Hamiltonian as a Hermitian random matrix H. Following von Neumann [45], we regard as given an orthogonal decomposition

$$\mathcal{H} = \bigoplus_{\nu} \mathcal{H}_\nu \tag{1}$$

of the system's Hilbert space $\mathcal{H}$ into subspaces $\mathcal{H}_\nu$ ("macro spaces") associated with different macro states ν [21,23,42,43]. The macroscopic behavior or macroscopic appearance of a vector $\psi \in \mathcal{H}$ is then represented by the weights given by ψ to different macro states ν, i.e., by the sizes of the contributions to ψ from the various $\mathcal{H}_\nu$, $\|P_\nu\psi\|^2$ with P_ν the projection to $\mathcal{H}_\nu$. Here we ask how, under the unitary time evolution $\psi_t = \exp(-iHt)\psi_0$, the macroscopic appearance of ψ_t typically evolves and equilibrates in the long run, more precisely, what $\|P_\nu\psi_t\|^2$ are for typical ψ_0 from a macro space and/or typical large t (where "typical ψ_0" refers to the uniform distribution over the unit sphere; more precise formulations later). Thereby, we extend and refine previous work [43] on the same physical question by means of a deeper mathematical analysis of the behavior of $\|P_\nu\psi_t\|^2$ for typical H (on top of typical ψ_0 and typical t) from suitable random matrix ensembles.

1.1. Normal Typicality

The macroscopic appearance of a system with $\psi \in \mathcal{H}_\nu$ is summarized by the macro state ν; of course, a general state is a superposition of contributions from different macro spaces, which makes the superposition weights $\|P_\nu\psi\|^2$ relevant. It can be useful to compare a given state ψ to a *purely random* vector ϕ, i.e., one that is uniformly distributed over the unit sphere $\mathbb{S}(\mathcal{H}) = \{\psi \in \mathcal{H} : \|\psi\| = 1\}$; ϕ has, with probability near 1, superposition weights approximately proportional to the dimension of $\mathcal{H}_\nu$,

$$\|P_\nu\phi\|^2 \approx \frac{d_\nu}{D}, \tag{2}$$

provided that $d_\nu := \dim \mathcal{H}_\nu$ and $D := \dim \mathcal{H}$ are large; see, e.g., Lemma 1 in [19] and the references therein. Such a behavior is sometimes called "normal," in analogy to the concept of a normal number [28].

In the cases that von Neumann considered in [45], it turned out that every $\psi_0 \in \mathbb{S}(\mathcal{H})$ evolves so that for most $t \in [0, \infty)$,

$$\|P_\nu\psi_t\|^2 \approx \frac{d_\nu}{D} \quad \forall \nu \tag{3}$$

[1] The physical background is that $\mathcal{H}$ is really the subspace of the full Hilbert space corresponding to a micro-canonical energy interval $[E - \Delta E, E]$ with ΔE the resolution of macroscopic energy measurements, and this interval contains a very large but finite number of eigenvalues of the Hamiltonian, realistically of the order $10^{10^{10}}$. However, in this paper we will not assume that all eigenvalues of H on $\mathcal{H}$ lie between $E - \Delta E$ and E.

if d_ν and D are sufficiently large. This behavior is called "normal typicality" and more elaborated in [19,21].

However, the cases von Neumann considered are not very realistic: His conditions on H are true with probability near 1 if the eigenbasis of H is chosen purely randomly (i.e., uniformly) among all orthonormal bases (and some further technical conditions that are not very restrictive). This can be regarded as expressing that the energy eigenbasis is *unrelated* to the orthogonal decomposition (1). In this case, the system would very rapidly go from any initial macro space $\mathcal{H}_\mu$ to the thermal equilibrium macro space $\mathcal{H}_{\rm eq}$ [16–18], which is unphysical.[2] That is why we are interested in generalizations of normal typicality that apply also to Hamiltonians whose eigenbasis is not unrelated to the decomposition (1).

In such more general scenarios, we showed in [43] that the following *generalized normal typicality* still holds: for most ψ_0 from the unit sphere $\mathbb{S}(\mathcal{H}_\mu)$ in the macro space associated with a (possibly nonequilibrium) macro state μ and most $t \in [0, \infty)$,

$$\|P_\nu \psi_t\|^2 \approx M_{\mu\nu} \quad \forall \nu \tag{4}$$

for suitable values $M_{\mu\nu}$, in fact (for nondegenerate H)

$$M_{\mu\nu} = \frac{1}{d_\mu} \sum_n \langle \phi_n | P_\mu | \phi_n \rangle \langle \phi_n | P_\nu | \phi_n \rangle \tag{5}$$

with $\phi_1, \dots, \phi_D$ an orthonormal eigenbasis of H. Put another way, generalized normal typicality means that the superposition weights $\|P_\nu \psi_t\|^2$ approach certain stable values and stay close to them for most times; we also say that the system *equilibrates*. See Figs. 1 and 2 for a numerical example.

One of our goals in this paper is to strengthen the results for suitable types of random Hamiltonians. We are particularly interested in random matrices with band structure (see Fig. 3) in a basis that diagonalizes the projections onto the macro spaces. This band structure makes it less likely that a state in a small macro space (far away from equilibrium) goes directly into the thermal equilibrium macro space without passing through some other macro spaces in between.

We showed in [43] that for general Hamiltonians under suitable assumptions the *absolute* errors

$$\big| \|P_\nu \psi_t\|^2 - M_{\mu\nu} \big| \tag{6}$$

are small (see Theorems 1, 2 in Sect. 2). However, if $d_\nu \ll D$ then we expect that $\|P_\nu \psi_t\|^2 \ll 1$ and $M_{\mu\nu} \ll 1$. Then, the absolute error would be small indeed, but this would not imply that $\|P_\nu \psi_t\|^2 / M_{\mu\nu} \approx 1$. Therefore, we want to show in the following that for suitable random Hamiltonians the *relative* errors

$$\frac{\big| \|P_\nu \psi_t\|^2 - M_{\mu\nu} \big|}{M_{\mu\nu}} \tag{7}$$

[2] Thermal equilibrium requires that energy (and other quantities) is rather evenly distributed over all degrees of freedom, and for getting an even distribution, it needs to get transported through space, which requires time and passage through other macro states.

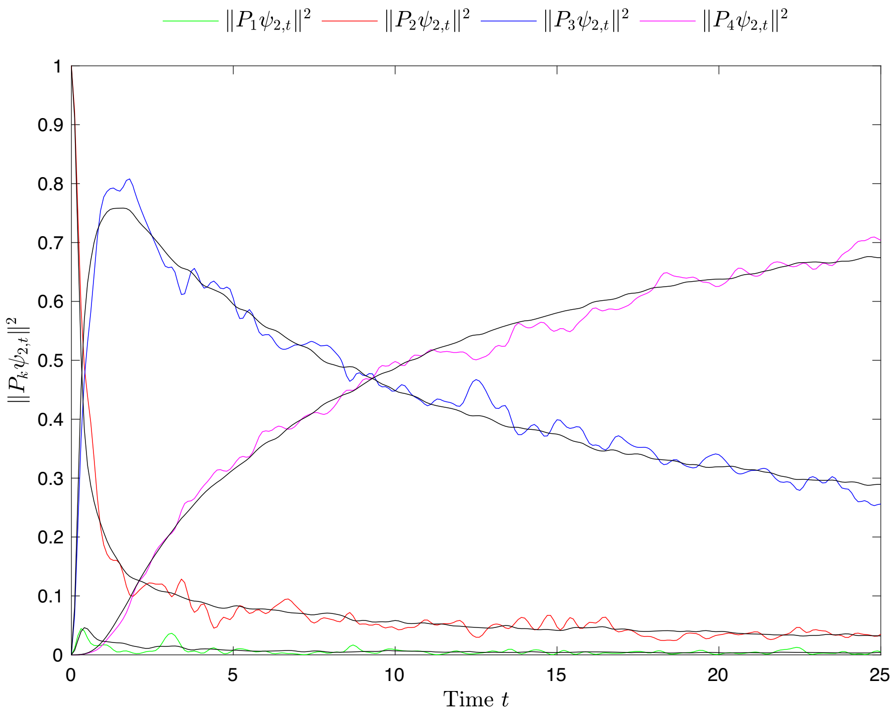

FIGURE 1. Numerical simulation of the functions $t \mapsto \|P_\nu \psi_t\|^2$ as detailed in Appendix B. A Hilbert space of dimension $D = 2222$ is decomposed into 4 macro spaces of dimensions $d_1 = 2$ (green curve), $d_2 = 20$ (red curve), $d_3 = 200$ (blue curve), and $d_4 = 2000$ (purple curve). At large times, see also Fig. 2, the equilibrium subspace $\mathcal{H}_4$ has the biggest contribution to ψ_t. The initial wave function ψ_0 was chosen purely randomly from the unit sphere $\mathbb{S}(\mathcal{H}_2)$, and the Hamiltonian is a random band matrix in a basis aligned with the macro spaces with a band that is wide enough such that the eigenfunctions are still delocalized. Then, parts of ψ_t reach $\mathcal{H}_4$ only after passing through the macro space $\mathcal{H}_3$; in this example, the blue curve increases first before it decreases and the purple curve increases. The black curves are the deterministic approximations $w_{2\nu}(t)$, see (34) according to dynamical typicality. (This figure already appeared in [43])

are small as well. This is the first goal of this paper: to show for certain tractable models and under suitable conditions on the d_ν that (4) holds in the sense that even the *relative* error is small. And this goal is indeed highly relevant, since we expect that $d_\nu \ll D$ for *all* nonequilibrium macro spaces $\mathcal{H}_\nu$.

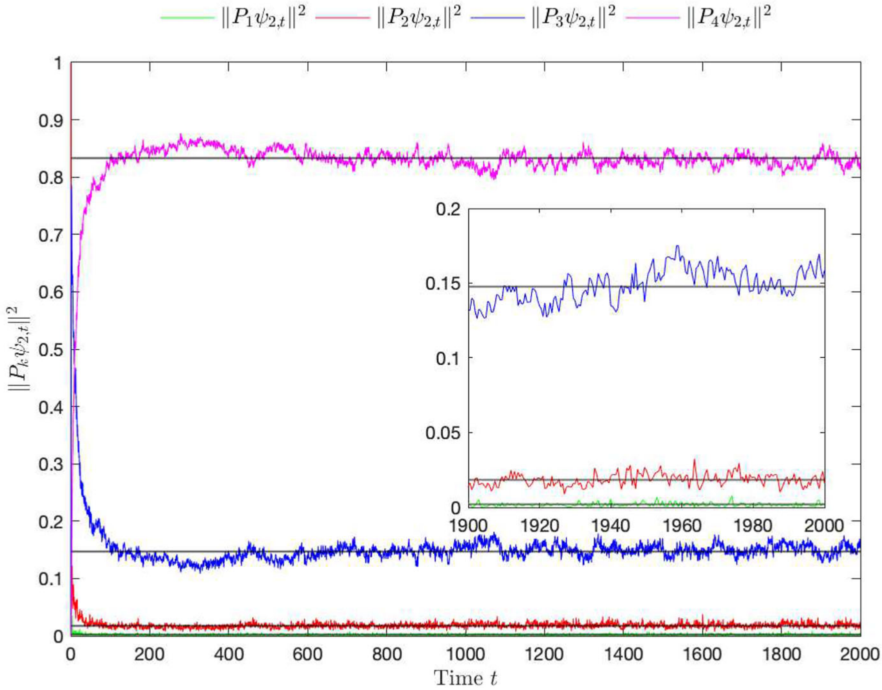

FIGURE 2. The same simulation as in Fig. 1, only for larger times. The inset shows a part of the figure in magnification. In the long run, the curves $t \mapsto \|P_\nu \psi_t\|^2$ reach the values $M_{\mu\nu}$ (indicated by the black horizontal lines) and stay close to them, up to either small or rare fluctuations. (This figure also already appeared in [43])

Rigorous statements are formulated in Sect. 3: In Theorem 4, we provide a lower bound on $M_{\mu\nu}$ under the assumption that $H = H_0 + V$ where H_0 is any deterministic Hermitian matrix and V is a Gaussian random matrix with entries that are independent up to Hermitian symmetry and have variances bounded away from 0 (so also far from the main diagonal, entries of H cannot be *too* small). From this lower bound and the upper bound on the absolute error provided in [43] (and recapitulated in Sect. 2), we obtain an upper bound on the relative error in Corollary 2, valid with high probability (i.e., for most H) for most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ and most $t \in [0, \infty)$. That is, the upshot is that for typical H from the ensemble considered, $\|P_\nu \psi_t\|^2$ is nearly independent of ψ_0 and in the long run nearly independent of t—it equilibrates.

This equilibration is related to the question of the increase in the quantum Boltzmann entropy observable

$$S_B := \sum_\nu S(\nu) P_\nu \tag{8}$$

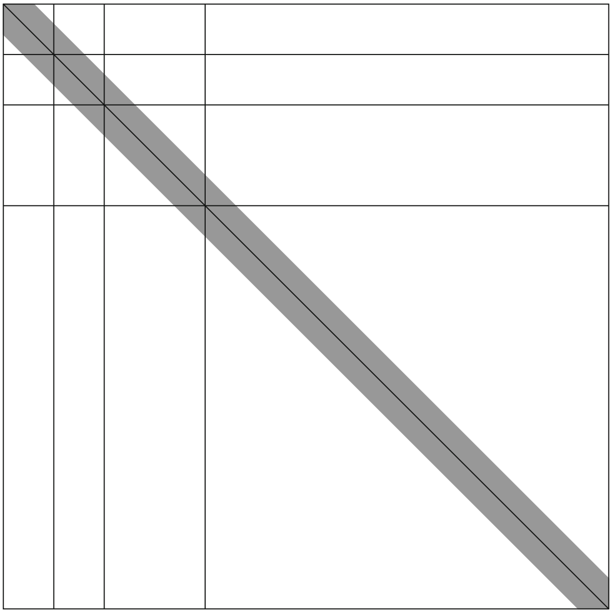

Figure 3. A band matrix has significantly nonzero entries only in the gray region near the main diagonal. The first macro space is spanned by the first d_1 basis vectors, the second macro space by the next d_2 basis vectors, etc.; in the figure, a line is drawn after the first d_1 columns and the first d_1 rows, etc.

with entropy values

$$S(\nu) := k_B \log(d_\nu)\,, \tag{9}$$

where k_B is the Boltzmann constant. Here is how: Physically, in every energy shell corresponding to a small energy interval $[E-\Delta E, E]$, there usually is one macro space $\mathcal{H}_\nu$, the one corresponding to thermal equilibrium, that has the overwhelming majority of dimensions, $\frac{d_\nu}{D} \approx 1$ [20]; we will denote it by $\mathcal{H}_{\mathrm{eq}}$. Now if normal typicality holds as in (3), or already if the $M_{\mu\nu}$ in (4) tend to be bigger for ν with bigger dimensions d_ν, then an initial state $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ from a nonequilibrium (i.e., comparatively small) macro space $\mathcal{H}_\mu$ will evolve so that at most times t, the biggest contribution to ψ_t lies in $\mathcal{H}_{\mathrm{eq}}$, and if $\|P_{\mathrm{eq}}\psi_t\|^2 \approx 1$, this implies further that $\|S_B\psi_t\| \approx S(\mathrm{eq}) \gg S(\mu) = \|S_B\psi_0\|$, that is, the state has evolved to one with higher quantum Boltzmann entropy.

In order to obtain lower bounds for the $M_{\mu\nu}$ (and therefore upper bounds for the relative errors), we need lower bounds on $\|P_\nu\phi_n\|$, where (ϕ_n) is again an orthonormal eigenbasis of H.

We take as fixed an orthonormal basis $(|j\rangle)_{j=1}^D$ of $\mathcal{H}$ that diagonalizes the P_ν (i.e., is such that each $|j\rangle$ lies in some $\mathcal{H}_\nu$). When we talk of H as a matrix, we refer to this basis. In terms of this basis,

$$\|P_\nu \phi_n\|^2 = \sum_{j\in I_\nu} |\langle \phi_n|j\rangle|^2 \tag{10}$$

with I_ν the appropriate subset of $[D] := \{1,\dots,D\}$. In order to obtain a lower bound on this quantity we ask whether ϕ_n (expressed as an element of $\mathbb{C}^D$ relative to the basis $(|j\rangle)_j$) is localized or delocalized.

Our first main result, Theorem 4, is based on results by Rudelson and Vershynin [37] about the *delocalization of eigenvectors* of a random matrix. While the delocalization of eigenvectors of different types of random matrices has been studied intensively in the literature in the recent years, e.g., [1,5,7,10–15,35,47], most results were concerned with delocalization in the sup-norm, meaning that

$$\|\phi_n\|_\infty = \sup_{j=1}^{D} |\langle \phi_n|j\rangle| \tag{11}$$

cannot be too large, so many entries of ϕ_n must be non-negligible. However, that would still allow a negligible $\|P_\nu\phi_n\|$ for some ν. Thus, for a lower bound on (10), we need that only few entries are negligible. For this kind of reasons, results about the sup-norm only yield upper bounds for the $M_{\mu\nu}$ and, in very special cases, lower bounds for some of the $M_{\mu\nu}$; for example, in Sect. 6 we show that with a sup-norm delocalization result from Ajanki, Erdős, and Krüger [1] we obtain useful lower bounds (Theorem 13) only if μ or ν is the thermal equilibrium macro state and d_{eq} is extremely large. The result from Rudelson and Vershynin [37], however, concerns another aspect of the delocalization of eigenvectors: it rules out gaps, i.e., no significant fraction of the coordinates of an eigenvector can carry only a negligible fraction of its mass. Their result gives lower bounds on (10) and enables us to prove nontrivial lower bounds for all $M_{\mu\nu}$. Unfortunately, the lower bounds for $M_{\mu\nu}$ obtained in this way seem to be far from optimal. More precisely, as long as one is in the delocalized regime, one expects that actually $M_{\mu\nu} \approx \frac{d_\nu}{D}$, as for normal typicality, even though the eigenvectors are not uniformly distributed over the sphere. (A matrix with uniformly chosen eigenvectors will be unlikely to have band structure, but a band matrix can very well have $M_{\mu\nu} \approx \frac{d_\nu}{D}$, as the simple example of the discrete Laplacian in 1d shows.) The bounds obtained by Rudelson and Vershynin were recently improved for matrices with independent entries [25,26] but since we are interested in Hermitian random matrices (since they can serve as Hamiltonians), these improved results are not applicable in our situation.

Our theorems in this paper are proved for a general deterministic Hermitian matrix B instead of a projection P_ν, in parallel to previous works such as [34,38,39]; in combination with the results from [43] we can also obtain a finite-time result (that concerns most $t\in[0,T]$ instead of most $t\in[0,\infty)$).

Sharper estimates for all $M_{\mu\nu}$ can be obtained for certain ensembles of Hermitian random matrices for which stronger statements about the delocalization of eigenvectors are available. However, such statements are apparently not currently available for random band matrices, but only for matrices that have very substantially nonzero entries also far from the diagonal. Nevertheless, we report here also what would follow about $M_{\mu\nu}$ for those matrices, based on two results in the literature. First, in Sect. 6, we use a delocalization result of Ajanki, Erdős, and Krüger [1] to obtain that $M_{\mu\nu} \approx d_\nu/D$ in case $\nu = \text{eq}$ or $\mu = \text{eq}$. Second, in Sect. 7, we consider a version of the ETH formulated by Cipolloni, Erdős, and Henheik [6] and show that it implies that $M_{\mu\nu} \approx d_\nu/D$ for *all* μ and ν.

1.2. Dynamical Typicality

In [43], we were not only concerned with generalized normal typicality but also with *dynamical typicality*, which means that for any given t, $\|P_\nu\psi_t\|^2$ is nearly independent of $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$. In Fig. 1, this phenomenon is visible in the proximity of the colored (exact) curves to the black ones. Here, we also provide an improved result about dynamical typicality for random H.

The name "dynamical typicality" was introduced by Bartsch and Gemmer [4] for the following phenomenon: Given an observable A and some value $a \in \mathbb{R}$, there is a function $a(t)$ such that for every $t \in \mathbb{R}$ and most initial wave functions $\psi_0 \in \mathbb{S}(\mathcal{H})$ with $\langle\psi_0|A|\psi_0\rangle \approx a$, it holds that $\langle\psi_t|A|\psi_t\rangle \approx a(t)$. A rigorous version of this fact was proved by Müller, Gross and Eisert [27] who considered initial wave functions $\psi_0 \in \mathbb{S}(\mathcal{H})$ with $\langle\psi_0|A|\psi_0\rangle = a$ for a given $a \in \mathbb{R}$. Reimann [32,33] argued that one also finds for most $\psi_0 \in \mathbb{S}(\mathcal{H})$ with $\langle\psi_0|A|\psi_0\rangle \approx a$ that $\langle\psi_t|B|\psi_t\rangle \approx b(t)$ for another observable B and suitable (ψ_0-independent) $b(t)$. For technical reasons, these results do not cover the case that A is a projection and $a = 1$. In [3], however, a quite general dynamical typicality result was proven that can also be applied to projections, and in [43] we provided a simple proof for this special case. The result for projections can also be obtained as a special case of Eq. (13) of Reimann [29] (applied to a Gaussian distribution with covariance P_μ, $K = 1$, and $A = e^{iHt}P_\nu e^{-iHt}$).

As for generalized normal typicality, we only obtained bounds for the *absolute* errors in [43] (recapitulated as Theorem 3 in Sect. 2). Unfortunately, we cannot provide here an upper bound on the *relative* errors either; but with the help of our improved version of the no-gaps delocalization result of Rudelson and Vershynin [37], we are able to bound what we call the *comparative error*, which is the absolute error divided by the time average, i.e.,

$$\frac{\left|\|P_\nu\psi_t\|^2 - \mathbb{E}_\mu\|P_\nu\psi_t\|^2\right|}{\overline{\|P_\nu\psi_t\|^2}}, \tag{12}$$

where $\mathbb{E}_\mu$ means the average over initial wave functions from the unit sphere in $\mathcal{H}_\mu$ and the overbar means time average over $[0,\infty)$. In Sect. 3.2, we formulate a rigorous result about (12) as Theorem 5. It provides an upper bound on (12) valid with high probability for random H (distributed as before) for every t and most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$. If the constants, in particular d_μ, d_ν, and D, are such that this bound is small, then this means that $\|P_\nu\psi_t\|^2$ is nearly deterministic. Moreover, Theorem 5 says that also the whole function $t \mapsto \|P_\nu\psi_t\|^2$ is close to the function $t \mapsto \mathbb{E}_\mu\|P_\nu\psi_t\|^2$ on any interval $[0,T]$ in the L^2 norm.

1.3. Contents

The remainder of this paper is organized as follows: In Sect. 2, we introduce some notation and recall our previous results from [43]. In Sect. 3, we formulate our main results. In Sect. 4, we describe the result of Rudelson and Vershynin [37] that we use (with minor corrections, see Theorem 6) and point out which Gaussian random matrices will satisfy their hypotheses (Theorems 8, 9). In Sect. 5, we prove our main results after formulating them in a somewhat more general version (Theorems 10, 11) that applies to arbitrary operators B instead of P_ν. In Sect. 6, we prove with the help of a delocalization result from Ajanki, Erdős, and Krüger [1] (quoted here as Theorem 12) some improved lower bounds for the $M_{\mu\nu}$ that are nontrivial if μ or ν is the thermal equilibrium macro state (Theorem 13). In Sect. 7 we show that sharper estimates for all $M_{\mu\nu}$ follow from the version of the ETH due to Cipolloni, Erdős, and Henheik [6]. In Appendix A, we include the proof that, with probability 1, a random H with continuous distribution has no degeneracies or resonances. In Appendix B, we give more detail about our numerical examples.

2. Prior Results

We quote here the main results of [43] about the absolute errors as Theorems 1, 2, and 3 alongside some related statements. We start with a few definitions and in particular make precise some notions that appeared in the introduction:

Definition 1. Most ψ, most t and time averages. Suppose that for each $\psi \in \mathbb{S}(\mathcal{H})$, the statement $s(\psi)$ is either true or false, and let $\varepsilon > 0$. We say that $s(\psi)$ is true for $(1-\varepsilon)$-*most* $\psi \in \mathbb{S}(\mathcal{H})$ if and only if

$$u\left(\{\psi \in \mathbb{S}(\mathcal{H}) : s(\psi)\}\right) \geq 1-\varepsilon, \tag{13}$$

where u is the normalized uniform measure over $\mathbb{S}(\mathcal{H})$.[3]

Suppose that for each $t \in [0,\infty)$, the statement $s(t)$ is either true or false, and let $T, \delta > 0$. We say that $s(t)$ is true for $(1-\delta)$-*most* $t \in [0,T]$ if and only if

$$\frac{1}{T}\lambda\left(\{t \in [0,T] : s(t)\}\right) \geq 1-\delta, \tag{14}$$

where λ denotes the Lebesgue measure on $\mathbb{R}$. Moreover, we say that $s(t)$ is true for $(1-\delta)$-*most* $t \in [0,\infty)$ if and only if

$$\liminf_{T\to\infty} \frac{1}{T}\lambda\left(\{t \in [0,T] : s(t)\}\right) \geq 1-\delta. \tag{15}$$

For any function $f : [0,\infty) \to \mathbb{C}$, define its *time average* as

$$\overline{f(t)} := \lim_{T\to\infty} \frac{1}{T}\int_0^T f(t)\, dt \tag{16}$$

whenever the limit exists.

[3] We do not consider here other ensembles of wave functions such as GAP [44] or the class of ensembles considered by Reimann in [29].

In the following, we consider Hamiltonians with spectral decomposition

$$H = \sum_{e\in\mathcal{E}} e\,\Pi_e, \tag{17}$$

where $\mathcal{E}$ is the set of distinct eigenvalues of H and Π_e the projection onto the eigenspace of H with eigenvalue e.

Definition 2. Relevant properties of the Hamiltonian. Let $\kappa > 0$. We define the *maximum degeneracy* of an eigenvalue as

$$D_E := \max_{e\in\mathcal{E}} \operatorname{tr}(\Pi_e) \tag{18}$$

and the *maximal number of gaps in an interval of length* κ as

$$G(\kappa) := \max_{E\in\mathbb{R}} \#\left\{(e,e') \in \mathcal{E}\times\mathcal{E} : e \neq e' \text{ and } e - e' \in [E, E+\kappa)\right\}. \tag{19}$$

Moreover, we define the *maximal gap degeneracy* as

$$D_G := \lim_{\kappa\to 0^+} G(\kappa). \tag{20}$$

Definition 3. Asymptotic superposition weights. Let $B \in \mathcal{L}(\mathcal{H})$. Given any $\psi \in \mathbb{S}(\mathcal{H})$ and any macro state μ, define

$$M_{\psi B} := \sum_e \operatorname{tr}\left(|\psi\rangle\langle\psi|\Pi_e B \Pi_e\right), \tag{21}$$

$$M_{\mu B} := \frac{1}{d_\mu}\sum_e \operatorname{tr}(P_\mu \Pi_e B \Pi_e). \tag{22}$$

In the special case that $B = P_\nu$ for some macro state ν we define

$$M_{\psi\nu} := M_{\psi P_\nu} \tag{23}$$

and

$$M_{\mu\nu} := M_{\mu P_\nu}. \tag{24}$$

In [43], we proved the following theorem concerning the absolute errors:

Theorem 1. *Let $B \in \mathcal{L}(\mathcal{H})$, let $\varepsilon, \delta, \kappa, T > 0$ and let μ be an arbitrary macro state. Then, $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that for $(1-\delta)$-most $t \in [0, T]$*

$$\left|\langle\psi_t|B|\psi_t\rangle - M_{\mu B}\right| \leq 4\left(\frac{D_E G(\kappa)\|B\|}{\delta\varepsilon d_\mu}\left(1 + \frac{8\log_2 d_E}{\kappa T}\right)\min\left\{\|B\|, \frac{\operatorname{tr}(|B|)}{d_\mu}\right\}\right)^{1/2}, \tag{25}$$

where D_E is the maximum degeneracy of an eigenvalue and D_G is the maximum degeneracy of an eigenvalue gap of H.

By setting $B = P_\nu$ for some macro state ν, choosing κ small enough such that $G(\kappa) = D_G$ and then taking the limit $T \to \infty$ we immediately obtain:

Theorem 2 (Generalized normal typicality: absolute errors). *Let $\varepsilon, \delta > 0$ and let μ, ν be two arbitrary macro states. Then, $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that for $(1-\delta)$-most $t \in [0, \infty)$*

$$\left|\|P_\nu\psi_t\|^2 - M_{\mu\nu}\right| \leq 4\left(\frac{D_E D_G}{\delta\varepsilon d_\mu}\min\left\{1, \frac{d_\nu}{d_\mu}\right\}\right)^{1/2}. \tag{26}$$

Theorem 2 tells us, roughly speaking, that as soon as

$$d_\mu \gg D_E D_G, \tag{27}$$

i.e., as soon as the dimension of $\mathcal{H}_\mu$ is huge and no eigenvalues or gaps are macroscopically degenerate, the superposition weight $\|P_\nu \psi_t\|^2$ will be close to the fixed value $M_{\mu\nu}$ for most times t and most initial states $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$.

Of course, the fixed value $M_{\mu\nu}$, or more generally $M_{\mu B}$, can be computed by taking the average of over time and over the sphere in $\mathcal{H}_\mu$. This is the content of the following well known proposition, see also [43].

Proposition 1. *Let $B \in \mathcal{L}(\mathcal{H})$. Then*

$$M_{\psi_0 B} = \overline{\langle \psi_t | B | \psi_t \rangle} \text{ and} \tag{28}$$

$$M_{\mu B} = \int_{\mathbb{S}(\mathcal{H}_\mu)} M_{\psi_0 B}\, u_\mu(d\psi_0) = \mathbb{E}_\mu M_{\psi_0 B}, \tag{29}$$

where u_μ is the normalized uniform measure over $\mathbb{S}(\mathcal{H}_\mu)$.

This motivates us to call $M_{\psi\nu}$ the *time average superposition weight* in $\mathcal{H}_\nu$ and $M_{\mu\nu}$ the *full average superposition weight* in $\mathcal{H}_\nu$.

For a system with N particles or more generally N degrees of freedom the dimension D is of order $\exp(N)$. For any macro state μ we define s_μ, the entropy per particle in the macro state μ, by

$$d_\mu =: \exp(s_\mu N / k_B), \tag{30}$$

where k_B is the Boltzmann constant.

With (30) for the dimensions of the macro spaces, we find the following corollary of Theorem 2 whose proof can also be found in [43]:

Corollary 1. *Assume (30) and let $\varepsilon, \delta > 0$. Then, for all macro states μ, ν_-, ν_+ with*

$$s_{\nu_-} \leq s_\mu \leq s_{\nu_+} \tag{31}$$

it holds for $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ for $(1-\delta)$-most $t \in [0,\infty)$

$$\left| \|P_{\nu_+}\psi_t\|^2 - M_{\mu\nu_+} \right| \leq \frac{4\sqrt{D_E D_G}}{\sqrt{\varepsilon\delta}} \exp\left(-\frac{s_\mu N}{2k_B} \right), \tag{32}$$

$$\left| \|P_{\nu_-}\psi_t\|^2 - M_{\mu\nu_-} \right| \leq \frac{4\sqrt{D_E D_G}}{\sqrt{\varepsilon\delta}} \exp\left(-\frac{\left(s_\mu - \frac{s_{\nu_-}}{2}\right) N}{k_B} \right). \tag{33}$$

Corollary 1 shows that fluctuations of the time-dependent superposition weights around their expected values are exponentially small in the number of particles, provided that no eigenvalues and gaps are macroscopically degenerate.

In addition to the theorem concerning generalized normal typicality, we also proved a result concerning dynamical typicality in [43]:

Theorem 3. *Let μ be an arbitrary macro state and let B be any operator on $\mathcal{H}$. Let $w_{\mu B} : \mathbb{R} \to [0, 1]$ be the function defined by*

$$w_{\mu B}(t) := \frac{1}{d_\mu} \operatorname{tr}\left[P_\mu \exp(iHt)B \exp(-iHt)\right] \tag{34}$$

Then, for every $t \in \mathbb{R}$ and every $\varepsilon > 0$, $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that

$$\left|\langle\psi_t|B|\psi_t\rangle - w_{\mu B}(t)\right| \leq \min\left\{\frac{\|B\|}{\sqrt{\varepsilon d_\mu}}, \sqrt{\frac{\|B\| \operatorname{tr}(|B|)}{\varepsilon d_\mu^2}}, \sqrt{\frac{18\pi^3 \log(4/\varepsilon)}{d_\mu}}\|B\|\right\}. \tag{35}$$

Moreover, for every μ and B, every $T > 0$, and $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$,

$$\frac{1}{T}\int_0^T \left|\langle\psi_t|B|\psi_t\rangle - w_{\mu B}(t)\right|^2 dt \leq \frac{\|B\|^2}{\varepsilon d_\mu}. \tag{36}$$

We note first that $w_{\mu B}(t)$ is exactly the ensemble average of $\langle\psi_t|B|\psi_t\rangle$, i.e., the average over $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$:

$$w_{\mu B}(t) = \mathbb{E}_\mu\langle\psi_t|B|\psi_t\rangle\,. \tag{37}$$

In particular, by Fubini's theorem,

$$\overline{w_{\mu B}(t)} = M_{\mu B}\,. \tag{38}$$

Equation (35) offers three estimates that can be useful depending on the sizes of B, ε, and d_μ. Specifically, if $\|B\|$ is of order 1 and $d_\mu \gg 1/\varepsilon$, then the first estimate implies that $\langle\psi_t|B|\psi_t\rangle$ is close to its average $w_{\mu B}(t)$. We note further that the first two estimates in (35) can be obtained by similar arguments as in the proof of Theorem 1 whereas the third estimate is a consequence of Lévy's lemma (see, e.g., [42, Sec. II.C]).

In the following, we can assume without loss of generality $D_E = D_G = 1$ since we consider random matrices whose joint distribution of their entries is absolutely continuous with respect to the Lebesgue measure. The set of matrices with degenerate eigenvalues or eigenvalue gaps—we also say that the matrix has *resonances*—has Lebesgue measure zero and therefore, with probability 1, $D_E = D_G = 1$. For the convenience of the reader, we give the proof of this statement in Appendix A.

3. Main Results

In this section, we present and discuss our main results concerning lower bounds for the $M_{\mu\nu}$ and therefore generalized normal typicality (Theorem 4) as well as a corollary thereof bounding the relative errors (Corollary 2) and a result bounding the comparative error for dynamical typicality (Theorem 5). In all the results presented in this section, it is assumed that the Hamiltonian is of a special form and we only consider projections P_ν onto a macro space $\mathcal{H}_\nu$. More general results and the proofs can be found in Sect. 5.

3.1. Generalized Normal Typicality

We make the following assumption:

Assumption 1. The Hamiltonian H is of the form $H = H_0 + V$, where H_0 is a (deterministic) Hermitian $D \times D$ matrix and V is a Hermitian random Gaussian perturbation, more precisely, $V = \frac{1}{\sqrt{2}}(A + A^*)$, where $A = (a_{ij})$ is a random $D \times D$ matrix with independent Gaussian entries with mean zero, i.e., all random variables $\operatorname{Re} a_{ij}, \operatorname{Im} a_{ij}, i, j \in [D]$, are independent and $\operatorname{Re} a_{ij}, \operatorname{Im} a_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2/2)$ for some $\sigma_{ij} > 0$ that fulfill $\sigma_{ij} = \sigma_{ji}$.

Note that because of Assumption 1 the matrix V is Hermitian and Gaussian, more precisely, $\operatorname{Re} v_{ij}, \operatorname{Im} v_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2)$ for $i \neq j$ and $v_{ii} = \operatorname{Re} v_{ii} \sim \mathcal{N}(0, \sigma_{ii}^2)$, i.e., V is a Hermitian Gaussian Wigner-type matrix.

For a Hamiltonian as in Assumption 1, we define

$$C_{H_0} := D^{-1/2} \max\{\|\operatorname{Re} H_0\|, \|\operatorname{Im} H_0\|\}. \tag{39}$$

Note that for a typical many-body Hamiltonian, we expect that $\|\operatorname{Re} H_0\|$, $\|\operatorname{Im} H_0\| \sim \log D$ and therefore C_{H_0} should be very small for large D.

Theorem 4. (Lower bounds for $M_{\mu\nu}$) *Let $\varepsilon' \in (0, \frac{1}{2})$ and let μ and ν be arbitrary macro states such that $d_\mu, d_\nu > \max\{166, 4|\log_2(\varepsilon'/\sqrt{2})|\}$. Let H satisfy Assumption 1. Let $\sigma_- := \min_{i,j} \sigma_{ij}$ and $\sigma_+ := \max_{i,j} \sigma_{ij}$. Then with probability at least $1 - \varepsilon'$,*

$$M_{\mu\nu} \geq \left(\sqrt{\varepsilon' C_\sigma} \, \frac{\max\{d_\mu, d_\nu\}}{D} \right)^{16} \min\left\{1, \frac{d_\nu}{d_\mu}\right\}, \tag{40}$$

where

$$C_\sigma := \frac{c_- \sigma_-}{c_+ \sigma_+ + C_{H_0}} \tag{41}$$

with C_{H_0} defined in (39) and absolute constants $c_-, c_+ > 0$.

The main application of this theorem is provided by our next result: the analogue of Corollary 1 for the relative errors. In addition to (30), define s_{mc} by

$$D = \exp\left(s_{\mathrm{mc}} N / k_B\right). \tag{42}$$

Since we expect that the dimension of the thermal equilibrium macro space d_{eq} is extremely large, it should hold that $s_{\mathrm{mc}} \approx s_{\mathrm{eq}}$.

Corollary 2 (Generalized normal typicality—relative errors). *Let $\varepsilon, \delta > 0$, $\varepsilon' \in (0, \frac{1}{2})$ and let μ and ν be macro states such that $d_\mu, d_\nu > \max\{166, 4| \log_2(\varepsilon'/\sqrt{2})|\}$. Let H be a random Hermitian $D \times D$ matrix as in Theorem 4. Then with probability at least $1 - \varepsilon'$, $(1 - \varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that for $(1 - \delta)$-most $t \in [0, \infty)$,*

$$\frac{\left|\|P_\nu\psi_t\|^2 - M_{\mu\nu}\right|}{M_{\mu\nu}} \leq \frac{4}{\sqrt{\varepsilon\delta}}(C_\sigma\varepsilon')^{-8}\exp\left(-\frac{N}{2k_B}\left(\min\{s_\mu,s_\nu\} - 32(s_{mc} - \max\{s_\mu,s_\nu\})\right)\right), \tag{43}$$

where C_σ *is defined in* (41).

The most important aspects of this bound are that it bounds the *relative* error, it shrinks exponentially as N increases, and the rate of shrinking can be given explicitly.

Remark 1 It is sometimes desirable to consider a sequence of systems with increasing dimension $D \to \infty$, corresponding, for example, to an increasing particle number $N \to \infty$, for which it makes sense to talk of the "same" macro states ν for each member of the sequence. In that case, suppose that C_σ is bounded below away from zero uniformly in D (as C_{H_0} is small in typical models, this is basically a condition on σ_- and σ_+). Then, the relative errors in Corollary 2 are small if $s_{\mathrm{mc}} < \max\{s_\mu, s_\nu\} + \min\{s_\mu, s_\nu\}/32$. Since we expect that $s_{\mathrm{mc}} \approx s_{\mathrm{eq}}$, one of the macro spaces $\mathcal{H}_\nu$ or $\mathcal{H}_\mu$ thus needs to have specific entropy not too far from the one of the equilibrium macro state. While this might seem quite restrictive at first sight, observe that even if we assume that the $M_{\mu\nu}$ scale like in the case of normal typicality, i.e.,

$$M_{\mu\nu} \approx \frac{d_\nu}{D} \approx \exp\left(-\frac{s_{\mathrm{mc}} - s_\nu}{k_B}N\right), \tag{44}$$

then the relative errors are small only if $s_{\mathrm{mc}} < \max\{s_\nu, s_\mu\} + \min\{s_\nu, s_\mu\}/2$ (recall Theorem 2 for the absolute errors), see also the discussion in [43].

Remark 2 Suppose again that C_{H_0} is bounded uniformly in D and that $\sigma_\pm = D^{\alpha_\pm}$ with $\alpha_- \leq \alpha_+$. In this case, the upper bound for the relative errors becomes

$$\begin{aligned}\frac{\left|\|P_\nu\psi_t\|^2 - M_{\mu\nu}\right|}{M_{\mu\nu}} &\leq \frac{4}{\sqrt{\varepsilon\delta}(c_-\varepsilon')^8}\left(c_+\exp(\alpha_+ s_{\mathrm{mc}}N/k_B) + C_{H_0}\right)^8 \\ &\quad\times \exp\left(-\frac{N}{2k_B}\left(\min\{s_\mu,s_\nu\} - 32\left(\left(1-\frac{\alpha_-}{2}\right)s_{\mathrm{mc}} - \max\{s_\mu,s_\nu\}\right)\right)\right).\end{aligned} \tag{45}$$

Suppose first that $\alpha_+ < 0$. Then, for large N, the second factor can be bounded by $(2C_{H_0})^8$ and the right-hand side is small if the expression in the exponential function is negative, i.e., if

$$\left(1-\frac{\alpha_-}{2}\right)s_{\mathrm{mc}} < \max\{s_\mu,s_\nu\} + \frac{1}{32}\min\{s_\mu,s_\nu\}. \tag{46}$$

Since $\alpha_- < 0$ the macro state μ or ν has to be even closer to the equilibrium macro state than in the case that the variances σ_-, σ_+ do not depend on the dimension D. If $\alpha_+ = 0$ then the second factor in (45) can also be bounded by some constant and we obtain again (46) as a condition for the relative errors to be small. If, however, $\alpha_+ > 0$, the second factor can be bounded by a quantity proportional to $\exp(8\alpha_+ s_{\mathrm{mc}} N/k_B)$ and therefore the right-hand side of (45) is small if

$$\left(1 + \frac{\alpha_+ - \alpha_-}{2}\right) s_{\mathrm{mc}} < \max\{s_\mu, s_\nu\} + \frac{1}{32}\min\{s_\mu, s_\nu\}. \tag{47}$$

Thus, if $\alpha_+ > \alpha_-$, then the macro state μ or ν again has to be closer to the equilibrium macro state than in the case discussed in Remark 1. Note that, of course, we can also assume that $\sigma_\pm = c_{\sigma_\pm} D^{\alpha_\pm}$ with constants $c_{\sigma_\pm} > 0$ and the conditions under which the relative errors are small remain the same since the upper bounds only change by multiplicative constants.

3.2. Dynamical Typicality

We now turn to our main result about dynamical typicality. We would like to give again a bound on the *relative error*, but this we can only provide for *most* (rather than *all*) times. But for *most* times, we already know from generalized normal typicality that $\|P_\nu \psi_t\|^2$ is close to $M_{\mu\nu}$ (as visible in Fig. 2). On the other hand, for dynamical typicality, we are particularly interested in times before normal equilibration has set in (as in Fig. 1). For this situation, we can provide a bound on what we call the *comparative error* and that is given by the quotient of the absolute error (which in this case is time-dependent) and the time average of the quantity considered. In this way, we compare the absolute error to a comparison value that expresses in a way what magnitude to expect of $\|P_\nu \psi_t\|^2$. Smallness of the comparative error at time t means that the absolute error at t is much smaller than the long-time value of $\|P_\nu \psi_t\|^2$, although not necessarily much smaller than the instantaneous ensemble average

$$\mathbb{E}_\mu \|P_\nu \psi_t\|^2 = w_{\mu P_\nu}(t) =: w_{\mu\nu}(t)\,. \tag{48}$$

This statement still gives us a handle on comparatively small $\mathcal{H}_\nu$ for which $\|P_\nu \psi_t\|^2$ is small in absolute terms for most t. Now recall from (28) that

$$M_{\psi_0 \nu} = \overline{\|P_\nu \psi_t\|^2}\,. \tag{49}$$

Theorem 5 (Dynamical typicality—comparative errors). *Let $\varepsilon > 0$, $\varepsilon' \in (0, \frac{1}{2})$ and let μ and ν be macro states such that $d_\nu > \max\{166, 4|\log_2 \varepsilon'|\}$. Let H be a random Hermitian $D \times D$ matrix as in Theorem 4. Then with probability at least $1 - \varepsilon'$, for each $t \in \mathbb{R}$, $(1 - \varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that*

$$\begin{aligned}\frac{\left|\|P_\nu \psi_t\|^2 - w_{\mu\nu}(t)\right|}{M_{\psi_0 \nu}} &\\ \leq \frac{1}{\sqrt{\varepsilon}} (C_\sigma \varepsilon')^{-8} \exp&\left(-\frac{N}{2k_B}\left(2s_\mu - \min\{s_\mu, s_\nu\} - 32(s_{mc} - s_\nu)\right)\right)\end{aligned} \tag{50}$$

where C_σ is defined in (41). Moreover, with probability $1-\varepsilon'$, for every $T>0$, $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that

$$\frac{1}{T}\int_0^T \frac{\left|\|P_\nu \psi_t\|^2 - w_{\mu\nu}(t)\right|^2}{M^2_{\psi_0 \nu}}\,dt \le \frac{1}{\varepsilon}(C_\sigma \varepsilon')^{-16} \exp\left(-\frac{N}{k_B}\left(s_\mu - 32(s_{mc}-s_\nu)\right)\right). \tag{51}$$

Again, key aspects are that the upper bounds in (50) and (51) shrink exponentially with increasing N at explicit rates. The bounds are small if $s_{\mathrm{mc}} < s_\nu + s_\mu/16 - \min\{s_\mu, s_\nu\}/32$ resp. $s_{\mathrm{mc}} < s_\nu + s_\mu/32$, i.e., s_ν must be close to s_{mc} and therefore close to s_{eq}. In that case, (51) shows that the curve $t \mapsto \|P_\nu \psi_t\|^2$ is close to the curve $t \mapsto \mathbb{E}_\mu \|P_\nu \psi_t\|^2$ in the L^2-sense even when compared to the (possibly small) time average $\overline{\|P_\nu \psi_t\|^2}$.

4. No-Gaps Delocalization and Band Matrices

Our main results are based on a "no-gaps delocalization" result of Rudelson and Vershynin (2016) [37, Theorem 1.3]. In this section, we state this result (with certain corrections, Theorem 6 in Sect. 4.1), provide an extension of it that we will use when proving our main results (Theorem 7 in Sect. 4.2), and provide examples of random matrices for which the hypotheses of Rudelson and Vershynin are satisfied (Theorems 8 and 9 in Sect. 4.3).

4.1. Statement of the No-Gaps Delocalization Result of Rudelson and Vershynin

Since explicit exponents and the dependence of constants on the parameters play an important role for our purposes, we have followed the exponents and constants through the proof of Rudelson and Vershynin in [37]. It turned out that in some places we arrived at different values than stated in [37]. In this subsection, we state the relevant result with these values, and provide a discussion of how we arrived at them.

In the following, we use the following abbreviation. For a random $D \times D$ matrix H and any number $J>0$, we introduce the "boundedness event"

$$\mathcal{B}_{H,J} := \left\{\|H\| \le J\sqrt{D}\right\}. \tag{52}$$

It will turn out in Sect. 4.3 that for relevant examples of random matrices H and large D, the event $\mathcal{B}_{H,J}$ occurs with high probability; this conclusion makes use of a result of Latala [22].

So here is our adjusted statement of Theorem 1.3 of [37].

Theorem 6 (No-gaps delocalization; Rudelson, Vershynin (2016)). *Let $H=(h_{ij})$ be a $D \times D$ random matrix such that for $i,j \in [D]$ the (continuous) random variable $\mathrm{Re}\, h_{ij}$ is independent of the other entries of $\mathrm{Re}\, H$ except possibly $\mathrm{Re}\, h_{ji}$, the densities $\varrho^{\mathrm{Re}}_{ij}$ of $\mathrm{Re}\, h_{ij}$ are bounded by some number $K \ge 1$ and the imaginary part is fixed. Choose $J \ge 1$ such that the boundedness event $\mathcal{B}_{H,J}$ holds with probability at least $1/2$. Let $\kappa \in (180/D, 1/2)$ and $s>0$. Then,* *conditionally on $\mathcal{B}_{H,J}$, the following holds with probability at least $1-(Cs)^{\kappa D}$: Every eigenvector v of H satisfies*

$$\|v_I\| \geq (\kappa s)^9 \|v\| \quad \textit{for all } I \subset [D] \textit{ with } |I| \geq \kappa D, \tag{53}$$

where

$$\|v_I\| := \left(\sum_{j \in I} |v_j|^2 \right)^{1/2} \tag{54}$$

and $C = C(K,J) \geq 1$.

Here is how this statement differs from Theorem 1.3 in [37]. The lower bound on ε is changed from $8/D$ to $180/D$ and the exponent of κs in (53) from 6 to 9. Therefore, the theorem tells us that subsets $I \subset [D]$ of at least 180 (instead of 8) coordinates of any eigenvector carry a non-negligible part of its mass. Moreover, the lower bound obtained for $\|v_I\|$ is slightly worse since $\kappa < 1$ by assumption and we can assume without loss of generality that also $s < 1$ because otherwise the lower bound on the probability, $1-(Cs)^{\kappa D}$, is negative.

Here is how we arrived at these exponents and constants. In the rest of this section, all numbers of theorems, subsections, etc., refer to the ones in [37] if not stated otherwise. We start from Theorem 5.1 (Sect. 5.1) to derive Theorem 1.3. Let C_2 be the constant C in Theorem 5.1 and define for $\delta, \kappa > 0$ the *gap event* (called the *localization event* in [37])

$$\mathrm{Gap}(H,\kappa,\delta) := \left\{ \exists \text{ eigenvector } v \in \mathbb{C}^D, \|v\| = 1, \exists I \subset [D], |I| = \kappa D : \|v_I\| < \delta \right\}, \tag{55}$$

i.e., the event that in an eigenvector v of H a fraction κ of basis vectors is underrepresented in v.

After an application of Proposition 4.1, one arrives at

$$\mathbb{P}\left(\mathrm{Gap}(H,\kappa,t/8J) \text{ and } \mathcal{B}_{H,J}\right) \leq 5 \left(\frac{8J}{t}\right)^2 (e/\kappa)^{\kappa D} p_0 \tag{56}$$

where $t \geq 0$ and $p_0 = (2C_2 K J t^{0.4} \kappa^{-1.4})^{\kappa D/2}$. Set $t = 8J(\kappa s)^\alpha$ for some $\alpha > 0$. (In [37], $\alpha = 6$ was stated, while we arrive at $\alpha = 9$ below, so let us keep the value unspecified for a little while.) What we want to show is that

$$\mathbb{P}\Big(\mathrm{Gap}(H,\kappa,(\kappa s)^\alpha) \Big| \mathcal{B}_{H,J}\Big) \leq (Cs)^{\kappa D}, \tag{57}$$

where C can depend on J and K but not on κ, s, or D. Since $\mathbb{P}(\mathcal{B}_{H,J}) \geq 1/2$, it suffices to bound $\mathbb{P}(\mathrm{Gap}(H,\kappa,(\kappa s)^\alpha) \cap \mathcal{B}_{H,J})$ in a similar way. From (56), we obtain that

$$\begin{aligned} &\mathbb{P}\left(\mathrm{Gap}(H,\kappa,(\kappa s)^\alpha) \text{ and } \mathcal{B}_{H,J}\right) \\ &\quad \leq 5(\kappa s)^{-2\alpha} (e/\kappa)^{\kappa D} \left(2C_2 K J t^{0.4} \kappa^{-1.4}\right)^{\kappa D/2} &(58a) \\ &\quad = \left(5^{1/\kappa D} \kappa^{-2\alpha/\kappa D + 0.2\alpha - 1.7} s^{-2\alpha/\kappa D + 0.2\alpha - 1} e \sqrt{2C_2 K J} (8J)^{0.2} s\right)^{\kappa D} &(58b) \end{aligned}$$

$$\leq \left(\kappa^{-2\alpha/\kappa D+0.2\alpha-1.7} s^{-2\alpha/\kappa D+0.2\alpha-1} 5e\sqrt{2C_2KJ}(8J)^{0.2}s\right)^{\kappa D}. \tag{58c}$$

So we would like a constant (independent of κ, s, D) as an upper bound for

$$\kappa^{-2\alpha/\kappa D+0.2\alpha-1.7} s^{-2\alpha/\kappa D+0.2\alpha-1}, \tag{59}$$

where $\kappa < 1/2$ and $s < 1$. However, in the limit $\kappa D \to 0$, it follows that $-2\alpha/\kappa D \to -\infty$ and $\kappa^{-2\alpha/\kappa D} \to \infty$, so we need to exclude κ values from a neighborhood of 0 and require, say, $\kappa \geq \ell/D$ for some $\ell \in \mathbb{N}$ (as explicitly done in [37] and in Theorem 6). But then κ can still be quite close to 0, and its exponent in (59) should be ≥ 0 or else (59) will not be bounded as $D \to \infty$, so we want that

$$-\frac{2\alpha}{\kappa D} + 0.2\alpha - 1.7 \geq 0\,. \tag{60}$$

In particular, for $\alpha = 6$ as in [37], this exponent would always be negative, and (59) for $\kappa = \ell/D$ would not be bounded. However, if we suppose that $\ell > 10$ and $\alpha \geq \frac{17\ell}{2(\ell-10)}$, then, for every $\kappa \geq \ell/D$,

$$-\frac{2\alpha}{\kappa D} + \frac{\alpha}{5} - \frac{17}{10} \geq -\frac{2\alpha}{\ell} + \frac{\alpha}{5} - \frac{17}{10} \tag{61a}$$

$$= \alpha\frac{\ell-10}{5\ell} - \frac{17}{10} \tag{61b}$$

$$\geq 0. \tag{61c}$$

In this case, we obtain that

$$\mathbb{P}\left(\mathrm{Gap}\left(H, \kappa, (\kappa s)^{\alpha}\right) \text{ and } \mathcal{B}_{H,J}\right) \leq \left(5e\sqrt{2C_2KN}(8J)^{0.2}s\right)^{\kappa D} \tag{62}$$

$$=: \left(\tilde{C}s\right)^{\kappa D} \tag{63}$$

with $\tilde{C} = \tilde{C}(K,J) = 5e\sqrt{2C_2KJ}(8J)^{0.2}$ being the desired constant.[4] Note that $\ell \mapsto \frac{17\ell}{2(\ell-10)}$ for $\ell > 10$ is monotonically decreasing and $\lim_{\ell\to\infty} \frac{17\ell}{2(\ell-10)} = 8.5$. Therefore, we can choose $\alpha = 9$ and $\ell = 180$ to get $-2\alpha/\kappa D + 0.2\alpha - 1.7 \geq 0$.

Remark 3 (Constant in Theorem 6) Since the constant $C(K,J)$ in Theorem 6 will appear in the upper bound for the relative errors and we want to allow the case that K and J depend on D, we are interested in the dependence of C on these two parameters. In [37], it is said that all appearing constants (denoted by $C, c, \ldots$) might depend on K and J, and therefore, we carefully investigated the constants in the theorems, propositions, lemmas, etc., needed for the proof of their main theorem (here Theorem 6). To avoid confusion with the constant C in the main theorem, we denote the other constants that are also denoted as C in [37] as $C_1, C_2, \ldots$ in the order of their appearance; other constants that appear multiple times with a different value as well are also numbered in the following in the order of their appearance. An inspection of the proofs reveals that

[4]More precisely, $C = 2\tilde{C}$, where C is the constant in Theorem 6. This is due to the fact that $\mathbb{P}\left(\mathrm{Gap}(H,\kappa,(\kappa s)^{\alpha}) \text{ and } \mathcal{B}_{H,J}\right)$ and not $\mathbb{P}\left(\mathrm{Gap}(H,\kappa,(\kappa s)^{\alpha})|\mathcal{B}_{H,J}\right)$ is considered in the proof of Theorem 1.3.

- $C_1 = e$ in Lemma 3.3,
- $C = 2\tilde{C}$, where $\tilde{C}$ is defined above,
- $C_2 = 2\max\left\{\left(\frac{4}{\tilde{c}_1}\right)^{0.9} C_6, C_4\sqrt{\tilde{c}_2}\right\}$ in Theorem 5.1, where $\tilde{c}_2$ is an absolute constant that bounds the constants $C(p)$ in (5.7) for $p \geq 2$, see also Theorem 3.21 in [41], and $\tilde{c}_1 \in (0, 0.2)$ is another absolute constant that can be chosen arbitrarily, see also Section 5.4.2 and 5.5.3 (note that in (5.9) a square is missing); the values of the other constants are given below,
- $C_3 = \frac{1}{4}C_2$ in (5.1),
- $C_4 = \sqrt{2}\hat{C}\bar{C}$ in Lemma 5.4, where $\hat{C}, \bar{C} > 0$ are absolute constants; more precisely, $\hat{C}$ appears in the upper bound for the density of a random vector PX where $X = (X_1, \dots, X_D)$ is a vector of real-valued independent random variables whose densities are bounded by some constant a.e. and P is an orthogonal projection onto a subspace of dimension d, see also [36] for details; moreover, the constant $\bar{C}$ appears in the upper bound of the volume of a d-dimensional ball of radius $\tau\sqrt{d}$ for some $\tau > 0$,[5]
- $C_{0,1} = 4$ in Lemma 5.5,
- $C_{0,2} = 4e$ in Lemma 5.6,
- $C_5 = 72e$ in Lemma 5.7,
- $C_6 = 72e$ in Lemma 5.8.

Thus, we see that the constant C_2 depends neither on K nor on J, and we conclude from our computation above that $C \sim \sqrt{K}J^{0.7}$.

4.2. An Extension

In Theorem 6, the imaginary part of the random matrix is fixed, and it is assumed that $J \geq 1$. In this section, we present and prove a theorem that covers matrices with random imaginary part as well, that also allows $0 < J \leq 1$ and we improve the exponent of κs in the lower bound for $\|v_I\|$ from 9 to 8. More precisely, we prove the following theorem:

Theorem 7 *Let $H = (h_{ij})$ be a $D \times D$ random matrix such that $Re\,H$ and $Im\,H$ are independent, for $i, j \in [D]$ the (continuous) random variable $Re\,h_{ij}$ is independent of the other entries of $Re\,H$ except possibly $Re\,h_{ji}$, for $i, j \in [D], i \neq j$ the (continuous) random variable $Im\,h_{ij}$ is independent of the other entries of $Im\,H$ except possibly $Im\,h_{ji}$, and the densities ϱ_{ij}^{Re} are bounded by some number $K > 0$. Choose $J > 0$ such that $JK \geq 1$ and the boundedness events $\mathcal{B}_{Re\,H,J}$ and $\mathcal{B}_{Im\,H,J}$ hold with probability at least $1 - \eta$ for some $0 < \eta \leq 1/2$. Let $\kappa \in (83/D, 1/2)$ and $0 < s \leq 1$. Then, the following holds with*

[5] The constant $\bar{C}$ can be chosen as $\bar{C} = \sqrt{2\pi e}$ which can be easily seen as follows: For $\tau > 0$, the volume of a d-dimensional ball with radius $\tau\sqrt{d}$ is given by

$$\frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}\left(\tau\sqrt{d}\right)^d \leq \left(\frac{e}{\frac{d}{2}+1}\right)^{d/2}\left(\sqrt{\pi}\tau\sqrt{d}\right)^d \leq \left(\sqrt{2\pi e}\tau\right)^d = (\bar{C}\tau)^d \qquad (64)$$

probability at least $\left(1-(c_c\sqrt{KJ}s)^{\kappa D}\right)(1-\eta)^4$*, where* $c_c \geq 1$ *is a universal constant: Every eigenvector* v *of* H *satisfies*

$$\|v_I\| \geq (\kappa s)^8 \|v\| \quad \textit{for all } I \subset [D] \textit{ with } |I| \geq \kappa D\,. \tag{65}$$

Proof We prove the theorem in three steps. The first step is to establish a modification of Theorem 6 where the imaginary part of the matrix H can also be random. The second step is to relax the condition $J \geq 1$ to $J > 0$ (in this case we have to additionally assume that $JK \geq 1$), which can be done via a simple scaling argument. Finally, in the third step, we show that by a small modification in the proof of Rudelson and Vershynin, the lower bound for $\|v_I\|_2$ can be slightly improved.

1. Step (the complex case) We first assume that $J, K \geq 1$ and that $\kappa \in (180/D, 1/2)$. Let

$$\begin{aligned} S := \{&\text{Every eigenvector } v \text{ of } H \text{ satisfies } \|v_I\| \geq (\kappa s)^9\|v\| \text{ for all } I \subset [D] \\ &\text{with } |I| \geq \kappa D\}, \end{aligned} \tag{66}$$

$$E := \left\{Q \in \mathbb{R}^{D\times D} : \mathbb{P}\left(\mathcal{B}_{\mathrm{Re}\, H+iQ,2J}\right) \geq \frac{1}{2}\right\}. \tag{67}$$

By the law of total expectation and the monotonicity of the conditional expectation, we obtain that

$$\begin{aligned} \mathbb{P}(S|\mathcal{B}_{H,2J}) &= \mathbb{E}\left(\mathbb{P}\left(S|\mathcal{B}_{H,2J}, \mathrm{Im}\, H\right)|\mathcal{B}_{H,2J}\right) && \text{(68a)}\\ &\geq \mathbb{E}\left(\mathbb{1}_{\{\mathrm{Im}\, H\in E\}}\mathbb{P}\left(S|\mathcal{B}_{H,2J}, \mathrm{Im}\, H\right)|\mathcal{B}_{H,2J}\right) && \text{(68b)}\\ &\geq \left(1-(Cs)^{\kappa D}\right)\mathbb{E}\left(\mathbb{1}_{\{\mathrm{Im}\, H\in E\}}|\mathcal{B}_{H,2J}\right) && \text{(68c)}\\ &= \left(1-(Cs)^{\kappa D}\right)\mathbb{P}\left(\mathrm{Im}\, H \in E|\mathcal{B}_{H,2J}\right) && \text{(68d)}\\ &\geq \left(1-(Cs)^{\kappa D}\right)\mathbb{P}\left(\{\mathrm{Im}\, H \in E\}\cap\mathcal{B}_{H,2J}\right), && \text{(68e)} \end{aligned}$$

where $C = C(K, 2J) \geq 1$ is the constant in Theorem 6, and Theorem 6 itself was applied in the third line. Next note that

$$\mathcal{B}_{\mathrm{Re}\, H,J} \cap \mathcal{B}_{\mathrm{Im}\, H,J} \subset \mathcal{B}_{H,2J}\,, \tag{69}$$

and thus

$$\begin{aligned} \mathbb{P}\left(\{\mathrm{Im}\, H\in E\}\cap\mathcal{B}_{H,2J}\right) &\geq \mathbb{P}\left(\{\mathrm{Im}\, H\in E\}\cap\mathcal{B}_{\mathrm{Re}\, H,J}\cap\mathcal{B}_{\mathrm{Im}\, H,J}\right) && \text{(70a)}\\ &= \mathbb{P}\left(\mathcal{B}_{\mathrm{Re}\, H,J}\right)\mathbb{P}\left(\{\mathrm{Im}\, H\in E\}\cap\mathcal{B}_{\mathrm{Im}\, H,J}\right) && \text{(70b)}\\ &\geq (1-\eta)\,\mathbb{P}\left(\{\mathrm{Im}\, H\in E\}\cap\mathcal{B}_{\mathrm{Im}\, H,J}\right), && \text{(70c)} \end{aligned}$$

where we used that $\mathrm{Re}\, H$ and $\mathrm{Im}\, H$ are independent. Next observe that for $Q \in \mathbb{R}^{D\times D}$ such that $\|Q\| \leq J\sqrt{D}$, we have that

$$\mathbb{P}\left(\mathcal{B}_{\mathrm{Re}\, H+iQ,2J}\right) \geq \mathbb{P}\left(\mathcal{B}_{\mathrm{Re}\, H,J}\right) \geq 1-\eta \geq \frac{1}{2}. \tag{71}$$

It follows that $\mathcal{B}_{\mathrm{Im}\, H,J} \subset \{\mathrm{Im}\, H \in E\}$ and thus

$$\mathbb{P}\left(\{\mathrm{Im}\, H\in E\}\cap\mathcal{B}_{H,2J}\right) \geq (1-\eta)\,\mathbb{P}\left(\mathcal{B}_{\mathrm{Im}\, H,J}\right) \geq (1-\eta)^2. \tag{72}$$

Therefore, we finally obtain that

$$\mathbb{P}\left(S|\mathcal{B}_{H,2J}\right) \geq \left(1-(Cs)^{\kappa D}\right)(1-\eta)^2\,, \tag{73}$$

which implies

$$\begin{aligned}\mathbb{P}(S) &\geq \mathbb{P}\left(S\cap\mathcal{B}_{H,2J}\right) = \mathbb{P}\left(S|\mathcal{B}_{H,2J}\right)\mathbb{P}\left(\mathcal{B}_{H,2J}\right) &\text{(74a)}\\ &\geq \mathbb{P}\left(S|\mathcal{B}_{H,2J}\right)\mathbb{P}\left(\mathcal{B}_{\mathrm{Re}\,H,J}\right)\mathbb{P}\left(\mathcal{B}_{\mathrm{Im}\,H,J}\right) &\text{(74b)}\\ &\geq \left(1-(Cs)^{\kappa D}\right)(1-\eta)^4. &\text{(74c)}\end{aligned}$$

This shows that for complex random $D\times D$ matrices H which satisfy the assumptions of this theorem with $J,K\geq 1$ (instead of $K,J>0$ such that $JK\geq 1$) and $\kappa\in(180/D,1/2)$ (instead of $\kappa\in(83/D,1/2)$), the following holds with probability at least $\left(1-(Cs)^{\kappa D}\right)(1-\eta)^4$: Every eigenvector v of H satisfies

$$\|v_I\| \geq (\kappa s)^9\|v\| \quad \text{for all } I\subset[D] \text{ with } |I|\geq\kappa D, \tag{75}$$

where $C=C(K,2J)\geq 1$ and $C(K,2J)\sim\sqrt{K}J^{0.7}$.

2. Step ($J>0$) As in the first step of the proof we assume that $\kappa\in(180/D,1/2)$. Note that if the density of a random variable X is bounded by some constant K, then the density of $J^{-1}X$ is bounded by JK. Therefore, applying the result of the first step to the matrix $\tilde{H}:=J^{-1}H$ immediately shows that for complex random $D\times D$ matrices H which satisfy the assumptions of this theorem with $\kappa\in(180/D,1/2)$ (instead of $\kappa\in(83/D,1/2)$), the following holds with probability at least $\left(1-(\tilde{c}s)^{\kappa D}\right)(1-\eta)^4$: Every eigenvector v of H satisfies

$$\|v_I\| \geq (\kappa s)^9\|v\| \quad \text{for all } I\subset[D] \text{ with } |I|\geq\kappa D, \tag{76}$$

where $\tilde{c}=C(KJ,2)\sim\sqrt{KJ}$.

3. Step (improved bound) We first assume that the imaginary part of H is fixed and that $K,J\geq 1$. In the proof of the Invertibility Theorem 5.1 at the end of Section 5 in [37], they arrive at the upper bound

$$\left[C_6KJ^{0.1}\kappa^{-0.05}\left(\frac{4Jt}{\tilde{c}_1\tau\kappa^{3/2}}\right)^{0.9}\right]^{\kappa D} + \left(C_4\sqrt{\tilde{c}_2}K\tau\right)^{\kappa D} \tag{77}$$

for the probability on the left-hand side of (5.18), where $t,\tau>0$ are arbitrary and where we used the notation for the constants introduced above. In their paper, they choose $\tau=\sqrt{t}$ and then they use that they can assume that $t\leq 1$ to bound $t^{0.45}$ and $t^{0.5}$ by $t^{0.4}$. The result can be slightly improved by choosing $\tau=t^{9/19}$ instead. With this choice, the upper bound (77) becomes

$$\left[\left(\frac{4}{\tilde{c}_1}\right)^{0.9}C_6KJ\kappa^{-1.4}t^{9/19}\right]^{\kappa D} + \left(C_4\sqrt{\tilde{c}_2}Kt^{9/19}\right)^{\kappa D} \leq \left(C_2KJ\kappa^{-1.4}t^{9/19}\right)^{\kappa D}, \tag{78}$$

where $C_2 = 2\max\left\{\left(\frac{4}{\tilde{c}_1}\right)^{0.9} C_6, C_4\sqrt{\tilde{c}_2}\right\}$. This shows that the exponent of t in Theorem 5.1 in [37] can be changed from 0.4 to 9/19. This change has implications for the derivation of the no-gaps delocalization theorem from Theorem 5.1: After an application of Proposition 4.1, one now arrives at

$$\mathbb{P}\left(\mathrm{Gap}(H,\kappa,t/8J) \text{ and } \mathcal{B}_{H,J}\right) \leq 5\left(\frac{8J}{t}\right)^2 (e/\kappa)^{\kappa D} p_0 \tag{79}$$

with $t \geq 0$ and $p_0 = (2C_2KJt^{9/19}\kappa^{-1.4})^{\kappa D/2}$. Again we set $t = 8J(\kappa s)^\alpha$ for some $\alpha > 0$. Then, a computation similar to the one in the proof of Theorem 6 shows that

$$\begin{aligned}&\mathbb{P}\left(\mathrm{Gap}(H,\kappa,(\kappa s)^\alpha) \text{ and } \mathcal{B}_{H,J}\right)\\ &\quad\leq \left(\kappa^{-2\alpha/\kappa D+9\alpha/38-1.7} s^{-2\alpha/\kappa D+9\alpha/38-1} 5e\sqrt{2C_2KJ}(8J)^{9/38}s\right)^{\kappa D}.\end{aligned} \tag{80}$$

In order to obtain an upper bound for $\kappa^{-2\alpha/\kappa D+9\alpha/38-1.7}s^{-2\alpha/\kappa D+9\alpha/38-1}$, we have to determine α such that $-2\alpha/\kappa D + 9\alpha/38 - 1.7 \geq 0$ since $\kappa < 1/2$ and $s \leq 1$. Suppose again that $\kappa \geq \frac{\ell}{D}$ for some $\ell \in \mathbb{N}, \ell > 8$, and $\alpha \geq \frac{323\ell}{5(9\ell-76)}$. Then,

$$-\frac{2\alpha}{\kappa D} + \frac{9\alpha}{38} - \frac{17}{10} \geq -\frac{2\alpha}{\ell} + \frac{9\alpha}{38} - \frac{17}{10} \tag{81a}$$

$$= \alpha\frac{9\ell-76}{38\ell} - \frac{17}{10} \tag{81b}$$

$$\geq 0. \tag{81c}$$

Thus, we arrive at

$$\mathbb{P}\left(\mathrm{Gap}(H,\kappa,(\kappa s)^\alpha) \text{ and } \mathcal{B}_{H,J}\right) \leq \left(5e\sqrt{2C_2KJ}(8J)^{9/38}s\right)^{\kappa D} \tag{82a}$$

$$= \left(\frac{cs}{2}\right)^{\kappa D} \tag{82b}$$

with $c = 10e\sqrt{2C_2KJ}(8J)^{9/38}$, i.e., $c \sim \sqrt{K}J^{14/19}$. Note that $\ell \mapsto \frac{323\ell}{5(9\ell-76)}$ for $\ell > 8$ is monotonically decreasing and $\lim_{\ell\to\infty}\frac{323\ell}{5(9\ell-76)} = \frac{323}{45}$. Therefore, we can choose $\alpha = 8$ and $\ell = 83$ to get $-2\alpha/\kappa D + 9\alpha/38 - 1.7 \geq 0$. For matrices whose imaginary part is not fixed, we proceed as in step 1 to obtain the result for complex matrices. In order to relax the condition $J \geq 1$ to $J > 0$ (and also $K \geq 1$ to $K > 0$), one can again apply a scaling result as in step 2, where one replaces K by JK and J by 1, to obtain $c \sim \sqrt{KJ}\,1^{14/19} = \sqrt{KJ}$. We thus arrive at the final result with $c_c = 10e\sqrt{2C_2}8^{9/38}$. Note that it depends on the value of C_2 whether $c_c \geq 1$ or merely $c_c > 0$; however, we can assume without loss of generality that $c_c \geq 1$ since replacing c_c by $\max\{10e\sqrt{2C_2}8^{9/38}, 1\}$ only results in a possibly smaller lower bound for the probability with which no-gaps delocalization at least holds. $\square$

Remark 4

1. Theorem 7 still holds if some entries of $\operatorname{Im} H$ are fixed. This observation will be important when we construct examples in Sect. 4.3 since for Hermitian matrices the diagonal entries of $\operatorname{Im} H$ are fixed to zero.
2. In the proof of Theorem 7, the choice $\alpha = 8$ is the smallest one possible if one wants $\alpha \in \mathbb{N}$. Otherwise, any other value $\alpha > \frac{323}{45} \approx 7.18$ is possible, however, at the cost of a larger value for ℓ and therefore κ.

Remark 5 (A different proof strategy) Another possibility (different from the scaling argument) to obtain no-gaps delocalization in the case that $0 < J < 1$ is to find out where the assumption $J \geq 1$ is used in the proof of Rudelson and Vershynin and to suitably modify the assumptions: The assumption $J \geq 1$ is used the first time (except for Sect. 5.1 to which we will turn at the end) in the proof of Lemma 5.7 (and Lemma 5.8 which follows from an application of Lemma 5.7). If $J \geq 1$, then $J/\sqrt{\kappa} > 1$ and since the constant C on the left-hand side of (5.15) and K are also greater than 1, the upper bound on the probability is only nontrivial if $\theta^{1-2\mu} < 1$. In view of how Lemma 5.7 is applied, we can safely assume that $\mu \leq 1/2$ and thus that $\theta < 1$. In the proof of Lemma 5.7 as well as Lemma 5.8, it is used that $\theta\sqrt{\kappa}/J < 1$ which is obviously fulfilled for this choice of parameter. If we allow that $J < 1$ but require $\kappa < J^2$, then we can draw the same conclusions and argue in the same way to obtain Lemma 5.7 and Lemma 5.8. This means that the upper bound for κ has to be replaced by $\min\{1/2, J^2\}$.

The next (and last) time the assumption $J \geq 1$ is used in the proof of the Invertibility Theorem 5.1 at the end of Sect. 5. Here, the improved estimate that was presented in the third step in the proof of Theorem 7 is necessary.

Note that the assumption $JK \geq 1$ is not needed in this case and we have to assume again that $K \geq 1$. However, in our examples below, we always need that $JK \geq 1$. Moreover, note that we cannot assume anymore that the constant $c \sim \sqrt{K} J^{14/19}$ in the proof of Theorem 7 is larger than one. An adaption of the proof of Theorem 10 shows that with probability at least

$$\left(1 - \min\left\{c_c \sqrt{K} J^{14/19}, \frac{1}{2}\right\}^{\min\{d_\nu, d_\mu, 2DJ^2\}/2}\right)(1-\eta)^4 \tag{83}$$

one finds the lower bound

$$M_{\mu\nu} \geq \left(\min\left\{1, \frac{1}{2c_c\sqrt{K}J^{14/19}}\right\} \min\left\{\frac{d_\mu}{2D}, \frac{d_\nu}{2D}, J^2\right\}\right)^{16} \max\left\{1, \frac{d_\nu}{d_\mu}\right\}. \tag{84}$$

Similarly, with probability at least $\left(1 - \min\{c_c\sqrt{K}J^{14/19}, 1/2\}^{\min\{d_\mu/2, DJ^2\}}\right)$ $(1-\eta)^4$ one has the lower bound

$$|M_{\mu B}| \geq \max\left\{b_{\min}^{\pm}, \left(\min\left\{\frac{d_\mu}{2D}, J^2\right\} \min\left\{1, \frac{1}{2c_c\sqrt{K}J^{14/19}}\right\}\right)^{16} \frac{\operatorname{tr}(B^{\pm})}{d_\mu}\right\}$$
$$- \min\left\{b_{\max}^{\mp}, \frac{\operatorname{tr}(B^{\mp})}{d_\mu}\right\}. \tag{85}$$

This lower bound is sometimes slightly better and sometimes slightly worse than the one we presented in Theorem 4 but yields more complicated expressions, e.g., in Corollary 2 and we therefore stick to the simpler lower bound.

4.3. Examples

In this section, we give some examples covered by Theorem 7. We will see that among these matrices are matrices with a band structure in a basis that diagonalizes the projections onto the macro spaces to which a "small" Gaussian perturbation is added. As discussed in the introduction, we are interested in such kind of matrices since, in contrast to Wigner matrices or matrices from the GOE/GUE, they can describe systems in which a nontrivial equilibration process passing through multiple macro states occurs.

Theorem 8 (Gaussian matrices, variances bounded by constants). *Let $A = (a_{ij})$ be a $D \times D$ random matrix with independent complex Gaussian entries with mean zero, i.e., all random variables $\mathrm{Re}\, a_{ij}$, $\mathrm{Im}\, a_{ij}$, $i, j \in [D]$, are independent and $\mathrm{Re}\, a_{ij}, \mathrm{Im}\, a_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2/2)$ for some $\sigma_{ij} > 0$. Let $\sigma_{ij} = \sigma_{ji}$, $\sigma_- := \min_{i,j} \sigma_{ij}$ and $\sigma_+ := \max_{i,j} \sigma_{ij}$. Then, the Hermitian matrix $H := \frac{1}{\sqrt{2}}(A + A^*)$ is Gaussian, more precisely, its entries h_{ij} satisfy $\mathrm{Re}\, h_{ij}, \mathrm{Im}\, h_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2/2)$ for $i \neq j$ and $h_{ii} = \mathrm{Re}\, h_{ii} \sim \mathcal{N}(0, \sigma_{ii}^2)$, and fulfills the assumptions in Theorem 7 with parameters $K = \frac{1}{\sqrt{2\pi}\sigma_-}$ and $J = \frac{4}{\eta}\hat{C}\sigma_+$ for arbitrary $\eta \in (0, \frac{1}{2})$, where $\hat{C}$ is a certain universal constant with the property (88).*

Proof The matrix H is obviously Hermitian, and since $\mathrm{Re}\, a_{ij}$ and $\mathrm{Re}\, a_{ji}$ are independent and normally distributed with zero expectation and variance $\sigma_{ij}^2 = \sigma_{ji}^2$ for $i \neq j$, we have

$$\mathrm{Re}\, h_{ij} = \frac{1}{\sqrt{2}} (\mathrm{Re}\, a_{ij} + \mathrm{Re}\, \overline{a_{ji}}) = \frac{1}{\sqrt{2}} (\mathrm{Re}\, a_{ij} + \mathrm{Re}\, a_{ji}) \sim \mathcal{N}(0, \sigma_{ij}^2/2). \tag{86}$$

Similarly, one has $\mathrm{Im}\, h_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2/2)$ and

$$\mathrm{Re}\, h_{ii} = \sqrt{2}\mathrm{Re}\, a_{ii} \sim \mathcal{N}(0, \sigma_{ii}^2). \tag{87}$$

By construction, $\mathrm{Re}\, h_{ij}$ is independent of the rest of the entries of $\mathrm{Re}\, H$ except $\mathrm{Re}\, h_{ji}$, the same holds true for the imaginary parts and obviously $\mathrm{Re}\, H$ and $\mathrm{Im}\, H$ are independent.

Latala [22, Theorem 2] showed that

$$\mathbb{E}\|(x_{ij})\| \leq \hat{C} \left(\max_i \sqrt{\sum_j \mathbb{E}x_{ij}^2} + \max_j \sqrt{\sum_i \mathbb{E}x_{ij}^2} + \sqrt[4]{\sum_{i,j} \mathbb{E}x_{ij}^4} \right) \tag{88}$$

for any finite matrix of independent mean zero random variables x_{ij}, where $\hat{C} > 0$ is a universal constant. Without loss of generality, we can assume that $\hat{C} \geq 1$.

With the help of (88), the expectation of the norm of $\mathrm{Re}\, H$ can be bounded in the following way:

$$\mathbb{E}\|\mathrm{Re}\, H\| \leq \frac{1}{\sqrt{2}} \mathbb{E} (\|\mathrm{Re}\, A\| + \|\mathrm{Re}\, A^*\|) \tag{89a}$$

$$\leq \hat{C}\left(\max_i \sqrt{\sum_j \sigma_{ij}^2} + \max_j \sqrt{\sum_i \sigma_{ij}^2} + \sqrt[4]{\sum_{i,j} 3\sigma_{ij}^4}\right) \tag{89b}$$

$$\leq \hat{C}\left(2\sqrt{D\sigma_+^2} + \sqrt[4]{3D^2\sigma_+^4}\right) \tag{89c}$$

$$\leq 4\hat{C}\sigma_+\sqrt{D}. \tag{89d}$$

For $0 < \eta \leq 1/2$ as in Theorem 7, we set $J := \frac{4}{\eta}\hat{C}\sigma_+$ and find with the help of Markov's inequality that

$$\mathbb{P}\left(\mathcal{B}_{\mathrm{Re}\,H,J}\right) = 1 - \mathbb{P}\left(\|\mathrm{Re}\,H\| > J\sqrt{D}\right) \geq 1 - \frac{\mathbb{E}\|\mathrm{Re}\,H\|}{J\sqrt{D}} \geq 1 - \eta, \tag{90}$$

i.e., the boundedness event $\mathcal{B}_{\mathrm{Re}\,H,J}$ holds with probability at least $1-\eta$. In the same way, we find that also $\mathcal{B}_{\mathrm{Im}\,H,J}$ holds with probability at least $1-\eta$.

Clearly, the densities of $\mathrm{Re}\,H$ are bounded by $K := \frac{1}{\sqrt{2\pi}\sigma_-}$. This is due to the fact that the density function f of the normal distribution with mean μ and variance σ^2 attains its maximum at $x = \mu$ with $f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$. and $\frac{1}{\sqrt{2\pi}\sigma_{ij}} \leq \frac{1}{\sqrt{2\pi}\sigma_-}$ for all $i, j \in [D]$. Moreover, the condition $JK \geq 1$ is obviously fulfilled for our choice of the parameters. □

Remark 6 By choosing the variance large close to the diagonal of the matrix and small far away from it, the matrix H in Theorem 8 has some kind of band structure. For example, we can fix a monotonically decreasing function $g : [0,1] \to [\sigma_-, \sigma_+]$, a "variance profile," and define

$$\sigma_{ij} := g\left(\frac{|i-j|}{D}\right) \qquad \forall i, j \in [D]. \tag{91}$$

So far we only considered matrices whose (Gaussian) entries have mean zero. In the following, we want to relax this condition and also allow entries with nonzero mean.

Theorem 9 (Gaussian matrices with nonzero mean). *Let $H_0 = (h_{ij}^0)$ be a (deterministic) Hermitian matrix with C_{H_0} as in (39) and let V be the $D \times D$ random matrix defined in Theorem 8 (there H). Define $H := H_0 + V$. Then, H is Hermitian with (non-centered) Gaussian independent (up to conjugate symmetry) entries; more precisely, its entries h_{ij} satisfy $\mathrm{Re}\,h_{ij} \sim \mathcal{N}(\mathrm{Re}\,h_{ij}^0, \sigma_{ij}^2/2)$, $\mathrm{Im}\,h_{ij} \sim \mathcal{N}(\mathrm{Im}\,h_{ij}^0, \sigma_{ij}^2/2)$ for $i \neq j$ and $h_{ii} = \mathrm{Re}\,h_{ii} \sim \mathcal{N}(h_{ii}^0, \sigma_{ii}^2)$. Moreover, H satisfies the assumptions in Theorem 7 with $K = \frac{1}{\sqrt{2\pi}\sigma_-}$ and $J = \frac{1}{\eta}(4\hat{C}\sigma_+ + C_{H_0})$ with $\hat{C}$ as before.*

Proof The claims concerning the distribution of the entries of H are obvious. With the help of the computation in the proof of Theorem 8, we find that

$$\mathbb{E}\|\mathrm{Re}\,H\| \leq \mathbb{E}\|\mathrm{Re}\,H_0\| + \|\mathrm{Re}\,V\| \tag{92a}$$

$$\leq \left(4\hat{C}\sigma_+ + C_{H_0}\right)\sqrt{D} \tag{92b}$$

For $0 < \eta \leq 1/2$ as in Theorem 7, we set $J := \frac{1}{\eta}\left(4\hat{C}\sigma_+ + C_{H_0}\right)$ and obtain

$$\mathbb{P}\left(\mathcal{B}_{\mathrm{Re}\,H,J}\right) \geq 1 - \eta, \tag{93}$$

i.e., the boundedness event $\mathcal{B}_{\mathrm{Re}\,H,J}$ holds with probability at least $1-\eta$. The parameter K has already been computed in Theorem 8 and with this choice the condition $KJ \geq 1$ is again automatically fulfilled. This finishes the proof. $\square$

Theorem 9 covers, for example, the case that H_0 is a matrix with a band structure in a basis that diagonalizes the projections P_μ and V is a Gaussian matrix with mean zero and small variances, i.e., a small Gaussian perturbation with entries small in the L^2-sense.

5. More General Results and Proofs

In this section, we state and prove a more general version of Theorem 4 (Theorem 10 in Sect. 5.1), some corollaries thereof and prove Theorem 4. We then state a more general version of Theorem 5 (Theorem 11 in Sect. 5.2) and prove Theorem 5.

5.1. Generalized Normal Typicality

Let $H = (h_{ij}) \in M_{D\times D}(\mathbb{C})$ be a random Hermitian matrix. Instead of Assumption 1, we now make the following weaker assumption (which follows if Assumption 1 holds):

Assumption 2 The random variables $(\mathrm{Re}\,h_{ij})_{i\leq j}$ and $(\mathrm{Im}\,h_{ij})_{i<j}$ are mutually independent and continuously distributed. The densities $\varrho_{ij}^{\mathrm{Re}}$ of $\mathrm{Re}\,h_{ij}$ are bounded.

Let B^+ and B^- denote the positive and negative part of B such that $B = B^+ - B^-$. Recall that d_μ and d_ν denote the dimensions of the macro spaces $\mathcal{H}_\mu$ and $\mathcal{H}_\nu$.

Theorem 10 *Let $\varepsilon' \in (0, \frac{1}{2})$ and let μ be an arbitrary macro state. Let $B \in M_{D\times D}(\mathbb{C})$ be a Hermitian $D\times D$ matrix and let $H = (h_{ij})$ be a random Hermitian $D\times D$ matrix such that Assumption 2 is satisfied. Let K be the least upper bound for the densities ϱ_{ij}^{Re}, i.e.*

$$K := \sup \bigcup_{i\leq j} \left\{ \varrho_{ij}^{Re}(x) : x \in \mathbb{R} \right\} < \infty. \tag{94}$$

Let $d_\mu > \max\{166, 2|\log_2 \varepsilon'|\}$. Moreover, let $\eta \in (0, \frac{1}{2})$ be the unique number that solves

$$1 - \varepsilon' = \left(1 - 2^{-d_\mu/2}\right)(1-\eta)^4 \tag{95}$$

and let $J^{Re}, J^{Im} \in (0,\infty)$ be the unique numbers such that

$$\mathbb{P}\left(\|Re\, H\| \leq J^{Re}\sqrt{D}\right) = 1-\eta, \tag{96}$$

$$\mathbb{P}\left(\|Im\, H\| \leq J^{Im}\sqrt{D}\right) = 1-\eta. \tag{97}$$

Set $J := \max\left\{K^{-1}, J^{Re}, J^{Im}\right\}$*. Then with probability at least* $1-\varepsilon'$*,*

$$|M_{\mu B}| \geq \max\left\{b^+_{\min}, \left(\frac{d_\mu}{4c_c\sqrt{KJD}}\right)^{16} \frac{\mathrm{tr}(B^+)}{d_\mu}\right\} - \min\left\{b^-_{\max}, \frac{\mathrm{tr}(B^-)}{d_\mu}\right\}, \tag{98}$$

where $M_{\mu B}$ *was defined in* (22)*,* $b^+_{\min}$ *and* $b^-_{\max}$ *denote the smallest and largest eigenvalue of* B^+ *and* B^-*, respectively, and* $c_c \geq 1$ *is the constant in Theorem* 7*. In particular, if* $B = P_\nu$ *for some macro state* ν*, then*

$$M_{\mu\nu} \geq \frac{d_\nu}{d_\mu}\left(\frac{d_\mu}{4c_c\sqrt{KJD}}\right)^{16}. \tag{99}$$

Moreover, if $d_\nu \geq 166$*, then for any* $\eta \in (0, \frac{1}{2})$ *(and* J *chosen as above) it holds with probability at least* $\left(1-2^{-d_\nu/2}\right)(1-\eta)^4$ *that*

$$M_{\mu\nu} \geq \left(\frac{d_\nu}{4c_c\sqrt{KJD}}\right)^{16}. \tag{100}$$

The lower bound for $|M_{\mu B}|$ is obviously nontrivial for positive (and negative) operators B but also for operators with positive and negative eigenvalues, provided that the spectrum satisfies certain assumptions, e.g., if $b^+_{\min} > b^-_{\max}$.

Note the second lower bound (100) is sharper than (99) if $d_\nu \geq d_\mu$. By combining the lower bounds in Theorem 10 with the upper bound from Theorem 1, keeping in mind that, with probability 1, $D_E = 1$ and $d_E = D$, we immediately obtain the following corollary:

Corollary 3 *Let* $\varepsilon, \delta, \kappa, T > 0, \varepsilon' \in (0, \frac{1}{2})$ *and let* μ *be an arbitrary macro state. Let* $B \in M_{D\times D}(\mathbb{C})$ *be a Hermitian* $D \times D$ *matrix and let* H *be a random Hermitian* $D \times D$ *matrix such that Assumption* 2 *is satisfied. Let* $d_\mu > \max\{166, 2|\log_2 \varepsilon'|\}$ *and let* $K, J > 0$ *and* $\eta \in (0, \frac{1}{2})$ *be defined as in Theorem* 10*. Then with probability at least* $1-\varepsilon'$*,* $(1-\varepsilon)$*-most* $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ *are such that for* $(1-\delta)$*-most* $t \in [0, T]$

$$\frac{\left|\langle\psi_t|B|\psi_t\rangle - M_{\mu B}\right|}{|M_{\mu B}|} \leq \frac{4\left(\frac{G(\kappa)\|B\|}{\delta\varepsilon d_\mu}\left(1+\frac{8\log_2 D}{\kappa T}\right)\min\left\{\|B\|, \frac{\mathrm{tr}(|B|)}{d_\mu}\right\}\right)^{1/2}}{\max\left\{b^+_{\min}, \left(\frac{d_\mu}{4c_c\sqrt{KJD}}\right)^{16}\frac{\mathrm{tr}(B^+)}{d_\mu}\right\} - \min\left\{b^-_{\max}, \frac{\mathrm{tr}(B^-)}{d_\mu}\right\}}, \tag{101}$$

whenever the denominator of the right-hand side is positive; here $c_c \geq 1$ *is the constant in Theorem* 7*.*

The next corollary for the special case that $B = P_\nu$ for some macro state ν follows from combining the lower bounds (99) and (100) for $M_{\mu\nu}$, which yield with probability at least $(1-2^{-d_\nu/2}-2^{-d_\mu/2})(1-\eta)^4$

$$M_{\mu\nu} \geq \frac{d_\nu}{\max\{d_\nu, d_\mu\}}\left(\frac{\max\{d_\nu, d_\mu\}}{4c_c\sqrt{KJD}}\right)^{16} = \left(\frac{\max\{d_\nu, d_\mu\}}{4c_c\sqrt{KJD}}\right)^{16}\min\left\{1, \frac{d_\nu}{d_\mu}\right\}, \tag{102}$$

with the upper bound for the absolute errors from Theorem 2, keeping again in mind that for random matrices with continuously distributed entries, it holds with probability 1 that $D_E = D_G = 1$.

Corollary 4 (Generalized normal typicality: relative errors). *Let $\varepsilon, \delta > 0$, $\varepsilon' \in (0, \frac{1}{2})$ and let μ, ν be two macro states such that $d_\mu, d_\nu > \max\{166, 2|\log_2(\varepsilon'/\sqrt{2})|\}$. Let H be a random Hermitian $D \times D$ matrix such that Assumption 2 is satisfied and let $K > 0$ be defined as in Theorem 10. Moreover, let $\eta \in (0, \frac{1}{2})$ be the unique number that solves*

$$1 - \varepsilon' = \left(1 - 2^{-d_\mu/2} - 2^{-d_\nu/2}\right)(1-\eta)^4 \tag{103}$$

and let $J > 0$ be defined as in Theorem 10. Then with probability at least $1-\varepsilon'$, $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that for $(1-\delta)$-most $t \in [0,\infty)$

$$\frac{\left|\|P_\nu \psi_t\|^2 - M_{\mu\nu}\right|}{M_{\mu\nu}} \leq \frac{4}{\sqrt{\varepsilon\delta \min\{d_\mu, d_\nu\}}} \left(\frac{4c_c\sqrt{KJ}D}{\max\{d_\mu, d_\nu\}}\right)^{16}, \tag{104}$$

where $c_c \geq 1$ is the constant in Theorem 7.

We now give the proofs of Theorem 10 and of one of our main theorems, Theorem 4.

Proof of Theorem 10 We can write the Hamiltonian as $H = \sum_n E_n |\phi_n\rangle\langle\phi_n|$, where (ϕ_n) is an orthonormal basis of eigenvectors of H with eigenvalues $E_n \in \mathbb{R}$. Moreover, we obtain with the reverse triangle inequality that

$$|M_{\mu B}| \geq \frac{1}{d_\mu}\left(\sum_n \langle\phi_n|P_\mu|\phi_n\rangle\langle\phi_n|B^+|\phi_n\rangle - \sum_m \langle\phi_m|P_\mu|\phi_m\rangle\langle\phi_m|B^-|\phi_m\rangle\right). \tag{105}$$

Note that we used here the fact that, with probability 1, the eigenvalues of H are distinct. Set $\kappa := \frac{d_\mu}{2D}$ and $s := \frac{1}{2c_c\sqrt{KJ}}$, where $c_c \geq 1$ is the constant in Theorem 7. With these choices all assumptions in Theorem 7 are fulfilled[6] and with

$$I_\mu = \{d_1 + \cdots + d_{\mu-1} + 1, \ldots, d_1 + \cdots + d_\mu\} \tag{106}$$

it follows that with probability at least

$$\left(1 - \left(c_c\sqrt{KJ}s\right)^{\kappa D}\right)(1-\eta)^4 = \left(1 - 2^{-d_\mu/2}\right)(1-\eta)^4 = 1 - \varepsilon' \tag{107}$$

it holds that

$$\langle\phi_n|P_\mu|\phi_n\rangle = \sum_{j \in I_\mu} |\phi_n(j)|^2 = \|(\phi_n)_{I_\mu}\|^2 \geq (\kappa s)^{16}. \tag{108}$$

Remember that we assume that the Hamiltonian is written in a basis that diagonalizes the projections onto the macro spaces as in Fig. 3. We find the following lower bounds for the first and upper bounds for the second sum in (105):

$$\sum_n \langle\phi_n|P_\mu|\phi_n\rangle\langle\phi_n|B^+|\phi_n\rangle \geq b^+_{\min} d_\mu, \tag{109a}$$

[6] Note that the factor 1/2 in the definition of κ ensures that $\kappa < 1/2$ since d_μ/D can be arbitrarily close (but not equal to) 1. This is because the trivial case in which there is only one macro space is excluded here since in this case we obviously have normal typicality and there is nothing to show at all.

$$\sum_n \langle\phi_n|P_\mu|\phi_n\rangle\langle\phi_n|B^+|\phi_n\rangle \geq (\kappa s)^{16}\,\mathrm{tr}(B^+), \tag{109b}$$

$$\sum_m \langle\phi_m|P_\mu|\phi_m\rangle\langle\phi_m|B^-|\phi_m\rangle \leq b^-_{\max} d_\mu, \tag{109c}$$

$$\sum_m \langle\phi_m|P_\mu|\phi_m\rangle\langle\phi_m|B^-|\phi_m\rangle \leq \mathrm{tr}(B^-). \tag{109d}$$

A combination of these bounds yields

$$|M_{\mu B}| \geq \max\left\{b^+_{\min}, (\kappa s)^{16}\frac{\mathrm{tr}(B^+)}{d_\mu}\right\} - \min\left\{b^-_{\max}, \frac{\mathrm{tr}(B^-)}{d_\mu}\right\} \tag{110}$$

$$= \max\left\{b^+_{\min}, \left(\frac{d_\mu}{4c_c\sqrt{KJD}}\right)^{16}\frac{\mathrm{tr}(B^+)}{d_\mu}\right\} - \min\left\{b^-_{\max}, \frac{\mathrm{tr}(B^-)}{d_\mu}\right\}. \tag{111}$$

In particular, if $B = P_\nu$ for some macro state ν, then $B^- = 0$, $b^+_{\min} = 0$, $\mathrm{tr}(B^+) = d_\nu$ and thus

$$M_{\mu\nu} \geq \frac{d_\nu}{d_\mu}\left(\frac{d_\mu}{4c_c\sqrt{KJD}}\right)^{16}. \tag{112}$$

Note that here no absolute value on the left-hand side is needed since it immediately follows from the definition of the $M_{\mu\nu}$ that $M_{\mu\nu} \geq 0$. If $d_\nu \geq 166$, we set $\kappa := \frac{d_\nu}{2D}$ and obtain with Theorem 7 with probability at least $\left(1 - 2^{-d_\nu/2}\right)(1-\eta)^4$

$$\langle\phi_n|P_\nu|\phi_n\rangle \geq (\kappa s)^{16} = \left(\frac{d_\nu}{4c_c\sqrt{KJD}}\right)^{16} \tag{113}$$

and therefore

$$M_{\mu\nu} = \frac{1}{d_\mu}\sum_n \langle\phi_n|P_\nu|\phi_n\rangle\langle\phi_n|P_\mu|\phi_n\rangle \geq \left(\frac{d_\nu}{4c_c\sqrt{KJD}}\right)^{16}. \tag{114}$$

Proof of Theorem 4 First note that since H satisfies Assumption 1 it obviously also satisfies Assumption 2 and that $d_\nu, d_\mu > \max\{166, 2|\log_2(\varepsilon'/\sqrt{2})|\}$ by assumption. Let $\eta \in (0, \frac{1}{2})$ be the unique number that solves

$$1 - \varepsilon' = \left(1 - 2^{-2_\mu/2} - 2^{-d_\nu/2}\right)(1-\eta)^4. \tag{115}$$

Then, it follows from Theorem 7 as in the proof of Theorem 10 that with probability at least $1 - \varepsilon'$,

$$M_{\mu\nu} \geq \left(\frac{\max\{d_\nu, d_\mu\}}{4c_c\sqrt{KJD}}\right)^{16}\min\left\{1, \frac{d_\nu}{d_\mu}\right\}, \tag{116}$$

see also (102). In order to arrive at the form (40), note first that with C_{H_0} as defined in (39) and with the help of Theorem 9, the factor KJ in (116) is given by

$$KJ = \frac{1}{\sqrt{2\pi}\sigma_-\eta}\left(4\hat{C}\sigma_+ + C_{H_0}\right). \tag{117}$$

Next we derive a lower bound for η in terms of ε' in order to eliminate η from (117). Therefore observe that $d_\nu, d_\mu > 4|\log_2(\varepsilon'/\sqrt{2})| = -4\log_2(\varepsilon'/\sqrt{2})$ and with (115) it follows that

$$1 - \varepsilon' \geq \left(1 - \varepsilon'^2\right)(1-\eta)^4. \tag{118}$$

Solving for η yields

$$\eta \geq 1 - \frac{1}{(1+\varepsilon')^{1/4}} \geq \frac{\varepsilon'}{6}. \tag{119}$$

In the last step, it was used that for $f, g : [0, \frac{1}{2}] \to \mathbb{R}$, $f(x) = 1 - \frac{1}{(1+x)^{1/4}}$, $g(x) = \frac{x}{6}$ one has $f \geq g$ which can easily be shown by standard arguments. Thus with (117), we find

$$c_c^2 K J \leq \frac{6c_c^2}{\sqrt{2\pi}\sigma_-\varepsilon'}\left(4\hat{C}\sigma_+ + C_{H_0}\right) =: \frac{1}{16\varepsilon'}\frac{c_+\sigma_+ + C_{H_0}}{c_-\sigma_-} \tag{120}$$

with $c_- := \frac{\sqrt{2\pi}}{96c_c^2}$ and $c_+ := 4\hat{C}$, i.e., $c \leq 1/(4\sqrt{\varepsilon' C_\sigma})$. Inserting this estimate into (116) finishes the proof. □

5.2. Dynamical Typicality

In order to find an upper bound for the relative error in the dynamical typicality theorem, Theorem 3, as well, we need a lower bound for $|w_{\mu B}(t)| = |\mathbb{E}_\mu \langle \psi_t | B | \psi_t \rangle|$. Without any further assumptions on the Hamiltonian, we immediately find the following proposition:

Proposition 2 *Let $B \in \mathcal{L}(\mathcal{H})$ be Hermitian such that $b := \max\{b^+_{\min} - b^-_{\max}, b^-_{\min} - b^+_{\max}\} > 0$, where $b^\pm_{\min}$ and $b^\pm_{\max}$ are again the smallest and largest eigenvalues of $B^\pm$ (the positive and negative parts of $B = B^+ - B^-$). Let $\varepsilon > 0$, $t \in [0, \infty)$, and let μ be an arbitrary macro state. Then,*

$$\left|\mathbb{E}_\mu \langle \psi_t | B | \psi_t \rangle\right| \geq b, \tag{121}$$

and therefore $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that

$$\frac{\left|\langle \psi_t | B | \psi_t \rangle - \mathbb{E}_\mu \langle \psi_t | B | \psi_t \rangle\right|}{\left|\mathbb{E}_\mu \langle \psi_t | B | \psi_t \rangle\right|} \leq b^{-1} \cdot (\textit{right-hand side of } (35)). \tag{122}$$

Moreover, for every μ and B, every $T > 0$, and $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$,

$$\frac{1}{T}\int_0^T \frac{\left|\langle \psi_t | B | \psi_t \rangle - \mathbb{E}_\mu \langle \psi_t | B | \psi_t \rangle\right|^2}{\left|\mathbb{E}_\mu \langle \psi_t | B | \psi_t \rangle\right|^2}\, dt \leq \frac{\|B\|^2}{b^2 \varepsilon d_\mu}. \tag{123}$$

Proof Let (φ_k) be an orthonormal basis of $\mathcal{H}_\mu$ and define $\varphi_{k,t} := e^{-itH}\varphi_k$. Then,

$$\begin{aligned}
\left|\mathbb{E}_\mu \langle \psi_t | B | \psi_t \rangle\right| &= \frac{1}{d_\mu}\left|\mathrm{tr}(P_\mu e^{itH} B e^{-itH})\right| && \text{(124a)}\\
&= \frac{1}{d_\mu}\left|\sum_k \langle \varphi_{k,t} | B | \varphi_{k,t} \rangle\right| && \text{(124b)}\\
&\geq \frac{1}{d_\mu}\max\left\{\sum_k \langle \varphi_{k,t} | B^+ - B^- | \varphi_{k,t} \rangle, \sum_k \langle \varphi_{k,t} | B^- - B^+ | \varphi_{k,t} \rangle\right\} && \text{(124c)}\\
&\geq \max\{b^+_{\min} - b^-_{\max}, b^-_{\min} - b^+_{\max}\}, && \text{(124d)}
\end{aligned}$$

i.e., $\left|\mathbb{E}_\mu \langle \psi_t | B | \psi_t \rangle\right| \geq b$. The remaining claims now follow immediately from Theorem 3. □

Proposition 2 yields a good lower bound for $\left|\mathbb{E}_\mu\langle\psi_t|B|\psi_t\rangle\right|$ and therefore useful upper bounds for the relative errors if, for example, B is a positive or negative operator (and in this case, $b = b^+_{\min} > 0$ and $b = b^-_{\min} > 0$, respectively). If B has both positive and nonpositive eigenvalues, the bounds are only useful if the spectrum of B is rather special since in most cases one has $b \leq 0$. In particular, if $B = P_\nu$ for some macro state ν we have $b = 0$ and Proposition 2 is not applicable. However, with the help of the corrected and improved no-gaps delocalization, Theorem 7, we are able to prove an upper bound for the comparative errors which also applies to more general operators and in particular to the case $B = P_\nu$:

Theorem 11 *Let $\varepsilon > 0$, $\varepsilon' \in (0, \frac{1}{2})$ and let μ and ν be macro states such that $d_\mu, d_\nu > \max\{166, 4|\log_2 \varepsilon'|\}$. Let B be a Hermitian $D \times D$ matrix and let H be a random Hermitian $D \times D$ matrix as in Theorem 10. Then with probability at least $1 - \varepsilon'$, for each $t \in \mathbb{R}$, $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that*

$$\frac{\left|\langle\psi_t|B|\psi_t\rangle - \mathbb{E}_\mu\langle\psi_t|B|\psi_t\rangle\right|}{\left|\overline{\langle\psi_t|B|\psi_t\rangle}\right|} \leq \mathrm{LB(B,\psi)}^{-1}\min\left\{\frac{\sqrt{2}\|B\|}{\sqrt{\varepsilon d_\mu}}, \sqrt{\frac{2\|B\|\operatorname{tr}(|B|)}{\varepsilon d_\mu^2}}, \sqrt{\frac{18\pi^3\log(8/\varepsilon)}{d_\mu}}\|B\|\right\}, \tag{125}$$

whenever

$$\mathrm{LB}(B,\psi) := \max\left\{b^+_{\min}, \left(\frac{d_\mu}{4c_c\sqrt{KJD}}\right)^{16}\frac{\operatorname{tr}(B^+)}{d_\mu}\right\} - \min\left\{b^-_{max}, \frac{\operatorname{tr}(B^-)}{d_\mu}\right\} - \sqrt{2}\left(\frac{\|B\|}{\varepsilon d_\mu}\min\left\{\|B\|, \frac{\operatorname{tr}(|B|)}{d_\mu}\right\}\right)^{1/2} \tag{126}$$

is positive. Moreover, for every μ and B, every $T > 0$, and $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$,

$$\frac{1}{T}\int_0^T \frac{\left|\langle\psi_t|B|\psi_t\rangle - \mathbb{E}_\mu\langle\psi_t|B|\psi_t\rangle\right|^2}{\left|\overline{\langle\psi_t|B|\psi_t\rangle}\right|^2}\,dt \leq \frac{2\|B\|^2}{\mathrm{LB}(B,\psi)^2\varepsilon d_\mu}. \tag{127}$$

Proof It follows from (124) and (125) in [43] that $\left(1 - \frac{\varepsilon}{2}\right)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that

$$\left|\overline{\langle\psi_t|B|\psi_t\rangle} - M_{\mu B}\right| \leq \sqrt{2}\left(\frac{\|B\|}{\varepsilon d_\mu}\min\left\{\|B\|, \frac{\operatorname{tr}(|B|)}{d_\mu}\right\}\right)^{1/2}. \tag{128}$$

Therefore, it follows from the triangle inequality and Theorem 10 that with probability at least $1 - \varepsilon'$, $\left(1 - \frac{\varepsilon}{2}\right)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that

$$\begin{aligned}\left|\overline{\langle\psi_t|B|\psi_t\rangle}\right| &\geq |M_{\mu B}| - \sqrt{2}\left(\frac{\|B\|}{\varepsilon d_\mu}\min\left\{\|B\|, \frac{\operatorname{tr}(|B|)}{d_\mu}\right\}\right)^{1/2} && (129)\\ &\geq \max\left\{b^+_{\min}, \left(\frac{d_\mu}{4c_c\sqrt{KJD}}\right)^{16}\frac{\operatorname{tr}(B^+)}{d_\mu}\right\} - \min\left\{b^-_{\max}, \frac{\operatorname{tr}(B^-)}{d_\mu}\right\}\\ &\quad - \sqrt{2}\left(\frac{\|B\|}{\varepsilon d_\mu}\min\left\{\|B\|, \frac{\operatorname{tr}(|B|)}{d_\mu}\right\}\right)^{1/2} && (130)\\ &= \mathrm{LB}(B,\psi). && (131)\end{aligned}$$

It follows from Theorem 3 that $\left(1-\frac{\varepsilon}{2}\right)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that

$$\left|\langle\psi_t|B|\psi_t\rangle - \mathbb{E}_\mu\langle\psi_t|B|\psi_t\rangle\right| \leq \min\left\{\frac{\sqrt{2}\|B\|}{\sqrt{\varepsilon d_\mu}}, \sqrt{\frac{2\|B\|\operatorname{tr}(|B|)}{\varepsilon d_\mu^2}}, \sqrt{\frac{18\pi^3\log(8/\varepsilon)}{d_\mu}}\|B\|\right\}. \tag{132}$$

A combination of (131) and (132) yields the first claim. Similarly, a combination of (131) and (36) in Theorem 3 gives the second claim. $\square$

Note that for large d_μ (and if $\|B\|$ does not depend on d_μ), the lower bound for $\left|\overline{\langle\psi_t|B|\psi_t\rangle}\right|$, $\mathrm{LB}(B,\psi)$, is up to a small error equal to the lower bound for $|M_{\mu B}|$ that was proved in Theorem 10.

Finally, we give the proof of Theorem 5 which yields a somewhat better bound than the previous theorem in the case that B is a projection P_ν onto some macro space $\mathcal{H}_\nu$:

Proof of Theorem 5 Let (ϕ_n) be an orthonormal basis of eigenvectors of H. With $\psi_0 = \sum_n c_n\phi_n$, we find that

$$\overline{\langle\psi_t|P_\nu|\psi_t\rangle} = \sum_{n,m} c_n^* c_m \overline{e^{it(E_n-E_m)}}\langle\phi_n|P_\nu|\phi_m\rangle = \sum_n |c_n|^2\langle\phi_n|P_\nu|\phi_n\rangle. \tag{133}$$

Since $\sum_n |c_n|^2 = 1$ and with the help of Theorem 7, we find with probability at least $1-\varepsilon'$ that

$$\overline{\langle\psi_t|P_\nu|\psi_t\rangle} \geq \left(\frac{d_\nu}{4c_c\sqrt{KJD}}\right)^{16} \geq \left(\frac{\sqrt{\varepsilon' C_\sigma}d_\nu}{D}\right)^{16} \tag{134}$$

see also the proof of Theorem 10 for the first step. In the second step, we used that $c_c\sqrt{KJ} \leq 1/(4\sqrt{\varepsilon' C_\sigma})$, which can be shown similarly as in the proof of Theorem 4. This implies together with Theorem 3 (by using the first two bounds in (35)) that $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that

$$\frac{\left|\|P_\nu\psi_t\|^2 - \mathbb{E}_\mu\|P_\nu\psi_t\|^2\right|}{\overline{\|P_\nu\psi_t\|^2}} \leq \frac{1}{\sqrt{\varepsilon}}(C_\sigma\varepsilon')^{-8}\frac{1}{d_\mu}\sqrt{\min\{d_\mu, d_\nu\}}\left(\frac{D}{d_\nu}\right)^{16}. \tag{135}$$

Inserting the definition of s_μ, s_ν and s_{mc} thus proves the first claim. For the second claim observe that it follows from Theorem 3 that for every $T > 0$, $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that

$$\frac{1}{T}\int_0^T \left|\|P_\nu\psi_t\|^2 - \mathbb{E}_\mu\|P_\nu\psi_t\|^2\right|^2 \mathrm{d}t \leq \frac{1}{\varepsilon d_\mu}. \tag{136}$$

Together with (135), this implies that with probability $1-\varepsilon'$, for every $T > 0$, $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that

$$\frac{1}{T}\int_0^T \frac{\left|\|P_\nu\psi_t\|^2 - \mathbb{E}_\mu\|P_\nu\psi_t\|^2\right|^2}{\overline{\|P_\nu\psi_t\|^2}}\mathrm{d}t \leq \frac{1}{\varepsilon}(C_\sigma\varepsilon')^{-16}\frac{1}{d_\mu}\left(\frac{D}{d_\nu}\right)^{32}. \tag{137}$$

Now also the second claim follows immediately from the definition of s_μ, s_ν and s_{mc}. $\square$

6. An Improved Lower Bound for $M_{\mu\,\mathrm{eq}}$, $M_{\mathrm{eq}\,\nu}$ and $M_{\mathrm{eq}\,\mathrm{eq}}$ for Wigner-Type Matrices

We expect that in many cases a significantly stronger lower bound for the $M_{\mu\nu}$ than the one obtained with the help of our improved version of the no-gaps delocalization result from Rudelson and Vershynin [37], Theorem 4, should hold. More precisely, it is expected that for band matrices H with sufficiently large band width, the eigenvectors are delocalized, and in that situation we expect that $M_{\mu\nu} \approx \frac{d_\nu}{D}$.

In this section, we show with the help of a result from Ajanki, Erdős and Krüger [1] concerning the delocalization of eigenvectors that for a special class of random matrices a much better lower bound for $M_{\mu\,\mathrm{eq}}$, $M_{\mathrm{eq}\,\nu}$ and $M_{\mathrm{eq}\,\mathrm{eq}}$ can be obtained provided that the equilibrium macro space is very dominant (in a sense made precise below). These matrices do not have a band structure. However, the results strengthen the expectation that often the lower bound in Theorem 4 can be significantly improved. We quote the relevant statement from [1] as Theorem 12 and obtain from it lower bounds for $M_{\mu\nu}$ in Theorem 13.

For using the result of [1], we need a certain shift of perspective. So far in this paper, we always regarded D as fixed and considered a single, randomly chosen $D\times D$ matrix H. In contrast, in this section and the next, we will consider a *sequence* of random matrices $H^{(D)}$, one for every $D \in \mathbb{N}$. In fact, our reasoning also applies if we merely allow an infinite set of D values (say, all powers of 2), and a random matrix $H^{(D)}$ for every D from that set; but for simplicity, we will pretend we have an $H^{(D)}$ for every $D \in \mathbb{N}$. More precisely, we assume that, for every $D \in \mathbb{N}$, we are given a probability distribution $\mathbb{P}^{(D)}$ over the Hermitian $D \times D$ matrices. We will show that, under suitable assumptions on $\mathbb{P}^{(D)}$, certain estimates hold for *sufficiently large* D, but we are not necessarily able to make explicit how large D has to be.

In order to state the result of [1], we need their notion of *stochastic domination*:

Definition 4 (Stochastic domination). For two sequences $X = (X^{(D)})_D$ and $Y = (Y^{(D)})_D$ of nonnegative random variables we say that X is stochastically dominated by Y if there exists a function $D_0 : (0,\infty)^2 \to \mathbb{N}$ such that for all $\tau > 0$ and $\alpha > 0$,

$$\mathbb{P}\left(X^{(D)} > D^{\tau} Y^{(D)}\right) \leq D^{-\alpha} \qquad \forall D \geq D_0(\tau,\alpha). \tag{138}$$

In this case we write $X \prec Y$.

That is, $X \prec Y$ means that for large D it has high probability that X is not much larger than Y. The key fact for our purposes will be that under certain assumptions on $\mathbb{P}^{(D)}$, every eigenvector ϕ_n of $H = H^{(D)} \sim \mathbb{P}^{(D)}$ satisfies

$$\|\phi_n\|_\infty \prec \frac{1}{\sqrt{D}} \tag{139}$$

(more precise statement around (147) below). That is, each component of ϕ_n is not much larger than $1/\sqrt{D}$ (the magnitude that a component would have to have if all components had the same magnitude). Since if some components were much smaller than $1/\sqrt{D}$, others would have to be larger, this also entails that not a large fraction of the components can be much smaller than $1/\sqrt{D}$. But this still allows that a *small* fraction of the components could be arbitrarily small; for example, if d_ν (despite being a large number) is a small fraction of D, then all components of ϕ_n in $\mathcal{H}_\nu$ could be arbitrarily small, so $\langle\phi_n|P_\nu|\phi_n\rangle$ could be arbitrarily small, and then the expression (5) suggests that $M_{\mu\nu}$ could be arbitrarily small, and we would not obtain a useful lower bound for $M_{\mu\nu}$. In fact, (139) provides such a bound only if either d_μ or d_ν is sufficiently close to D, as the detailed analysis below confirms.

We now turn to describing the matrices considered in [1] and in this section.

Assumption 3 For every D, $H^{(D)} = (h_{ij}) \sim \mathbb{P}^{(D)}$ is a Hermitian $D \times D$ matrix of *Wigner type*, i.e., its entries h_{ij} are centered random variables and the entries $(h_{ij})_{i\le j}$ are independent. The matrix of variances $S = (\sigma_{ij}^2)$, defined by

$$\sigma_{ij}^2 := \mathbb{E}|h_{ij}|^2, \tag{140}$$

is *flat*, i.e.,

$$\sigma_{ij}^2 \le \frac{1}{D}, \qquad i,j = 1,\dots,D. \tag{141}$$

Let $p, P > 0$ and $L \in \mathbb{N}$ be parameters independent of D and let $\mu = (\mu_1, \mu_2, \dots)$ be a sequence of nonnegative real numbers.

Assumption 4 The matrix S is *uniformly primitive*, i.e.

$$(S^L)_{ij} \ge \frac{p}{D}, \qquad i,j = 1,\dots,D. \tag{142}$$

It can be shown that the corresponding *vector Dyson equation*

$$-\frac{1}{m_i(z)} = z + \sum_{j=1}^{D} \sigma_{ij}^2 m_j(z), \qquad \text{for all } i = 1,\dots,D \text{ and } z \in \mathbb{H} \tag{143}$$

for a function $m = (m_1, \dots, m_D) : \mathbb{H} \to \mathbb{H}^D$ on the complex upper half plane $\mathbb{H} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ has a unique solution [2].

Assumption 5 The matrix S induces a bounded solution of the vector Dyson equation, i.e., the unique solution of (143) corresponding to S is bounded,

$$|m_i(z)| \le P, \qquad i = 1,\dots,D, \quad z \in \mathbb{H}. \tag{144}$$

Assumption 6 The entries h_{ij} of the random matrix H have *bounded moments*, i.e.

$$\mathbb{E}|h_{ij}|^k \le \mu_k \sigma_{ij}^k, \qquad k \in \mathbb{N}, \quad i,j = 1,\dots,D. \tag{145}$$

Sufficient conditions for Assumption 5 are given in [2] Theorem 6.1. The reason we make these assumptions is that under these conditions, Ajanki, Erdős and Krüger (2017) [1] proved the following theorem concerning the delocalization of the eigenvectors of H:

Theorem 12 ([1] Corollary 1.14). *Consider a sequence $(\mathbb{P}^{(D)})_{D\in\mathbb{N}}$ of probability distributions over the Hermitian $D \times D$ matrices with $H^{(D)} \sim \mathbb{P}^{(D)}$ satisfying Assumptions 3–6. Let $E_1^{(D)} \le \dots \le E_D^{(D)}$ be the eigenvalues of the random matrix $H^{(D)}$ and $\phi_n^{(D)} \in \mathbb{C}^D$ a normalized eigenvector of $H^{(D)}$ with eigenvalue $E_n^{(D)}$. Then for every sequence of unit vectors $b^{(D)} \in \mathbb{C}^D$ and every sequence $n^{(D)} \in \{1,\dots,D\}$,*

$$\left|\left\langle b^{(D)} \middle| \phi_{n^{(D)}}^{(D)} \right\rangle\right| \prec \frac{1}{\sqrt{D}}. \tag{146}$$

In particular, the eigenvectors are completely delocalized, i.e.,

$$\left\|\phi_{n^{(D)}}^{(D)}\right\|_\infty \prec \frac{1}{\sqrt{D}}. \tag{147}$$

The function $D_0(\tau, \alpha)$ implicit in the $\prec$ symbol in (147) depends only on the constants $p, P, L, (\mu_k)_{k\in\mathbb{N}}$ from Assumptions 3–6.

With the help of Theorem 12, we find the following lower bounds for the $M_{\mu B}$:

Theorem 13 (Lower bounds for $|M_{\mu B}|$). *Consider a sequence $(\mathbb{P}^{(D)})_{D\in\mathbb{N}}$ of probability distributions over the Hermitian $D\times D$ matrices with $H^{(D)}\sim\mathbb{P}^{(D)}$ satisfying Assumptions 3-6 and let $D_0:(0,\infty)^2\to\mathbb{N}$ be the function provided by Theorem 12 for (147). Let $\tau>0,\alpha>1$, $D\geq D_0(\tau,\alpha)$, and let B be a Hermitian $D\times D$ matrix. Then with probability at least $1-D^{-\alpha+1}$, it holds for every macro state μ that*

$$|M_{\mu B}|\geq \max\left\{b^+_{\min},\frac{\operatorname{tr}(B^+)}{d_\mu}\left(1-\frac{D-d_\mu}{D^{1-2\tau}}\right)\right\}-\min\left\{b^-_{\max},\frac{\operatorname{tr}(B^-)}{d_\mu}\right\}. \tag{148}$$

In particular, if $B=P_\nu$ for some macro state ν, then

$$M_{\mu\nu}\geq\frac{d_\nu}{d_\mu}\left(1-\frac{D-d_\mu}{D^{1-2\tau}}\right), \tag{149}$$

and, moreover,

$$M_{\mu\nu}\geq 1-\frac{D-d_\nu}{D^{1-2\tau}}\,. \tag{150}$$

Proof Let $\tau>0,\alpha>0,D\geq D_0(\tau,\alpha)$. Since D_0 does not depend on the sequence $n^{(D)}$,

$$\mathbb{P}^{(D)}\left(\|\phi_n^{(D)}\|_\infty>D^\tau D^{-1/2}\right)\leq D^{-\alpha} \tag{151}$$

for all $n=1,\ldots,D$. Writing $\mathbb{P}$ for $\mathbb{P}^{(D)}$ and ϕ_n for $\phi_n^{(D)}$, we obtain that

$$\mathbb{P}\left(\forall n:\|\phi_n\|_\infty\leq D^{-1/2+\tau}\right)=1-\mathbb{P}\left(\exists n:\|\phi_n\|_\infty>D^{-1/2+\tau}\right) \tag{152a}$$

$$\geq 1-\sum_{n=1}^{D}\mathbb{P}\left(\|\phi_n\|_\infty>D^{-1/2+\tau}\right) \tag{152b}$$

$$\geq 1-D^{-\alpha+1}. \tag{152c}$$

Now assume that $\|\phi_n\|_\infty\leq D^{-1/2+\tau}$ for all $n=1,\ldots,D$. Then,

$$\langle\phi_n|P_\mu|\phi_n\rangle=1-\sum_{\mu'\neq\mu}\langle\phi_n|P_{\mu'}|\phi_n\rangle \tag{153a}$$

$$=1-\sum_{l\in[D]\setminus I_\mu}|\phi_n(l)|^2 \tag{153b}$$

$$\geq 1-\frac{D-d_\mu}{D^{1-2\tau}} \tag{153c}$$

where $I_\mu=\{d_1+\cdots+d_{\mu-1}+1,\ldots,d_1+\cdots+d_\mu\}$.

As in the proof of Theorem 4, we have that

$$|M_{\mu B}|\geq\frac{1}{d_\mu}\left(\sum_n\langle\phi_n|P_\mu|\phi_n\rangle\langle\phi_n|B^+|\phi_n\rangle-\sum_m\langle\phi_m|P_\mu|\phi_m\rangle\langle\phi_m|B^-|\phi_m\rangle\right), \tag{154}$$

where B^+ and B^- denote the positive and negative part of B, respectively.

We find that

$$\sum_n\langle\phi_n|P_\mu|\phi_n\rangle\langle\phi_n|B^+|\phi_n\rangle\geq\operatorname{tr}(B^+)\left(1-\frac{D-d_\mu}{D^{1-2\tau}}\right), \tag{155}$$

and together with (109a), (109c) and (109d) we obtain that

$$|M_{\mu B}|\geq \max\left\{b^+_{\min},\frac{\operatorname{tr}(B^+)}{d_\mu}\left(1-\frac{D-d_\mu}{D^{1-2\tau}}\right)\right\}-\min\left\{b^-_{\max},\frac{\operatorname{tr}(B^-)}{d_\mu}\right\}. \tag{156}$$

In particular, if $B = P_\nu$ for some macro state ν, then $B^+ = P_\nu$, $B^- = 0$, $\mathrm{tr}(B^+) = d_\nu$, $b^\pm_{\min} = b^-_{\max} = 0$, $b^+_{\max} = 1$, and thus

$$M_{\mu\nu} \geq \frac{d_\nu}{d_\mu}\left(1 - \frac{D - d_\mu}{D^{1-2\tau}}\right). \tag{157}$$

Alternatively, in the case that $B = P_\nu$ we can apply the estimate in (153c) to $\langle\phi_n|P_\nu|\phi_n\rangle$, which immediately yields

$$M_{\mu\nu} = \frac{1}{d_\mu}\sum_n \langle\phi_n|P_\mu|\phi_n\rangle\langle\phi_n|P_\nu|\phi_n\rangle \geq 1 - \frac{D - d_\nu}{D^{1-2\tau}}. \tag{158}$$

□

If the macro state μ resp. ν is such that $(D - d_\mu)D^{-1+2\tau} > 1$ resp. $(D - d_\nu)D^{-1+2\tau} > 1$, then the lower bound in (149) resp. (150) becomes negative and therefore useless since we always have the trivial bound $M_{\mu\nu} \geq 0$. However, if μ or ν is the equilibrium macro space in the sense that the corresponding macro space is extremely dominant, more precisely, if $d_{\mathrm{eq}} = D - o(D^{1-2\tau})$, then the lower bounds for the $M_{\mu\nu}$ are nontrivial. If $\nu = \mathrm{eq}$, then (150) implies that $M_{\mu\nu} \gtrsim 1 \approx \frac{d_{\mathrm{eq}}}{D}$ and if $\mu = \mathrm{eq}$, then (149) shows that $M_{\mu\nu} \gtrsim \frac{d_\nu}{d_\mu} \approx \frac{d_\nu}{D}$, in agreement with our expectations.

7. Consequences of the Eigenstate Thermalization Hypothesis

Another result, due to Cipolloni, Erdős and Henheik (CEH) [6], shows that Wigner matrices (i.e., $H = H^*$ such that h_{ij} for $i \leq j$ are centered, independent random variables with bounded moments, the h_{ij} for $i < j$ are identically distributed, and the h_{ii} are identically distributed) satisfy a version of the eigenstate thermalization hypothesis (ETH) that implies that the eigenvectors are delocalized. Although the matrices we are most interested in, the band matrices, are not Wigner matrices, we make explicit in this section which lower bounds on $M_{\mu\nu}$ (essentially versions of $M_{\mu\nu} \approx d_\nu/D$) would follow from the ETH in the version formulated by CEH (Proposition 3) and what they would imply about the relative error of generalized normal typicality (Corollary 5). After all, as mentioned in the beginning of Sect. 6, it is believed that for band matrices with sufficiently wide band, all eigenvectors are delocalized.

We begin by formulating the precise condition:

Definition 5 (ETH according to CEH [6]). We say that a sequence $(\mathbb{P}^{(D)})_{D\in\mathbb{N}}$ of probability distributions over the Hermitian $D\times D$ matrices satisfies the CEH-version of ETH if for every sequence $(B^{(D)})_{D\in\mathbb{N}}$ of $D\times D$ matrices with $\|B^{(D)}\| \leq 1$ and every $\gamma > 0$ and $\xi > 0$, there is $\tilde{D}_0 \in \mathbb{N}$ such that for $D \geq \tilde{D}_0$, it has probability at least $1 - D^{-\gamma}$ that

$$\max_{i,j\in[D]}\left|\langle\phi_i^{(D)}|B^{(D)}|\phi_j^{(D)}\rangle - \frac{\mathrm{tr}(B^{(D)})}{D}\delta_{ij}\right| \leq \frac{D^\xi}{D}\,\mathrm{tr}(|\mathring{B}^{(D)}|^2)^{1/2}, \tag{159}$$

where $\phi_1^{(D)},\ldots,\phi_D^{(D)}$ is any orthonormal eigenbasis of $H^{(D)} \sim \mathbb{P}^{(D)}$ and $\mathring{B}^{(D)} = B^{(D)} - \mathrm{tr}(B^{(D)})/D$ denotes the traceless part of B.

In that case, we obtain in particular that for any sequence $(B^{(D)})$ of Hermitian $D\times D$ matrices, any $\xi > 0$ and $\gamma > 0$, sufficiently large D, $B = B^{(D)}$ and every orthonormal eigenbasis $\phi_1, \ldots, \phi_D$ of $H = H^{(D)}$,

$$\langle \phi_n | B | \phi_n \rangle \geq \frac{\operatorname{tr}(B)}{D} - \frac{D^\xi}{D} \operatorname{tr}(|\mathring{B}|^2)^{1/2} \tag{160}$$

for all $n = 1, \ldots, D$ simultaneously with probability at least $1 - D^{-\gamma}$. Now recall that, if H is nondegenerate, then

$$M_{\mu B} = \frac{1}{d_\mu} \sum_n \langle \phi_n | P_\mu | \phi_n \rangle \langle \phi_n | B | \phi_n \rangle \,. \tag{161}$$

Proposition 3 (Lower bound for $|M_{\mu B}|$). *Let $\xi > 0$, let H be a Hermitian $D \times D$ matrix with orthonormal eigenbasis $\{\phi_1, \ldots, \phi_D\}$, and let B be a Hermitian $D \times D$ matrix such that (160) is satisfied for all $n = 1, \ldots, D$. Then, for every macro state μ,*

$$|M_{\mu B}| \geq \frac{|\operatorname{tr}(B)|}{D} - \frac{D^\xi}{D} \operatorname{tr}(|\mathring{B}|^2)^{1/2}. \tag{162}$$

In particular,

$$M_{\mu\nu} \geq \frac{d_\nu}{D} \left(1 - \frac{D^\xi}{\sqrt{d_\nu}} \right). \tag{163}$$

for every macro state ν. Moreover,

$$M_{\mu\nu} \geq \frac{d_\nu}{D} \left(1 - \frac{D^\xi}{\sqrt{d_\mu}} \right). \tag{164}$$

Proof From (160), we obtain that

$$M_{\mu B} = \frac{1}{d_\mu} \sum_n \langle \phi_n | P_\mu | \phi_n \rangle \langle \phi_n | B | \phi_n \rangle \geq \frac{\operatorname{tr}(B)}{D} - \frac{D^\xi}{D} \operatorname{tr}(|\mathring{B}|^2)^{1/2}. \tag{165}$$

Similarly, one finds that

$$M_{\mu B} \leq \frac{\operatorname{tr}(B)}{D} + \frac{D^\xi}{D} \operatorname{tr}(|\mathring{B}|^2)^{1/2} \tag{166}$$

and therefore

$$\left| M_{\mu B} - \frac{\operatorname{tr}(B)}{D} \right| \leq \frac{D^\xi}{D} \operatorname{tr}(|\mathring{B}|^2)^{1/2}. \tag{167}$$

With the help of the reverse triangle inequality we obtain (162). If $B = P_\nu$ for some macro state ν we have $\operatorname{tr}(B) = d_\nu$, $|\mathring{B}|^2 = P_\nu - 2P_\nu d_\nu/D + d_\nu^2/D^2$, therefore $\operatorname{tr}(|\mathring{B}|^2) = d_\nu - 2d_\nu^2/D + d_\nu^2/D^2 \leq d_\nu$ and (163) thus follows immediately from (162).

Concerning the last bound, observe that (160) for $B = P_\mu$ implies

$$M_{\mu\nu} = \frac{1}{d_\mu} \sum_n \langle \phi_n | P_\mu | \phi_n \rangle \langle \phi_n | P_\nu | \phi_n \rangle \geq \frac{d_\nu}{d_\mu} \left(\frac{d_\mu}{D} - \frac{D^\xi \sqrt{d_\mu}}{D} \right). \tag{168}$$

□

For large dimensions D, we find for any macro states μ and ν that $M_{\mu\nu} \gtrsim \frac{d_\nu}{D}$ (provided that d_ν is large enough such that $\frac{d_\nu}{D} \gg D^{\xi-1}\sqrt{d_\nu}$), in agreement with our expectations. For the relative errors of the $\|P_\nu \psi_t\|^2$, we obtain the following.

Corollary 5 *Suppose that* $(\mathbb{P}^{(D)})$ *satisfies the CEH-version of ETH, and that it is a continuous distribution for every* D*. Let* $\varepsilon, \delta, \xi, \gamma > 0$*, and let* μ *and* ν *be arbitrary macro states such that* $\sqrt{d_\nu} \geq 2D^\xi$ *or* $\sqrt{d_\mu} \geq 2D^\xi$*, i.e.,*

$$s_\nu \geq 2\xi s_{\mathrm{mc}} + \frac{2k_B}{N}\ln 2 \quad \text{or} \quad s_\mu \geq 2\xi s_{\mathrm{mc}} + \frac{2k_B}{N}\ln 2. \tag{169}$$

Then, for sufficiently large D *and with probability at least* $1 - D^{-\gamma}$*,* $(1-\varepsilon)$*-most* $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ *are such that for* $(1-\delta)$*-most* $t \in [0, \infty)$*,*

$$\frac{\left|\|P_\nu \psi_t\|^2 - M_{\mu\nu}\right|}{M_{\mu\nu}} \leq \frac{8}{\sqrt{\varepsilon\delta}} \exp\left(-\frac{N}{k_B}\left(\max\{s_\mu, s_\nu\} - s_{\mathrm{mc}} + \frac{1}{2}\min\{s_\nu, s_\mu\}\right)\right). \tag{170}$$

Proof It follows immediately from (163) that

$$M_{\mu\nu} \geq \frac{d_\nu}{2D}. \tag{171}$$

Together with the upper bound for the absolute error from Theorem 2, we obtain that $(1-\varepsilon)$-most $\psi_0 \in \mathbb{S}(\mathcal{H}_\mu)$ are such that for $(1-\delta)$-most $t \in [0, \infty)$,

$$\frac{\left|\|P_\nu \psi_t\|^2 - M_{\mu\nu}\right|}{M_{\mu\nu}} \leq \frac{8}{\sqrt{\varepsilon\delta}} \frac{D}{d_\nu\sqrt{d_\mu}} \left(\min\left\{1, \frac{d_\nu}{d_\mu}\right\}\right)^{1/2} \tag{172}$$

The claim now follows immediately from the definition of s_μ, s_ν and s_{mc}. $\square$

The relative errors in Corollary 5 are small for large N if $s_{\mathrm{mc}} < \max\{s_\mu, s_\nu\} + \min\{s_\mu, s_\nu\}/2$, i.e., we recover the condition we found in the case of normal typicality, see also Remark 1.

Acknowledgements

We thank László Erdős and Roman Vershynin for helpful discussions. C.V. gratefully acknowledges financial support by the German Academic Scholarship Foundation. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—TRR 352—Project-ID 470903074.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors have no conflict of interest.

A. Appendix—No Resonances

In this appendix, we provide a proof of the fact that Hermitian matrices whose joint distribution of their entries is absolutely continuous with respect to the Lebesgue measure have, with probability 1, no degeneracies and no resonances (i.e., also the eigenvalue gaps are nondegenerate). This fact is widely known, but we could not find a proof in the literature.

Proposition 4 (No resonances). *Let H be a random Hermitian $D \times D$ matrix with eigenvalues $\lambda_1, \dots, \lambda_D$ such that the joint distribution of its entries is absolutely continuous with respect to the Lebesgue measure. Then,*

$$\mathbb{P}\left(\lambda_i - \lambda_j = \lambda_k - \lambda_l \textit{ for some } (i \neq j \textit{ or } k \neq l) \textit{ and } (i \neq k \textit{ or } j \neq l)\right) = 0. \tag{173}$$

We begin with some preparations for the proof. Let $S(D)$ denote the set of Hermitian $D \times D$ matrices and $U(D)$ the unitary group of $D \times D$ matrices, here regarded as orthonormal bases of the underlying Hilbert space. We define $\chi : \mathbb{R}^D \times U(D) \to S(D)$ by

$$\chi(\lambda_1, \dots, \lambda_D, \psi_1, \dots, \psi_D) = \sum_{j=1}^{D} \lambda_j |\psi_j\rangle\langle\psi_j|, \tag{174}$$

where ψ_j denotes the jth column of the unitary matrix $[\psi_1, \dots, \psi_D]$. Since the matrix $\chi(\lambda_1, \dots, \lambda_D, \psi_1, \dots, \psi_D)$ remains constant when the phase of the ψ_i is changed, χ defines a mapping $\varphi : \mathbb{R}^D \times U(D)/U(1)^D \to S(D)$.

Lemma 1 *The map φ defined above is smooth.*

Proof First note that the function χ in (174) defined on $\mathbb{R}^D \times \mathbb{C}^{D\times D} \to \mathbb{C}^{D\times D}$ (extended in the obvious way) is smooth since all components of $\chi(\lambda_1, \dots, \lambda_D, \psi_1, \dots, \psi_D)$ are polynomials in the λ_j and the $(\psi_j)_k$. Since $U(D)$ is a (embedded) submanifold of $\mathbb{C}^{D\times D}$, the restriction of χ to $\mathbb{R}^D \times U(D)$ is also smooth [24, Prop. 5.27]. It remains to show that also the replacing of $U(D)$ by the quotient $U(D)/U(1)^D$ does not destroy the smoothness. To this end, consider the map

$$f : \mathbb{C}^D \backslash \{0\} \to \mathbb{C}^{D\times D},\ \psi \mapsto f(\psi) = |\psi\rangle\langle\psi|. \tag{175}$$

This map is obviously smooth and remains smooth when considered as a map on $(\mathbb{C}^D \backslash \{0\})/U(1)$. Thus, we can conclude that φ is smooth. $\square$

Proof of Proposition 4 Since $U(1)^D$ is a closed subset of $U(D)$ (because limits of diagonal matrices are diagonal), and since $U(1)^D$ is a Lie subgroup of $U(D)$, it follows that the set $U(D)/U(1)^D$ of left cosets is a smooth manifold [24, Thm. 21.17]. It is well known that $\dim U(D) = D^2$, $\dim U(D)/U(1)^D = D^2 - D$, $\dim S(D) = D^2$ and thus $\dim\left(\mathbb{R}^D \times (U(D)/U(1)^D)\right) = \dim S(D)$. The spaces $U := \mathbb{R}^D \times (U(D)/U(1)^D)$ and $V := S(D)$ are therefore smooth manifolds of the same dimension and $\varphi|_{\mathbb{R}^D \times (U(D)/U(1)^D)} : U \to V$ is smooth by Lemma 1. Since the set

$$N_{\text{res}} := \Big\{(x_1, \dots, x_D) \in \mathbb{R}^D :$$

$$x_i - x_j = x_k - x_l \text{ for some } (i \neq j \text{ or } k \neq l) \text{ and } (i \neq k \text{ or } j \neq l)\Big\} \tag{176}$$

is a null set in $\mathbb{R}^D$, it follows that $N_{\mathrm{res}} \times (U(D)/U(1)^D)$ is a null set in $\mathbb{R}^D \times (U(D)/U(1)^D)$. Generally, if $f : U \to V$ is a smooth mapping between smooth manifolds U, V of equal dimension and $N \subset U$ is a null set, then $f(N)$ is a null set in V [24, Thm. 6.9]. Thus,

$$\varphi(N_{\mathrm{res}} \times (U(D)/U(1)^D)) = \chi(N_{\mathrm{res}} \times U(D)) \tag{177}$$

is a null set in $S(D)$. Altogether we have shown that the set of Hermitian matrices that have resonances has measure zero. Finally, the absolute continuity of the joint distribution of the entries of H with respect to the Lebesgue measure immediately proves the claim. □

B. Appendix—Numerical Examples

We briefly describe the numerical simulations we used to create Figs. 1 and 2. These simulations serve for illustrating our results and stating some further conjectures. We partition the D-dimensional Hilbert space $\mathcal{H} := \mathbb{C}^D = \bigoplus_{\nu=1}^4 \mathbb{C}^{d_\nu}$ into four macro spaces $\mathcal{H}_\nu$ of dimension d_ν with $d_\nu \ll d_{\nu+1}$ for $\nu = 1, 2, 3$; more precisely, we choose $d_\nu = 2 \times 10^{\nu-1}$. Then $\mathcal{H}_1$ is spanned by the first d_1 canonical basis vectors, $\mathcal{H}_2$ by the next d_2 canonical basis vectors and so on, and $\mathcal{H}_4$, the largest macro space in the decomposition, corresponds to the "equilibrium space." The Hamiltonian $H = (h_{jk})$ is modeled by a Hermitian random $D \times D$ matrix with a band structure which means that it couples neighboring macro spaces more strongly than distant ones, see also Fig. 3. The entries of H satisfy $h_{jj} \sim \mathcal{N}(0, \sigma_{jj}^2)$ and $h_{jk} \sim \mathcal{N}(0, \sigma_{jk}^2/2) + i\mathcal{N}(0, \sigma_{jk}^2/2)$ for $j \neq k$, where

$$\sigma_{jk}^2 := \exp(-s|j-k|) \tag{178}$$

with some parameter $s > 0$ that controls the bandwidth of the random matrix, i.e., the variances decrease exponentially in the distance from the diagonal. Note that this model is not covered by our examples in Sect. 4.3 since the σ_{jk} cannot be bounded by a positive D-independent constant or by D^α for some $\alpha \in \mathbb{R}$ from below. However, it also suggests that similar results should hold in more general situations than in the ones we were able to study.

References

[1] Ajanki, O.H., Erdős, L., Krüger, T.: Universality for general Wigner-type matrices. Probab. Theory Relat. Fields **169**, 667–727 (2017). arXiv:1506.05098

[2] Ajanki, O.H., Erdős, L., Krüger, T.: Quadratic Vector Equations On Complex Upper Half-Plane. Memoirs of the American Mathematical Society, 261, (2019). arXiv:1506.05095

[3] Balz, B., Richter, J., Gemmer, J., Steinigeweg, R., Reimann, P.: Dynamical typicality for initial states with a preset measurement statistics of several commuting observables. In Binger, F., Correa, L.A., Gogolin, C., Anders, J., Adesso, G., editors, Thermodynamics in the Quantum Regime, chapter 17, pages 413–433. Springer, Cham, (2019). arXiv:1904.03105

[4] Bartsch, C., Gemmer, J.: Dynamical typicality of quantum expectation values. Phys. Rev. Lett. **102**, 110403 (2009). arXiv:0902.0927

[5] Bourgade, P., Yau, H.-T., Yin, J.: Random band matrices in the delocalized phase I: quantum unique ergodicity and universality. Commun. Pure Appl. Math. **73**, 1526–1596 (2020). arXiv:1807.01559

[6] Cipolloni, G., Erdős, L., Henheik, J.: Eigenstate thermalisation at the edge for Wigner matrices, (2023). Preprint. arXiv:2309.05488

[7] Cipolloni, G., Erdős, L., Henheik, J., Kolupaiev, O.: Gaussian fluctuations in the Equipartition Principle for Wigner matrices. Forum of Mathematics Sigma **11**, E74 (2023). arXiv:2301.05181

[8] Cipolloni, G., Erdős, L., Schröder, D.: Eigenstate thermalization hypothesis for Wigner matrices. Commun. Math. Phys. **388**, 1005–1048 (2021). arXiv:2012.13215

[9] Deutsch, J.M.: Quantum statistical mechanics in a closed system. Phys. Rev. A **43**, 2046–2049 (1991)

[10] Erdős, L., Knowles, A.: Quantum diffusion and delocalization for band matrices with general distribution. Annales de l'Institut Henri Poincaré **12**, 1227–1319 (2011). arXiv:1005.1838

[11] Erdős, L., Knowles, A.: Quantum diffusion and eigenfunction delocalization in a random matrix band model. Commun. Math. Phys. **303**, 509–554 (2011). arXiv:1002.1695

[12] Erdős, L., Krüger, T., Schröder, D.: Random matrices with slow correlation decay. Forum of Mathematics, Sigma **7**, E8 (2019). arXiv:1705.10661

[13] Erdős, L., Schlein, B., Yau, H.-T.: Local semicircle law and complete delocalization for Wigner random matrices. Commun. Math. Phys. **287**, 641–655 (2009). arXiv:0803.0542

[14] Erdős, L., Schlein, B., Yau, H.-T.: Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. Ann. Probab. **37**, 815–852 (2009). arXiv:0711.1730

[15] Erdős, L., Yau, H.-T., Yin, J.: Bulk universality for generalized Wigner matrices. Probab. Theory Relat. Fields **154**, 341–407 (2012). arXiv:1001.3453

[16] Goldstein, S., Hara, T., Tasaki, H.: Time Scales in the Approach to Equilibrium of Macroscopic Quantum Systems. Phys. Rev. Lett. **111**, 140401 (2013). arXiv:1307.0572

[17] Goldstein, S., Hara, T., Tasaki, H.: The approach to equilibrium in a macroscopic quantum system for a typical nonequilibrium subspace, (2014). Preprint, arXiv:1402.3380

[18] Goldstein, S., Hara, T., Tasaki, H.: Extremely quick thermalization in a macroscopic quantum system for a typical nonequilibrium subspace. New J. Phys. **17**, 045002 (2015). arXiv:1402.0324

[19] Goldstein, S., Lebowitz, J.L., Mastrodonato, C., Tumulka, R., Zanghì, N.: Normal typicality and von Neumann's quantum ergodic theorem. Proc. Royal Soc. A **466**(2123), 3203–3224 (2010). arXiv:0907.0108

[20] Goldstein, S., Lebowitz, J.L., Mastrodonato, C., Tumulka, R., Zanghì, N.: On the approach to thermal equilibrium of macroscopic quantum systems. Phys. Rev. E **81**, 011109 (2010). arXiv:0911.1724

[21] Goldstein, S., Lebowitz, J.L., Tumulka, R., Zanghì, N.: Long-time behavior of macroscopic quantum systems. Euro. Phys. J. H **35**, 173–200 (2010). arXiv:1003.2129

[22] Latala, R.: Some estimates of norms of random matrices. Proc. Am. Math. Soc. **133**, 1273–1282 (2005)

[23] Lebowitz, J.L.: Macroscopic laws, microscopic dynamics, time's arrow and Boltzmann's entropy. Phys. A **194**, 1–27 (1993)

[24] Lee, J.M.: Introduction to Smooth Manifolds, 2nd edn. Springer, Berlin (2013)

[25] Luh, K., O'Rourke, S.: Eigenvector delocalization for non-Hermitian random matrices and applications. Random Struct. Algorithms **57**, 169–210 (2020). arxiv:1810.00489

[26] Lytova, A., Tikhomirov, K.: On delocalization of eigenvectors of random non-Hermitian matrices. Probab. Theory Relat. Fields **177**, 465–524 (2020). arxiv:1810.01590

[27] Müller, M.P., Gross, D., Eisert, J.: Concentration of measure for quantum states with a fixed expectation value. Commun. Math. Phys. **303**, 785–824 (2011). arxiv:1003.4982

[28] Normal number. In: Wikipedia, the free encyclopedia (accessed 12/12/2021). http://en.wikipedia.org/wiki/Normal_number

[29] Reimann, P.: Typicality for generalized microcanonical ensembles. Phys. Rev. Lett. **99**, 160404 (2007). arxiv:0710.4214

[30] Reimann, P.: Foundations of statistical mechanics under experimentally realistic conditions. Phys. Rev. Lett. **101**, 190403 (2008). arxiv:0810.3092

[31] Reimann, P.: Generalization of von Neumann's approach to thermalization. Phys. Rev. Lett. **115**, 010403 (2015). arxiv:1507.00262

[32] Reimann, P.: Dynamical typicality approach to eigenstate thermalization. Phys. Rev. Lett. **120**, 230601 (2018). arxiv:1806.03193

[33] Reimann, P.: Dynamical typicality of isolated many-body quantum systems. Phys. Rev. E **97**, 062129 (2018). arxiv:1805.07085

[34] Reimann, P., Gemmer, J.: Why are macroscopic experiments reproducible? Imitating the behavior of an ensemble by single pure states. Phys. A **552**, 121840 (2020). arxiv:2005.14626

[35] Rudelson, M., Vershynin, R.: Delocalization of eigenvectors of random matrices with independent entries. Duke Math. J. **164**, 2507–2538 (2015). arxiv:1306.2887

[36] Rudelson, M., Vershynin, R.: Small ball probabilities for linear images of high-dimensional distributions. Int. Math. Res. Not. **2015**(19), 9594–9617 (2015). arxiv:1402.4492

[37] Rudelson, M., Vershynin, R.: No-gaps delocalization for general random matrices. Geom. Funct. Anal. **26**, 1716–1776 (2016). arxiv:1506.04012

[38] Short, A.J.: Equilibration of quantum systems and subsystems. New J. Phys. **13**, 053009 (2011). arxiv:1012.4622

[39] Short, A.J., Farrelly, T.C.: Quantum equilibration in finite time. New J. Phys. **14**, 013063 (2012)

[40] Srednicki, M.: Chaos and quantum thermalization. Phys. Rev. E **50**, 888–901 (1994). arXiv:cond-mat/9403051

[41] Stein, E.M., Weiss, G.: Introduction to Fourier analysis on Euclidean spaces (PMS-32). Princeton University Press, USA (1971)

[42] Strasberg, P., Winter, A., Gemmer, J., Wang, J.: Classicality, Markovianity, and local detailed balance from pure state dynamics. Phys. Rev. A **108**, 012225 (2023). arXiv:2209.07977

[43] Teufel, S., Tumulka, R., Vogel, C.: Time evolution of typical pure states from a macroscopic Hilbert subspace. J. Stat. Phys. **190**, 69 (2023). arxiv:2210.10018

[44] Teufel, S., Tumulka, R., Vogel, C.: Canonical Typicality for Other Ensembles than Micro-Canonical. Annales Henri Poincaré, (2024). arXiv:2307.15624

[45] von Neumann, J.: Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik. Zeitschrift für Physik, 57:30–70, 1929. English translation: European Physical Journal H, **35**: 201–237, (2010). arXiv:1003.2133

[46] Wigner, E.P.: Characteristic vectors of bordered matrices with infinite dimensions. Ann. Math. **62**(3), 548–564 (1955)

[47] Yang, F., Yau, H.-T., Yin, J.: Delocalization and quantum diffusion of random band matrices in high dimensions I: Self-energy renormalization, (2021). Preprint. arXiv:2104.12048

Stefan Teufel, Roderich Tumulka and Cornelia Vogel
Mathematics Institute
Eberhard Karls University Tübingen
Auf der Morgenstelle 10
72076 Tübingen
Germany
e-mail: stefan.teufel@uni-tuebingen.de;
roderich.tumulka@uni-tuebingen.de;
cornelia.vogel@uni-tuebingen.de

Communicated by Christian Maes.

Received: April 5, 2023.

Accepted: November 23, 2024.