Generalization of von Neumann's approach to thermalization: Supplemental material

Peter Reimann
Universität Bielefeld, Fakultät für Physik, 33615 Bielefeld, Germany

PACS numbers: 05.30.-d, 03.65.-w

I. GENERAL FRAME

For the sake of convenience, we recall here once again the general set-up from the main paper: We consider a Hamiltonian with eigenvalues E_n and eigenvectors $|n\rangle$ on a D-dimensional Hilbert space $\mathcal{H}$,

$$H = \sum_{n=1}^{D} E_n |n\rangle\langle n| \ , \tag{1}$$

and it is assumed that the differences $E_m - E_n$ are finite and mutually different for all pairs (m, n) with $m \neq n$. System states are described by density operators ρ with matrix elements $\rho_{mn} := \langle m|\rho|n\rangle$, evolving in time according to the standard Liouville-von-Neumann equation. Given the initial condition $\rho(0)$, the solution $\rho(t)$ of the latter equation can be written by means of the propagator (unitary time evolution operator) $\mathcal{U}_t := e^{-iHt}$ ($\hbar = 1$) in the well-known form

$$\rho(t) = \mathcal{U}_t \rho(0) \mathcal{U}_t^\dagger \ , \tag{2}$$

implying for the matrix elements that

$$\rho_{mn}(t) = \rho_{mn}(0)\, e^{i[E_n - E_m]t} \ . \tag{3}$$

In particular $\rho_{nn}(t) = \rho_{nn}(0)$ for all t and n (conserved quantities).

Observables are modeled by Hermitian operators $A : \mathcal{H} \to \mathcal{H}$ with eigenvalues a_ν and eigenvectors $|\nu\rangle$,

$$A = \sum_{\nu=1}^{D} a_\nu |\nu\rangle\langle \nu| \ , \tag{4}$$

and their expectation values are given by

$$\langle A\rangle_{\rho(t)} := \mathrm{Tr}\{\rho(t)A\} \ . \tag{5}$$

The largest and smallest eigenvalues of A are denoted as

$$a_{max} := \max_\nu a_\nu \tag{6}$$

$$a_{min} := \min_\nu a_\nu \ , \tag{7}$$

where the maximization and the minimization are over all $\nu = 1, \ldots, D$. Accordingly, the range Δ_A of A is defined as

$$\Delta_A := a_{max} - a_{min} \ . \tag{8}$$

As in the main paper, the identity operator on $\mathcal{H}$ is denoted as

$$P := \sum_{n=1}^{D} |n\rangle\langle n| \ , \tag{9}$$

the microcanonical density operator as

$$\rho_{\mathrm{mc}} := P/D = \frac{1}{D}\sum_{n=1}^{D} |n\rangle\langle n| \tag{10}$$

and the corresponding microcanonical expectation values as

$$\langle A\rangle_{\mathrm{mc}} := \mathrm{Tr}\{\rho_{\mathrm{mc}} A\} = \frac{1}{D}\mathrm{Tr}\{A\} \ . \tag{11}$$

II. DERIVATION OF EQ. (4) FROM THE MAIN PAPER

In this section, we provide the derivation of the relations

$$\overline{\sigma^2(t)} = \sum_{m\neq n}^{D} |\rho_{mn}(0)|^2\, |A_{mn}|^2 \leq \max_{m\neq n} |A_{mn}|^2 \ , \tag{12}$$

which are identical to Eq. (4) in the main paper. We remark that the same or very similar calculations are already contained, e.g., in [1–7], but they are repeated here to keep the paper self-contained.

Evaluating the trace in (5) by means of the energy basis $|n\rangle$ and exploiting (3) it follows that

$$\begin{aligned}\langle A\rangle_{\rho(t)} &= \sum_{m,n=1}^{D} \rho_{mn}(t) A_{nm} \\ &= \sum_{m,n=1}^{D} e^{i[E_n - E_m]t} \rho_{mn}(0) A_{nm} \ ,\end{aligned} \tag{13}$$

where $A_{mn} := \langle m|A|n\rangle$. Indicating time averages over arbitrary functions $f(t)$ by

$$\overline{f(t)} := \lim_{T\to\infty} \frac{1}{T}\int_0^T dt\, f(t) \tag{14}$$

and exploiting the assumption that $E_m - E_n \neq 0$ for all $m \neq n$ (see below Eq. (1)) it follows that

$$\overline{\langle A\rangle_{\rho(t)}} = \sum_{n=1}^{D} \rho_{nn}(0) A_{nn} = \mathrm{Tr}\{\bar\rho A\} = \langle A\rangle_{\bar\rho} \ , \tag{15}$$

where $\bar\rho := \overline{\rho(t)}$ and thus

$$\bar\rho_{mn} = \delta_{mn}\rho_{mn}(0) \ . \tag{16}$$

Along the same line of reasoning one finds that

$$\begin{aligned}[\langle A\rangle_{\rho(t)}]^2 &= \sum_{m,n=1}^{D} e^{i[E_n-E_m]t}\rho_{mn}(0)A_{nm} \\ &\times \sum_{j,k=1}^{D} e^{i[E_k-E_j]t}\rho_{jk}(0)A_{kj}\end{aligned}$$

$$\begin{aligned}&= \sum_{m\neq n}^{D}\rho_{mn}(0)A_{nm}\sum_{j,k=1}^{D} e^{i[E_n-E_m+E_k-E_j]t}\rho_{jk}(0)A_{kj} \\ &+\sum_{n=1}^{D}\rho_{nn}(0)A_{nn}\sum_{j,k=1}^{D} e^{i[E_k-E_j]t}\rho_{jk}(0)A_{kj} \ .\end{aligned} \tag{17}$$

Upon averaging over time, in the second summand only terms with $j = k$ survive since $E_j - E_k \neq 0$ for all $j \neq k$. Hence, this summand can be rewritten by means of (15) as $[\langle A\rangle_{\bar\rho}]^2$. Upon averaging over time in the first summand, only terms with $n = j$ and $m = k$ survive as a consequence of our above mentioned assumption (see below Eq. (1)) that the differences $E_m - E_n$ are finite and mutually different for all pairs (m, n) with $m \neq n$. With the definition

$$\sigma^2(t) := \left[\langle A\rangle_{\rho(t)} - \langle A\rangle_{\bar\rho}\right]^2 \ , \tag{18}$$

which is identical to Eq. (3) in the main paper, it follows that

$$\begin{aligned}\overline{\sigma^2(t)} &= \overline{\left[\langle A\rangle_{\rho(t)} - \langle A\rangle_{\bar\rho}\right]^2} = \overline{\left[\langle A\rangle_{\rho(t)}\right]^2} - [\langle A\rangle_{\bar\rho}]^2 \\ &= \sum_{m\neq n}^{D} |\rho_{mn}(0)|^2\,|A_{mn}|^2 \ .\end{aligned} \tag{19}$$

This is the first part of (12). The second part follows from

$$\sum_{m\neq n}^{D} |\rho_{mn}(0)|^2\,|A_{mn}|^2 \leq \max_{m\neq n}|A_{mn}|^2 \sum_{m\neq n}^{D}|\rho_{mn}(0)|^2 \tag{20}$$

in combination with

$$\begin{aligned}\sum_{m\neq n}^{D}|\rho_{mn}(0)|^2 &\leq \sum_{m,n=1}^{D}|\rho_{mn}(0)|^2 \\ &= \sum_{m,n=1}^{D}\langle m|\rho(0)|n\rangle\langle n|\rho(0)|m\rangle \\ &= \sum_{m}^{D}\langle m|\rho(0)\left(\sum_{n=1}^{D}|n\rangle\langle n|\right)\rho(0)|m\rangle \\ &= \sum_{m}^{D}\langle m|\rho^2(0)|m\rangle \\ &= \mathrm{Tr}\{\rho^2(0)\} \leq 1 \ .\end{aligned} \tag{21}$$

III. DERIVATION OF EQ. (5) FROM THE MAIN PAPER

In this section, we provide the derivation of the relation

$$\mu_U\left(\overline{\sigma^2(t)} \geq \epsilon\right) \leq 4\,\exp\left\{-\frac{\epsilon D}{18\pi^3\Delta_A^2} + 2\ln D\right\} \tag{22}$$

for any $\epsilon > 0$, which is identical to Eq. (5) in the main paper.

As in the main paper, $\mu_U(X)$ denotes the fraction (normalized measure) of all unitary transformations $U : \mathcal{H} \to \mathcal{H}$ (uniformly distributed according to the Haar measure [8–11]) which exhibit a certain property X. Moreover, let us denote by $\langle f(U)\rangle_U$ the average over all those U's, where $f(U)$ is an arbitrary, real-valued function of U. It follows that

$$\mu_U(f(U) \geq \epsilon) = \langle\theta(f(U) - \epsilon)\rangle_U \tag{23}$$

for any $\epsilon \in \mathbb{R}$, where $\theta(x) := \int_{-\infty}^{x}\delta(y)\mathrm{d}y$ is the Heaviside step function.

Next, we consider an arbitrary number of real-valued functions f_i of U (for convenience, arguments U are omitted, while the exact range of i is irrelevant and hence not specified). One readily verifies that

$$\theta(\max_i\{f_i\} - \epsilon) \leq \sum_i \theta(f_i - \epsilon) \tag{24}$$

for arbitrary f_i, $\epsilon \in \mathbb{R}$. By means of (23) and (24) one can infer that

$$\begin{aligned}&\mu_U(\max_i\{f_i\} \geq \epsilon) = \langle\theta(\max_i\{f_i\} - \epsilon)\rangle_U \\ &\leq \langle\sum_i \theta(f_i - \epsilon)\rangle_U = \sum_i\langle\theta(f_i - \epsilon)\rangle_U \\ &= \sum_i \mu_U(f_i \geq \epsilon)\end{aligned} \tag{25}$$

for arbitrary real-valued functions f_i of U and any $\epsilon \in \mathbb{R}$.

Similarly as in (24), one sees that

$$\theta(\sum_i f_i - \sum_i \epsilon_i) \leq \sum_i \theta(f_i - \epsilon_i) \tag{26}$$

for arbitrary f_i, $\epsilon_i \in \mathbb{R}$. Similarly as in (25), it follows that

$$\begin{aligned}&\mu_U(\sum_i f_i \geq \sum_i \epsilon_i) = \langle\theta(\sum_i f_i - \sum_i \epsilon_i)\rangle_U \\ &\leq \langle\sum_i \theta(f_i - \epsilon_i)\rangle_U = \sum_i\langle\theta(f_i - \epsilon_i)\rangle_U \\ &= \sum_i \mu_U(f_i \geq \epsilon_i) \ .\end{aligned} \tag{27}$$

Given an arbitrary but fixed pair $m \neq n$, we consider the four normalized vectors:

$$\begin{aligned}|\phi_1\rangle &:= (|m\rangle + |n\rangle)/\sqrt{2} \\ |\phi_2\rangle &:= (|m\rangle - |n\rangle)/\sqrt{2} \\ |\phi_3\rangle &:= (|m\rangle + i|n\rangle)/\sqrt{2} \\ |\phi_4\rangle &:= (|m\rangle - i|n\rangle)/\sqrt{2}\end{aligned} \tag{28}$$

With the abbreviation

$$A_{\phi\phi} := \langle\phi|A|\phi\rangle \ , \tag{29}$$

and exploiting that $\langle m|n\rangle = 0$, one readily verifies that

$$A_{\phi_1\phi_1} - A_{\phi_2\phi_2} - iA_{\phi_3\phi_3} + iA_{\phi_4\phi_4} = 2A_{mn} \tag{30}$$

and hence that

$$A_{mn} \ = \ (\Delta_1 - \Delta_2 - i\Delta_3 + i\Delta_4)/2 \tag{31}$$

$$\Delta_i \ := \ A_{\phi_i\phi_i} - \langle A\rangle_{\mathrm{mc}} \ , \ i = 1,\ldots,4 \tag{32}$$

where $\langle A\rangle_{\mathrm{mc}}$ is defined in (11). It follows that $|A_{mn}| \leq \sum_{i=1}^4 |\Delta_i|/2$ and thus

$$\mu_U(|A_{mn}| \geq \epsilon) \leq \mu_U\Big(\sum_{i=1}^4 |\Delta_i| \geq 2\,\epsilon\Big) \tag{33}$$

Choosing $\epsilon_i = \epsilon/2$ it follows with (27) that

$$\mu_U(|A_{mn}| \geq \epsilon) \leq \sum_{i=1}^4 \mu_U\Big(|\Delta_i| \geq \epsilon/2\Big) \ . \tag{34}$$

This is a quite interesting result in itself, admitting one to draw conclusions about off-diagonal matrix elements from related properties of the diagonal elements.

To quantitatively exploit this virtue of (34), we invoke the descendant of Levy's Lemma from Eq. (7) of the main paper (see also Ref. [12] and, in particular, Lemma 3, Lemma 5, and Eq. (64) in Ref. [13]), stating that

$$\mu_U\Big(|A_{\phi\phi} - \langle A\rangle_{\mathrm{mc}}| \geq \epsilon\Big) \leq 2\ \exp\left\{-\frac{2\,\epsilon^2 D}{9\pi^3\Delta_A^2}\right\} \tag{35}$$

for any normalized $|\phi\rangle := \sum_{n=1}^D c_n|n\rangle$. Recalling the definitions (28) and (32) it follows from (34) and (35) that

$$\mu_U(|A_{mn}| \geq \epsilon) \leq 8\ \exp\left\{-\frac{\epsilon^2 D}{18\pi^3\Delta_A^2}\right\} \ . \tag{36}$$

In view of (25) we thus can conclude that

$$\mu_U\left(\max_{m\neq n}|A_{mn}| \geq \epsilon\right) \leq \frac{D(D-1)}{2} 8\ \exp\left\{-\frac{\epsilon^2 D}{18\pi^3\Delta_A^2}\right\} \ . \tag{37}$$

The factor $D(D-1)/2$ is due to the fact that $|A_{nm}| = |A_{mn}|$, i.e., there are only $D(D-1)/2$ independent pairs (m,n) over which one has to maximize. Upon rewriting (37) as

$$\mu_U\left(\max_{m\neq n}|A_{mn}|^2 \geq \epsilon\right) \leq 4\ \exp\left\{-\frac{\epsilon D}{18\pi^3\Delta_A^2} + 2\ln D\right\} \ . \tag{38}$$

and observing (12), we recover (22).

IV. DERIVATION OF EQ. (10) FROM THE MAIN PAPER

In this section, we provide the derivation of the relation

$$\mu_U\left(|B| \geq \epsilon\right) \leq 2\ \exp\left\{-\frac{2}{9\pi^3}\frac{\epsilon^2 D}{\Delta_A^2} + \ln D\right\} \tag{39}$$

for any $\epsilon > 0$, which is identical to Eq. (10) in the main paper.

Since all necessary steps have already been prepared in the last section, the argument is very short: Choosing $|\phi\rangle = |n\rangle$ in the descendant (35) of Levy's Lemma and exploiting (25) implies

$$\mu_U\Big(\max_n|A_{nn} - \langle A\rangle_{\mathrm{mc}}| \geq \epsilon\Big) \leq 2\,D\ \exp\left\{-\frac{2\,\epsilon^2 D}{9\pi^3\Delta_A^2}\right\} \tag{40}$$

In combination with the relation

$$|B| \leq \max_n|A_{nn} - \langle A\rangle_{\mathrm{mc}}| \ , \tag{41}$$

which is identical to Eq. (9) in the main paper, we thus recover (39).

V. COMPARISON WITH THE ESTIMATES BY VON NEUMANN AND BY PAULI AND FIERZ

In this section, we compare von Neumann's main estimates, obtained in the Appendix of his work [8], and their improvements by Pauli and Fierz in Ref. [14], with Levy's Lemma and its descendants, utilized as a basic ingredient in the previous two sections.

Since von Neumann's entire approach is formulated in terms of projection operators, let us denote by P_d an arbitrary projector onto a d-dimensional subspace of $\mathcal{H}$. Referring to the equation numbers in the English translation by R. Tumulka [8], we first note that the concomitant matrix elements $P_{mn}^{(d)} := \langle m|P_d|n\rangle$ correspond to von Neumann's $e_{\rho\sigma}$ with $\rho := m$, $\sigma := n$, see Eq. (142) in [8]. Likewise, our present D and d are named S and s in the Appendix of [8]. Recalling that $\mu_U(X)$ indicates the fraction of all U with property X (see below (22)), relation (192) in [8] takes the form

$$\mu_U\left(|e_{\rho\sigma}|^2 \geq a\right) \leq \exp\{-4a(S-5/2)\} \tag{42}$$

for any $a > 0$ and $\rho \neq \sigma$. In particular, this result is independent of s. Returning to our present notation, we thus obtain

$$\mu_U\left(|P_{mn}^{(d)}| \geq \epsilon\right) \leq \exp\{-4\epsilon^2(D-5/2)\} \tag{43}$$

for any $\epsilon > 0$, $m \neq n$, and independently of d. Upon observing that $\Delta_A = 1$ when A is a projector (cf. Eq. (8)), we see that (36) is indeed very similar to (43).

In the same vein, von Neumann's result for the diagonal matrix elements $P_{nn}^{(d)}$ from Eq. (162) in [8] (note that actually two summands of the form (162) contribute to the total probability in Eq. (152) of [8]) can be rewritten in our present notation as

$$\mu_U\left(|P^{(d)}_{nn} - d/D| \geq \epsilon\right) \leq \frac{2\,D}{e\sqrt{2\pi d}}\,\exp\left\{-\frac{\vartheta\epsilon^2 D^2}{2d}\right\}\ , \tag{44}$$

provided $1 \ll d \ll D$, where ϑ is some number slightly smaller than one, and where ϵ must satisfy $D^{-1} \leq \epsilon \ll d/D$. Observing that $d/D = \langle P_d\rangle_{\rm mc}$ (cf. Eq. (11)) and that $\Delta_A = 1$ when A is a projector (cf. Eq. (8)), we see that von Neumann's bound (44) is somewhat better than the descendant (35) of Levy's Lemma. On the other hand, (44) is restricted to projectors P_d with $1 \ll d \ll D$, while (35) applies to arbitrary A.

Without any further calculation, we can immediately deduce from (44) by exploiting (34) the bound

$$\mu_U\left(|P^{(d)}_{mn}| \geq \epsilon\right) \leq \frac{8\,D}{e\sqrt{2\pi d}}\,\exp\left\{-\frac{\vartheta\epsilon^2 D^2}{8d}\right\}\ . \tag{45}$$

The comparison with (36) and with (43) is analogous to the discussion below (44).

Finally, we turn to the improvement of von Neumann's estimate (44) by Pauli and Fierz in Appendix 2 of Ref. [14]. We first note that our present quantity $P^{(d)}_{nn}$ corresponds to the quantity $C^\nu_{\rho\rho}$ with $\rho := n$ in [14], see Eq. (11) therein (the index ν labels different projectors and is of no relevance here). Likewise, our present D and d are named S and s_ν in [14], and s_ν/S is abbreviated as g_ν (see p. 579 in [14]). Recalling that $\mu_U(X)$ indicates the fraction of all U with property X, relation (28) in [14] takes the form

$$\begin{aligned}&\mu_U\left(C^\nu_{\rho\rho} - g_\nu \geq \sqrt{g_\nu a}\right) \leq \exp\{-\kappa\sqrt{aS} + \kappa + \ln S\}\\ &\kappa := 1 - \ln 2 \simeq 0.3068\end{aligned} \tag{46}$$

for all $a > 2/(S-2)$ and $S \geq 3$, and likewise

$$\mu_U\left(C^\nu_{\rho\rho} - g_\nu \leq -\sqrt{g_\nu a}\right)\ \leq\ \exp\{-\kappa\sqrt{aS} + \kappa + \ln S\}\ .$$

It follows that

$$\begin{aligned}&\mu_U\left([C^\nu_{\rho\rho} - g_\nu]^2 \leq g_\nu a\right) \leq 2\exp\{-\kappa\sqrt{aS} + \kappa + \ln S\}\\ &= \exp\{-\kappa\sqrt{aS} + 1 + \ln S\}\ ,\end{aligned} \tag{47}$$

for all $a > 2/(S-2)$ and $S \geq 3$, where we exploited (46) in the last relation. Returning to our present notation, we thus obtain

$$\mu_U\left(|P^{(d)}_{nn} - d/D| \geq \epsilon\right) \leq e\,D\,\exp\left\{-\frac{\kappa\epsilon D}{\sqrt{d}}\right\}\ , \tag{48}$$

for all $\epsilon > \sqrt{2/(D-2)}$ and $D \geq 3$. While this estimate is in fact somewhat weaker than (44), it represents a substantial improvement in so far as there is no restriction with respect to d. Similar conclusions apply to the estimate following from (34) and (48), namely

$$\mu_U\left(|P^{(d)}_{mn}| \geq \epsilon\right) \leq 4\,e\,D\,\exp\left\{-\frac{\kappa\epsilon D}{2\sqrt{d}}\right\}\ . \tag{49}$$

We finally remark that it is possible to deduce directly from von Neumann's estimate for projectors (43) a relation similar to the one in (22) for general observables A. Likewise, one can deduce from the estimate by Pauli and Fierz (48) a relation similar to (39). The main idea is to "truncate" (or round) the eigenvalues a_ν of A in (4) after, say, 20 relevant digits (cf. Eq. (2) in the main paper). On the one hand, this does not measurably change the expectation value $\langle A\rangle_\rho$ for any arbitrary but fixed ρ. On the other hand, this approximation for A then only exhibits a number of mutually different eigenvalues which is (approximately) bounded by $\Delta_A/\delta A$ according to Eq. (2) of the main paper, and which is thus relatively small compared to the dimensionality D of $\mathcal{H}$. Consequently, one is left with a relatively small number of pertinent projectors, to each of which the above estimates by von Neumann and by Pauli and Fierz can be applied (at this point it is important that those estimates are valid for arbitrary d). However, the technical details are quite lengthy and therefore omitted here.

In conclusion, quite similar results can be obtained either by invoking Levy's Lemma or by exploiting the estimates by von Neumann and by Pauli and Fierz, but the calculations are considerably shorter in the first case.

VI. DERIVATION OF EQ. (11) FROM THE MAIN PAPER

In this section, we provide the derivation of the relation

$$\mu_V\left(S[\rho_{\rm mc}] - S[\bar\rho] \geq s\right) \leq k_B/s \tag{50}$$

for any $s > 0$, which is identical to Eq. (11) in the main paper.

We start by recalling the definition

$$S[\rho] := -k_B{\rm Tr}\{\rho\ln\rho\} \tag{51}$$

of the von Neumann entropy of an arbitrary density operator ρ, where k_B is Boltzmann's constant. Furthermore, we can rewrite (16) as $\bar\rho_{mn} = \delta_{mn}\,p_n$, where the level populations

$$p_n := \rho_{nn}(0) \tag{52}$$

satisfy $0 \leq p_n \leq 1$ and $\sum_{n=1}^D p_n = 1$. It readily follows that

$$S[\bar\rho]\ =\ -k_B\sum_{n=1}^D p_n\ln p_n \tag{53}$$

$$\sum_{n=1}^D p_n^l\ =\ {\rm Tr}\{\bar\rho^l\} \tag{54}$$

for any $l \in \mathbb{N}$.

Introducing the microcanonical density operator from (10) into (51) yields

$$S[\rho_{\rm mc}] = -k_B\sum_{n=1}^D(1/D)\ln(1/D) = k_B\ln D\ . \tag{55}$$

It follows with (53) that

$$\Delta S := \frac{S[\rho_{\rm mc}] - S[\bar{\rho}]}{k_B} = \sum_{n=1}^{D} p_n \left(\ln p_n - \ln(1/D)\right) \ . \quad (56)$$

Since $\sum_{n=1}^{D}(-p_n + 1/D) = 0$ we can conclude that [14]

$$\Delta S = \sum_{n=1}^{D} L(p_n, 1/D) \quad (57)$$

$$L(x,y) := x\left(\ln x - \ln y\right) - x + y \ , \quad (58)$$

where $x, y > 0$ is tacitly assumed. Rewriting $L(x,y)$ as $yf(x/y)$ with $f(x) := x\ln x - x + 1$, one readily verifies that $f(x) \geq 0$ for all $x > 0$ and hence that $L(x,y) \geq 0$ for all $x, y > 0$. Furthermore, one can infer from (58) that $yL(x,y) + xL(y,x) = (x-y)^2$. Since $xL(y,x) \geq 0$ it follows that

$$0 \leq L(x,y) \leq (x-y)^2/y \quad (59)$$

for all $x, y > 0$ [14]. Introducing this result into (57) yields

$$0 \leq \Delta S \leq \sum_{n=1}^{D} D(p_n - 1/D)^2$$

$$= \left(D \sum_{n=1}^{D} p_n^2\right) - 1 = D\mathrm{Tr}\{\bar{\rho}^2\} - 1 \quad (60)$$

where we used (54) in the last step.

Denoting by r_j and $|\psi_j\rangle$ the eigenvalues and eigenvectors of the density operator $\rho(0)$ implies that

$$\rho(0) = \sum_{j=1}^{D} r_j |\psi_j\rangle\langle\psi_j| \quad (61)$$

$$\sum_{j=1}^{D} r_j^l = \mathrm{Tr}\{\rho^l(0)\} \quad (62)$$

for any $l \in \mathbb{N}$.

Recalling that V represents the unitary basis transformation between the eigenvectors of $\rho(0)$ and those of H (see main paper) and that those eigenvectors are denoted as $|\psi_j\rangle$ and $|n\rangle$, the transformation matrix elements are given by

$$V_{nj} := \langle n|\psi_j\rangle \ . \quad (63)$$

Furthermore, (52) can be rewritten as

$$p_n = \langle n|\rho(0)|n\rangle = \sum_{j=1}^{D} \langle n|\rho(0)|\psi_j\rangle\langle\psi_j|n\rangle$$

$$= \sum_{j=1}^{D} r_j \langle n|\psi_j\rangle(\langle n|\psi_j\rangle)^* = \sum_{j=1}^{D} r_j |V_{nj}|^2 \quad (64)$$

and hence

$$p_n^2 = \sum_{j,k=1}^{D} r_j r_k |V_{nj} V_{nk}|^2 \ . \quad (65)$$

As in the main paper, $\mu_V(X)$ indicates the fraction (normalized measure) of all unitary transformations $V : \mathcal{H} \to \mathcal{H}$ (uniformly distributed according to the Haar measure [8–11]) which exhibit a certain property X. Moreover, let us denote by $\langle f(V)\rangle_V$ the average over all those V's, where $f(V)$ is an arbitrary real-valued function of V. Similarly as in (23), it follows that

$$\mu_V(|f(V)| \geq \epsilon) = \langle \theta(|f(V)| - \epsilon)\rangle_V \quad (66)$$

for any $\epsilon > 0$, where $\theta(x) := \int_{-\infty}^{x} \delta(y) \mathrm{d}y$ is the Heaviside step function. Hence,

$$\langle \theta(|f(V)| - \epsilon)\rangle_V \leq \langle |f(V)/\epsilon|^a \theta(|f(V)| - \epsilon)\rangle_V$$

$$\leq \langle |f(V)/\epsilon|^a\rangle_V \quad (67)$$

for any $a > 0$. By comparison with (66) we thus recover Markov's inequality

$$\mu_V(|f(V)| \geq \epsilon) \leq \epsilon^{-a}\langle |f(V)|^a\rangle_V \quad (68)$$

for any real valued function $f(V)$ and arbitrary ϵ, $a > 0$.

In the following, we will prove (50) with the help of (68). To begin with, we average (65) over V, yielding

$$\langle p_n^2\rangle_V = \sum_{j,k=1}^{D} r_j r_k \langle |V_{nj} V_{nk}|^2\rangle_V \ . \quad (69)$$

In doing so, we have exploited that only the eigenbases of H and $\rho(0)$ change relatively to each other upon variation of V, while the eigenvalues r_j are kept fixed.

Averages as those on the right hand side of (69) have been evaluated repeatedly and often independently of each other in the literature, see e.g. [15–20]. Specifically, the two results (A6) and (A8) from the Appendix in Ref. [15] (or Eq. (2.4) from [19]) can be readily unified and rewritten as

$$\langle |V_{nj} V_{nk}|^2\rangle_V = \frac{1 + \delta_{jk}}{D(D+1)} \ . \quad (70)$$

Introducing this result into (69) and exploiting (62) yields

$$\langle p_n^2\rangle_V = \frac{\left(\sum_{j=1}^{D} r_j\right)^2 + \sum_{j=1}^{D} r_j^2}{D(D+1)} =$$

$$= \frac{(\mathrm{Tr}\{\rho(0)\})^2 + \mathrm{Tr}\{\rho^2(0)\}}{D(D+1)} \leq \frac{2}{D(D+1)} \quad (71)$$

Upon averaging over V in (60) we thus can conclude that

$$\langle \Delta S\rangle_V \leq \left(D \sum_{n=1}^{D} \frac{2}{D(D+1)}\right) - 1 = \frac{D-1}{D+1} \leq 1 \quad (72)$$

Finally, by choosing $f(V) := \Delta S$, $\epsilon := s/k_B$, and $a := 1$ in (68) we can conclude from (56) and (72) that

$$\mu_V\left(|S[\rho_{\rm mc}] - S[\bar{\rho}]| \geq s\right) \leq k_B/s \tag{73}$$

From (60) and (56) we recover the well-known relation $S[\rho_{\rm mc}] \geq S[\rho]$ for all ρ. As a consequence, (73) implies (50).

VII. DERIVATION OF EQ. (12) FROM THE MAIN PAPER

In this section, we first demonstrate the relation

$$\overline{\sigma^2(t)} \;\leq\; (\Delta_A^2/4)\,\mathrm{Tr}\{\bar{\rho}^2\}\;, \tag{74}$$

mentioned in the main paper above Eq. (12) therein. In a second step, we then provide the derivation of the result

$$\mu_V\left(\overline{\sigma^2(t)} \geq \epsilon\right) \leq \Delta_A^2/(2\epsilon D) \tag{75}$$

for any $\epsilon > 0$, which is identical to Eq. (12) in the main paper. We remark that (74) has already been obtained, e.g., in [4–7], but to make the paper self-contained, we derive here the relation once again.

We define $\tilde{A} := A - \alpha P$ for any $\alpha \in \mathbb{R}$, where P is the identity operator from (9). It follows that $\tilde{A}_{mn} = A_{mn}$ for all $m \neq n$ and with (12) that

$$\begin{aligned}\overline{\sigma^2(t)} \;&=\; \sum_{m\neq n}^{D} |\rho_{mn}(0)|^2\, |\tilde{A}_{mn}|^2 \\ &\leq\; \sum_{m,n=1}^{D} |\rho_{mn}(0)|^2\, |\tilde{A}_{mn}|^2\;. \end{aligned}\tag{76}$$

It furthermore follows that the eigenvalues $\tilde{a}_\nu$ of $\tilde{A}$ and the eigenvalues a_ν of A are connected via $\tilde{a}_\nu = a_\nu - \alpha$ for all $\nu = 1, \ldots, D$. Choosing

$$\alpha := (a_{max} + a_{min})/2\;, \tag{77}$$

where a_{max} and a_{min} are defined in (6), (7), we can conclude from (8) that

$$|a_\nu - \alpha| \leq \Delta_A/2 \tag{78}$$

for all ν. It follows that

$$\langle\psi|\tilde{A}^2|\psi\rangle \leq \Delta_A^2/4 \tag{79}$$

for any normalized vector $|\psi\rangle \in \mathcal{H}$.

Since $\rho(0)$ is a non-negative operator it follows by Cauchy-Schwarz's inequality that $|\rho_{mn}(0)|^2 \leq \rho_{mm}(0)\rho_{nn}(0) = p_n p_m$, where we used (52) in the last relation. We thus can rewrite (76) as

$$\begin{aligned}\overline{\sigma^2(t)} \;&\leq\; \sum_{m,n=1}^{D} p_m p_n |\tilde{A}_{mn}|^2 \\ &\leq\; \sum_{m,n=1}^{D} b_{mn} c_{mn} \end{aligned}\tag{80}$$

$$b_{mn} \;:=\; p_n\,|\tilde{A}_{mn}| \tag{81}$$

$$c_{mn} \;:=\; p_m\,|\tilde{A}_{mn}| \tag{82}$$

Invoking Cauchy-Schwarz's inequality once more it follows that

$$\begin{aligned}\overline{\sigma^2(t)} \;&\leq\; \sqrt{\sum_{m,n=1}^{D} |b_{mn}|^2 \sum_{m,n=1}^{D} |c_{mn}|^2} \\ &=\; \sum_{m,n=1}^{D} |b_{mn}|^2 \\ &=\; \sum_{n=1}^{D} p_n^2\,\langle n|\tilde{A}\left(\sum_{m=1}^{D} |m\rangle\langle m|\right)\tilde{A}|n\rangle \\ &=\; \sum_{n=1}^{D} p_n^2\,\langle n|\tilde{A}^2|n\rangle \\ &\leq\; \frac{\Delta_A^2}{4}\sum_{n=1}^{D} p_n^2\;, \end{aligned}\tag{83}$$

where we used (79) in the last step. Exploiting (54), one readily recovers (74).

With (71) one can conclude from (83) that

$$\langle\overline{\sigma^2(t)}\rangle_V \;\leq\; \frac{\Delta_A^2}{4}\sum_{n=1}^{D}\frac{2}{D(D+1)} \leq \frac{\Delta_A^2}{2D} \tag{84}$$

Choosing $f(V) := \overline{\sigma^2(t)}$ and $a := 1$ in (68) we finally recover (75).

VIII. DERIVATION OF EQS. (14) AND (15) FROM THE MAIN PAPER

This section provides the derivation of the relations (14) and (15) in the main paper, namely

$$\mu_W\left(\sigma_{\rm mc}^2(t) \geq \epsilon\right) \leq \Delta_A^2/(\epsilon D) \tag{85}$$

for arbitrary t and $\epsilon > 0$, and

$$\mu_W\left(\frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \sigma_{\rm mc}^2(t)\,\mathrm{d}t \geq \epsilon\right) \leq \Delta_A^2/(\epsilon D) \tag{86}$$

for arbitrary $t_1 < t_2$.

We consider an arbitrary but fixed density operator ρ and denote its eigenvalues and eigenvectors as r_n and

$|\psi_n\rangle$. Similarly as in (61), (62) it follows that

$$\rho = \sum_{n=1}^{D} r_n |\psi_n\rangle\langle\psi_n| \tag{87}$$

$$\sum_{n=1}^{D} r_n^l = \mathrm{Tr}\{\rho^l(t)\} \tag{88}$$

for any $l \in \mathbb{N}$, and likewise for A from (4):

$$\sum_{\nu=1}^{D} a_\nu^l = \mathrm{Tr}\{A^l\} \tag{89}$$

We denote by $\tilde{W}$ the unitary basis transformation between the eigenvectors of ρ and those of A with matrix elements

$$\tilde{W}_{n\nu} := \langle\psi_n|\nu\rangle \ . \tag{90}$$

It follows with (87) that

$$\begin{aligned}\mathrm{Tr}\{\rho A\} &= \sum_{n=1}^{D} \langle\psi_n|\rho A|\psi_n\rangle \\ &= \sum_{n=1}^{D} r_n \langle\psi_n|A|\psi_n\rangle\end{aligned} \tag{91}$$

and hence with (4) and (90) that

$$\begin{aligned}\mathrm{Tr}\{\rho A\} &= \sum_{n,\nu=1}^{D} r_n \langle\psi_n|\nu\rangle\langle\nu|A|\psi_n\rangle \\ &= \sum_{n,\nu=1}^{D} r_n a_\nu \tilde{W}_{n\nu}\tilde{W}_{n\nu}^* \ .\end{aligned} \tag{92}$$

Similarly as in (69) this yields upon averaging over $\tilde{W}$ (uniformly distributed according to the Haar measure [8–11]) the result

$$\langle \mathrm{Tr}\{\rho A\}\rangle_{\tilde{W}} = \sum_{n,\nu=1}^{D} r_n a_\nu \langle|\tilde{W}_{n\nu}|^2\rangle_{\tilde{W}} \ . \tag{93}$$

Likewise, squaring (92) and then averaging over $\tilde{W}$ yields

$$\langle[\mathrm{Tr}\{\rho A\}]^2\rangle_{\tilde{W}} = \sum_{m,n,\mu,\nu=1}^{D} r_m r_n a_\mu a_\nu \langle|\tilde{W}_{m\mu}\tilde{W}_{n\nu}|^2\rangle_{\tilde{W}} \ . \tag{94}$$

As already mentioned above (70), averages as those appearing on the right hand side of (93) and (94) are well-known. For example, formulae (A5), (A6), (A8), (A10), (A12) in the Appendix of Ref. [15] can be readily recast into the form

$$\langle|\tilde{W}_{n\nu}|^2\rangle_{\tilde{W}} = \frac{1}{D} \tag{95}$$

$$\langle|\tilde{W}_{n\mu}\tilde{W}_{n\nu}|^2\rangle_{\tilde{W}} = \frac{1+\delta_{\mu\nu}}{D(D+1)} \tag{96}$$

$$\langle|\tilde{W}_{m\mu}\tilde{W}_{n\nu}|^2\rangle_{\tilde{W}} = \frac{1-\delta_{\mu\nu}/D}{D^2-1} \quad \text{if} \quad m \neq n \ . \tag{97}$$

By exploiting (95), (88), and (89), one finds from (93) that

$$\langle \mathrm{Tr}\{\rho A\}\rangle_{\tilde{W}} = \frac{1}{D}\sum_{n,\nu=1}^{D} r_n a_\nu = \frac{1}{D}\mathrm{Tr}\{\rho\}\mathrm{Tr}\{A\} \ . \tag{98}$$

The evaluation of (94) is similar but more tedious. We start by splitting the sum into contributions with $m = n$ and with $m \neq n$,

$$\langle[\mathrm{Tr}\{\rho A\}]^2\rangle_{\tilde{W}} = \Sigma_1 + \Sigma_2 \tag{99}$$

$$\Sigma_1 := \sum_{n=1}^{D}\sum_{\mu,\nu=1}^{D} r_n^2 a_\mu a_\nu \langle|\tilde{W}_{n\mu}\tilde{W}_{n\nu}|^2\rangle_{\tilde{W}} \tag{100}$$

$$\Sigma_2 := \sum_{m\neq n}^{D}\sum_{\mu,\nu=1}^{D} r_m r_n a_\mu a_\nu \langle|\tilde{W}_{m\mu}\tilde{W}_{n\nu}|^2\rangle_{\tilde{W}} \tag{101}$$

Introducing (96) into (100) yields

$$\begin{aligned}\Sigma_1 &= \sum_{n=1}^{D} r_n^2 \frac{\sum_{\mu,\nu=1}^{D} a_\mu a_\nu + \sum_{\nu=1}^{D} a_\nu^2}{D(D+1)} \\ &= \mathrm{Tr}\{\rho^2\}\,\frac{[\mathrm{Tr}\{A\}]^2 + \mathrm{Tr}\{A^2\}}{D(D+1)} \ ,\end{aligned} \tag{102}$$

where we exploited (88) and (89) in the last step. Analogously, one finds with (97) that

$$\begin{aligned}\Sigma_2 &= \sum_{m\neq n}^{D} r_m r_n \frac{\sum_{\mu,\nu=1}^{D} a_\mu a_\nu - \sum_{\nu=1}^{D} a_\nu^2/D}{D^2-1} \\ &= \left([\mathrm{Tr}\{\rho\}]^2 - \mathrm{Tr}\{\rho^2\}\right)\frac{[\mathrm{Tr}\{A\}]^2 - \mathrm{Tr}\{A^2\}/D}{D^2-1}\end{aligned} \tag{103}$$

Finally, a straightforward but somewhat lengthy calculation yields

$$\begin{aligned}&\langle[\mathrm{Tr}\{\rho A\}]^2\rangle_{\tilde{W}} - [\langle \mathrm{Tr}\{\rho A\}\rangle_{\tilde{W}}]^2 = \\ &= \frac{\left(\mathrm{Tr}\{\rho^2\} - [\mathrm{Tr}\{\rho\}]^2/D\right)\left(\mathrm{Tr}\{A^2\} - [\mathrm{Tr}\{A\}]^2/D\right)}{D^2-1} \ .\end{aligned} \tag{104}$$

This result is still symmetric with respect to interchanging ρ and A since we only exploited so far the fact that ρ is a Hermitian operator but not yet any of its special properties as a density operator.

By exploiting (10) we can conclude from (98) that

$$\langle \mathrm{Tr}\{\rho A\}\rangle_{\tilde{W}} = \mathrm{Tr}\{\rho_{\mathrm{mc}} A\} = \langle \mathrm{Tr}\{\rho_{\mathrm{mc}} A\}\rangle_{\tilde{W}} \tag{105}$$

and hence that

$$\begin{aligned}&\langle[\mathrm{Tr}\{\rho A\}]^2\rangle_{\tilde{W}} - [\langle \mathrm{Tr}\{\rho A\}\rangle_{\tilde{W}}]^2 = \\ &= \langle[\mathrm{Tr}\{\rho A\} - \mathrm{Tr}\{\rho_{\mathrm{mc}} A\}]^2\rangle_{\tilde{W}} \ .\end{aligned} \tag{106}$$

Furthermore, the last term in (104) can be rewritten as

$$\mathrm{Tr}\{A^2\} - [\mathrm{Tr}\{A\}]^2/D = D\,[\langle A^2\rangle_{\mathrm{mc}} - \langle A\rangle_{\mathrm{mc}}^2] \ . \tag{107}$$

Likewise, one finds that

$$\begin{aligned}\mathrm{Tr}\{\rho^2\} - [\mathrm{Tr}\{\rho\}]^2/D &= \mathrm{Tr}\{\rho^2\} - 1/D \\ &= \mathrm{Tr}\{(\rho - \rho_{\mathrm{mc}})^2\} \ . \end{aligned} \quad (108)$$

Hence, this term is non-negative, vanishes if and only if $\rho = \rho_{\mathrm{mc}}$, and can be estimated from above for arbitrary ρ by $1 - 1/D = (D-1)/D$. Altogether, we thus can conclude from (104) that

$$\begin{aligned}\langle[\mathrm{Tr}\{\rho A\} - \mathrm{Tr}\{\rho_{\mathrm{mc}} A\}]^2\rangle_{\tilde{W}} &\leq \frac{\langle A^2\rangle_{\mathrm{mc}} - \langle A\rangle^2_{\mathrm{mc}}}{D} \\ = \frac{\langle[A - \langle A\rangle_{\mathrm{mc}}]^2\rangle_{\mathrm{mc}}}{D} \ . \end{aligned} \quad (109)$$

The numerator on the right hand side represents the thermal equilibrium fluctuations of A in the microcanonical ensemble. In particular, this term is non-negative and can be estimated from above by Δ_A^2 according to (8), i.e.

$$\langle[\mathrm{Tr}\{\rho A\} - \mathrm{Tr}\{\rho_{\mathrm{mc}} A\}]^2\rangle_{\tilde{W}} \leq \Delta_A^2/D \ . \quad (110)$$

So far, ρ was still an arbitrary density operator. Next, we choose $\rho = \rho(t)$ for an arbitrary but fixed t. With the definition

$$\sigma^2_{\mathrm{mc}}(t) := \left[\langle A\rangle_{\rho(t)} - \langle A\rangle_{\mathrm{mc}}\right]^2 \ , \quad (111)$$

which is identical to Eq. (13) in the main paper, we thus can rewrite (110) as

$$\langle\sigma^2_{\mathrm{mc}}(t)\rangle_{\tilde{W}} \leq \Delta_A^2/D \ . \quad (112)$$

As in the main paper, we denote by W the unitary basis transformation between the eigenvectors of $\rho(0)$ and those of A. On the other hand, $\tilde{W}$ was defined as the basis transformation between the eigenvectors of $\rho = \rho(t)$ and those of A. The connection between $\rho(t)$ and $\rho(0)$ is provided by (2), where the propagator $\mathcal{U}_t : \mathcal{H} \to \mathcal{H}$ is a unitary operator for any given t. In fact, the same basic relation (2) is well-known to apply even for time dependent Hamiltonians $H(t) : \mathcal{H} \to \mathcal{H}$. The salient point is that $\mathcal{U}_t$ unitarily transforms the basis of $\rho(0)$ into the basis of $\rho(t)$, and that for a given t, the very same transformation $\mathcal{U}_t$ applies to any $\rho(0)$, in particular independently of the basis of $\rho(0)$. As a consequence, averaging uniformly (i.e. according to the Haar measure) over all bases of $\rho(0)$ is equivalent [11] to averaging uniformly over all bases of $\rho(t)$ in (112), i.e.

$$\langle\sigma^2_{\mathrm{mc}}(t)\rangle_W \leq \Delta_A^2/D \ . \quad (113)$$

Since t was arbitrary but fixed, the latter result is valid for any choice of t.

Exploiting that averaging and integrating are commuting operations, we can infer from (113) that

$$\begin{aligned}&\left\langle \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \sigma^2_{\mathrm{mc}}(t)\,\mathrm{d}t \right\rangle_W = \\ &= \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \langle\sigma^2_{\mathrm{mc}}(t)\rangle_W\,\mathrm{d}t \leq \Delta_A^2/D \end{aligned} \quad (114)$$

for arbitrary $t_1 < t_2$.

Denoting by $\mu_W(X)$, as usual, the fraction (normalized measure) of all unitary transformations $W : \mathcal{H} \to \mathcal{H}$ which exhibit a certain property X, and by $\langle f(W)\rangle_W$ the average over all those W's for an arbitrary real valued function $f(W)$, one recovers, similarly as in (66)-(68), Markov's inequality

$$\mu_V(|f(W)| \geq \epsilon) \leq \epsilon^{-a}\langle|f(W)|^a\rangle_W \quad (115)$$

for any ϵ, $a > 0$. Upon choosing $f(W) := \sigma^2_{\mathrm{mc}}(t)$ and $a := 1$, Eq. (113) implies (85). Similarly, Eq. (114) implies (86).

We finally note that (75) actually includes (74) as special cases when $t_1 \uparrow t$ and $t_2 \downarrow t$.

---

[1] A. Peres, Phys. Rev. A **30**, 504 (1984)
[2] H. Tasaki, Phys. Rev. Lett. **80**, 1373 (1998)
[3] M. Srednicki, J. Phys. A: Math. Gen **32**, 1163 (1999)
[4] P. Reimann, Phys. Rev. Lett. **101**, 190403 (2008)
[5] N. Linden, S. Popescu, A. J. Short, and A. Winter, Phys. Rev. E **79**, 061103 (2009)
[6] P. Reimann, New J. Phys. **12**, 055027 (2010)
[7] A. J. Short, New J. Phys. **13**, 053009 (2011)
[8] J. von Neumann, Z. Phys. **57**, 30 (1929); English translation by R. Tumulka, Eur. J. Phys. H **35**, 201 (2010)
[9] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Eur. J. Phys. H **35**, 173 (2010)
[10] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, Phys. Rev. E **81**, 011109 (2010)
[11] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, Proc. R. Soc. London Ser. A **466**, 3203 (2010)
[12] S. Popescu, A. J. Short, and A. Winter, Nature Phys. **2**, 754 (2006)
[13] S. Popescu, A. J. Short, and A. Winter, `arXiv:quant-ph/0511225v3` (2006)
[14] W. Pauli and M. Fierz, Z. Phys. **106**, 572 (1937)
[15] I. E. Farquhar and P. T. Landsberg, Proc. Royal Soc. London **239**, 134 (1957)
[16] P. Bocchieri and A. Loinger, Phys. Rev. **111**, 668 (1958)
[17] P. Bocchieri and A. Loinger, Phys. Rev. **114**, 948 (1959)
[18] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong, Rev. Mod. Phys. **53**, 385 (1981), Sect. VIII.B.
[19] P. Pechukas, J. Math. Phys. **25**, 534 (1984)
[20] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics* (2nd edition, Springer, Berlin, Heidelberg, 2009), Appendix A.