J Stat Phys (2008) 132: 921–935
DOI 10.1007/s10955-008-9576-1

Typicality of Pure States Randomly Sampled According to the Gaussian Adjusted Projected Measure

Peter Reimann

Received: 18 January 2008 / Accepted: 29 May 2008 / Published online: 25 June 2008
© Springer Science+Business Media, LLC 2008

Abstract Consider a mixed quantum mechanical state, describing a statistical ensemble in terms of an arbitrary density operator ρ of low purity, $\mathrm{tr}\,\rho^2 \ll 1$, and yielding the ensemble averaged expectation value $\mathrm{tr}(\rho A)$ for any observable A. Assuming that the given statistical ensemble ρ is generated by randomly sampling pure states $|\psi\rangle$ according to the corresponding so-called Gaussian adjusted projected measure (Goldstein et al. in J. Stat. Phys. 125:1197, 2006), the expectation value $\langle\psi|A|\psi\rangle$ is shown to be extremely close to the ensemble average $\mathrm{tr}(\rho A)$ for the overwhelming majority of pure states $|\psi\rangle$ and any experimentally realistic observable A. In particular, such a 'typicality' property holds whenever the Hilbert space $\mathcal{H}$ of the system contains a high dimensional subspace $\mathcal{H}_+ \subset \mathcal{H}$ with the property that all $|\psi\rangle \in \mathcal{H}_+$ are realized with equal probability and all other $|\psi\rangle \in \mathcal{H}$ are excluded.

Keywords Quantum statistical mechanics · Quantum ensemble theory · Gaussian measures · Microcanonical density matrices

1 Introduction

Spheres in high dimensional Euclidean spaces exhibit astonishing geometrical properties, as discussed in detail e.g. in basic Statistical Physics lectures: two randomly drawn vectors, each connecting the center of the sphere with any point at its surface, are practically orthogonal with extremely high probability; almost the entire volume of the sphere is contained within an extremely thin surface layer of the sphere; the latter in turn exhibits an extreme concentration of its volume around a very narrow 'equatorial belt', and so on. In quantum mechanics, pure states live on unit spheres in Hilbert spaces of usually very high dimension, and hence one naturally may wonder about their corresponding peculiarities. One of them is

P. Reimann (✉)
Fakultät für Physik, Universität Bielefeld, 33615 Bielefeld, Germany
e-mail: reimann@Physik.Uni-Bielefeld.DE

the subject of our present paper. Namely, we will show the following main result: Consider a mixed state, describing a statistical ensemble in terms of a density operator ρ with low purity, $\mathrm{tr}\,\rho^2 \ll 1$, meaning that the mixed state is very 'far' from resembling any pure state. Yet, the statistical ensemble ρ can be thought of as arising by randomly sampling pure states $|\psi\rangle$ according to some probability distribution. In fact, it is well known (see Sect. 2 for more details), that there are many different probability distributions of pure states $|\psi\rangle$ which give rise to the same mixed state ρ. Here we show that for any given ρ of low purity there exists at least one such probability distribution with the following quite astonishing property: Given an observable A, the expectation value $\langle\psi|A|\psi\rangle$ for the overwhelming majority of pure states $|\psi\rangle$ is extremely close to the ensemble averaged expectation value $\mathrm{tr}(\rho A)$ compared to the full range of *a priori* possible expectation values $\max_{|\psi\rangle}\langle\psi|A|\psi\rangle - \min_{|\psi\rangle}\langle\psi|A|\psi\rangle$ (the latter difference is tacitly assumed to be finite, as is the case for any experimentally realistic observable A, see Sect. 3).

For this kind of property, the term 'typicality' has been coined in [1]. While such 'typicality' results are applicable in principle to general quantum mechanical systems, they are obviously of particular interest with respect to the foundation of statistical physics of macroscopic systems at equilibrium, as discussed in detail e.g. in [1–9]. Further related works include [10–17]. With our present study we extend previous results from [1, 2, 4–9] to yet another important class of probability distributions of the pure states $|\psi\rangle$, namely the so-called Gaussian adjusted projected measure (GAP), recently introduced in [4].

2 Outline of the Problem

We consider a quantum mechanical system with (separable) Hilbert space $\mathcal{H}$ of dimension $N \leq \infty$. The system is assumed to be in a mixed state (statistical ensemble) described by a density matrix ρ. Let $\{|n\rangle\}_{n=1}^{N}$ be an orthonormal basis of eigenvectors of ρ and p_n the corresponding eigenvalues,

$$\rho = \sum_{n=1}^{N} p_n |n\rangle\langle n| \tag{1}$$

with the usual properties

$$p_n \geq 0 \tag{2}$$

$$\sum_{n=1}^{N} p_n = 1. \tag{3}$$

In the context of equilibrium statistical mechanics, $|n\rangle$ will usually be the eigenstates of the system Hamiltonian, but we will not make use of such a property anywhere in this paper. A particularly simple and important example is the *microcanonical density operator* with

$$\rho_{mic} = \frac{1}{N_+} \sum_{n\in S} |n\rangle\langle n| \tag{4}$$

or, equivalently, with

$$p_n = 1/N_+ \quad \text{if } n \in S, \qquad p_n = 0 \quad \text{if } n \notin S, \tag{5}$$

where S is a subset of $\{1, \ldots, N\}$, consisting of N_+ elements ($1 \leq N_+ \leq N$).

Given a density matrix of the general form (1)–(3), an arbitrary, normalized pure state (e.g. a wave function) can be written in the form

$$|\psi\rangle = \sum_{n=1}^{N} z_n |n\rangle \tag{6}$$

where $z_n := \langle n|\psi\rangle$ are complex coefficients, satisfying the normalization condition

$$\|\mathbf{z}\| = 1 \tag{7}$$

with the standard definitions

$$\mathbf{z} := (z_1, z_2, \ldots, z_N) \tag{8}$$

$$\|\mathbf{z}\| := \left(\sum_{n=1}^{N} |z_n|^2 \right)^{1/2}. \tag{9}$$

Next, we assume that the statistical ensemble ρ is generated by randomly sampling pure states (6) according to some probability density $p(\mathbf{z})$. The corresponding ensemble average of an arbitrary function $f(\mathbf{z})$ is denoted by

$$\overline{f(\mathbf{z})} := \int d\mathbf{z} f(\mathbf{z}) p(\mathbf{z}), \tag{10}$$

where $d\mathbf{z}$ represents the natural, uniform measure for the N-dimensional complex argument $\mathbf{z}$,

$$d\mathbf{z} := \prod_{n=1}^{N} d(\mathrm{Re} z_n) d(\mathrm{Im} z_n). \tag{11}$$

For infinite dimensional systems, well defined limits $N \to \infty$ are tacitly taken for granted in (10) and in similar expressions later on.

Put differently, by averaging pure states (6), represented as projectors $|\psi\rangle\langle\psi|$, according to the probability density $p(\mathbf{z})$, the given statistical ensemble ρ has to be reproduced, i.e.

$$\overline{|\psi\rangle\langle\psi|} = \sum_{m,n=1}^{N} \overline{z_m^* z_n}\, |n\rangle\langle m| = \sum_{n=1}^{N} p_n\, |n\rangle\langle n| = \rho \tag{12}$$

where the star indicates complex conjugation and where we exploited (1) in the last identity. Hence, the second moments of the distribution $p(\mathbf{z})$ are fixed by the given statistical ensemble ρ via

$$\overline{z_m^* z_n} = \delta_{mn} p_n. \tag{13}$$

In turn, every $p(\mathbf{z})$ with second moments (13) reproduces the given ρ in (12). We thus recover the well known fact that *a given density matrix ρ does not uniquely fix the distribution of pure states $p(\mathbf{z})$*. For more details, explicit examples, and further references see e.g. [4].

Next we consider an arbitrary observable $A = A^\dagger : \mathcal{H} \to \mathcal{H}$ with eigenvectors $|\nu\rangle$ and eigenvalues a_ν, i.e.

$$A = \sum_{\nu=1}^{N} a_\nu |\nu\rangle\langle\nu|. \tag{14}$$

Here and in the following we use the convention that Greek labels ν and μ implicitly refer to the eigenvectors of A, which is convenient but somewhat ambiguous in so far as, e.g., $|\nu = 3\rangle$ is not the same vector as $|n = 3\rangle$. According to (1), the ensemble averaged expectation value of A is given by

$$\langle A\rangle := \mathrm{tr}(\rho A) = \sum_{n=1}^{N} p_n A_{nn} \tag{15}$$

$$A_{mn} := \langle m|A|n\rangle. \tag{16}$$

Further, any given pure state $|\psi\rangle$ gives rise to an expectation value $\langle\psi|A|\psi\rangle \in \mathbb{R}$. The random distribution of those expectation values, induced by the distribution $p(\mathbf{z})$ of pure states, is the quantity of central interest in our present work. In particular, we will be interested in identifying conditions under which the expectation values $\langle\psi|A|\psi\rangle$ will be very close to the ensemble average (15) with very high probability, i.e. for a large majority of pure states $|\psi\rangle$.

3 Relevant Hilbert Space and Observables

Without loss of generality, we assume that the indices n in (1) are ordered so that

$$p_n > 0 \quad \text{for } n \le N_+ \quad \text{and} \quad p_n = 0 \quad \text{for } n > N_+ \tag{17}$$

for some integer N_+ with $1 \le N_+ \le N$. The N_+-dimensional sub Hilbert space spanned by the basis vectors $\{|n\rangle\}_{n=1}^{N_+}$ is denoted by $\mathcal{H}_+$ and the projector onto this subspace by

$$P_+ := \sum_{n=1}^{N_+} |n\rangle\langle n|. \tag{18}$$

In particular, P_+ is at the same time the identity operator on $\mathcal{H}_+$ and the microcanonical density operator (4) takes the form

$$\rho_{mic} = P_+/N_+. \tag{19}$$

From (13) and (17) we conclude [4] that (with probability one) $z_n = 0$ for $n > N_+$ and hence $|\psi\rangle \in \mathcal{H}_+$ according to (6). As pointed out in Sect. 2, our main goal is to determine the probability distribution of the expectation values $\langle\psi|A|\psi\rangle$. It follows that with respect to this goal only the restriction/projection

$$A_+ := P_+ A P_+ \tag{20}$$

of the observable A to the subspace $\mathcal{H}_+$ is relevant. Equivalently, *whenever m or n exceeds N_+ then the matrix element A_{mn} is of no relevance for our purpose* and thus can be set to zero without loss of generality.

The full range of possible expectation values $\langle\psi|A|\psi\rangle$ is quantified by

$$\begin{aligned}\Delta_A &:= \max_{|\psi\rangle\in\mathcal{H}_+} \langle\psi|A_+|\psi\rangle - \min_{|\psi\rangle\in\mathcal{H}_+} \langle\psi|A_+|\psi\rangle \\ &= \max_{|\psi\rangle\in\mathcal{H}_+} \langle\psi|A|\psi\rangle - \min_{|\psi\rangle\in\mathcal{H}_+} \langle\psi|A|\psi\rangle \\ &\le \max_{|\psi\rangle\in\mathcal{H}} \langle\psi|A|\psi\rangle - \min_{|\psi\rangle\in\mathcal{H}} \langle\psi|A|\psi\rangle.\end{aligned} \tag{21}$$

The second relation follows from (20) and the fact that P_+ is the projector onto the subspace $\mathcal{H}_+$ and the last relation from the fact that $\mathcal{H}_+ \subset \mathcal{H}$. Alternatively, Δ_A can thus be identified with the difference between the largest and the smallest eigenvalues of A_+ and is bounded from above by the difference between the largest and the smallest eigenvalues of A, cf. (14).

Clearly, any given real experimental apparatus has a finite range and hence the corresponding range of possible expectation values Δ_A from (21) is finite. This range is specific to the given measurement device, but is (practically) independent of the properties (e.g. the size) of the observed system. Here and in the following we restrict ourselves to observables A in the above sense. For instance, the energy of a harmonic oscillator is not an observable in this sense: in principle, the energy of the oscillator may become arbitrarily large (albeit with extremely small probability), but no real device would be able to display its value beyond a certain upper limit. Rather, all energies beyond this limit will yield one and the same measurement result (e.g. a blow up of the device), and hence only the corresponding 'truncated' energy operator would be an admissible 'observable'.

We emphasize again that the above restriction regarding the admissible observables A is of a purely mathematical/formal nature, it does not exclude any relevant observable corresponding to a realistic physical measurement. Indeed, it is well known that any realistic observable can be built up by means of suitable projection operators, and each such projector P only has eigenvalues zero and unity and hence $\Delta_P = 1$. Essentially, the same 'restriction' in fact also applies to the seminal prior works [1, 2, 9]. The measure of distance employed in [2] is the trace-norm, involving a maximization over all operators with operator norm bounded by unity. Hence, the estimates from [2] become worse and worse, as the maximally admissible norm of the considered operators increases. Similar conclusions apply for the estimates from [1]. The most explicit discussion of this issue is contained in [9].

To summarize, only the sub Hilbert space $\mathcal{H}_+$ and the projected observables (20) are of relevance for our purpose. Accordingly, we can and will for simplicity assume in some of the following sections temporarily that $\mathcal{H} = \mathcal{H}_+$ and thus $A = A_+$. In other words, all subscripts '+' will be omitted and it will be taken for granted that $p_n > 0$ for all n. Furthermore, we can and will focus on observables of finite range Δ_A according to (21).

4 The Gaussian Adjusted Projected Measure (GAP)

To avoid unnecessary technical complications, we temporarily restrict ourselves to finite dimensions N. In the final results of our calculations, the limit $N \to \infty$ can be readily performed. Furthermore, we assume $p_n > 0$ for all n without loss of generality, see end of Sect. 3.

Taking for granted the above assumptions that $N < \infty$ and $p_n > 0$, we define

$$p(\mathbf{z}) := \int d\mathbf{y} \mathcal{N} \exp\left(-\sum_{n=1}^{N} \frac{|y_n|^2}{p_n}\right) \|\mathbf{y}\|^2 \delta(\mathbf{z} - \mathbf{y}/\|\mathbf{y}\|). \tag{22}$$

Here, $d\mathbf{y}$ is defined like in (11), i.e. we are dealing with an integral over $2N$ real integration variables, and $\mathcal{N} := \prod_{n=1}^{N} (\pi p_n)^{-1}$ is a normalization constant (see below). The delta-function is by definition understood in such a way that the ensemble average of an arbitrary function $f(\mathbf{z})$ from (10) takes the form

$$\overline{f(\mathbf{z})} = \int d\mathbf{y} \mathcal{N} \exp\left(-\sum_{n=1}^{N} \frac{|y_n|^2}{p_n}\right) \|\mathbf{y}\|^2 f(\mathbf{y}/\|\mathbf{y}\|) \tag{23}$$

Equation (22) defines the Gaussian adjusted projected (GAP) measure [4] associated with the density matrix (1), written in the form of a probability density $p(\mathbf{z})$ with respect to the natural measure (11). The word 'Gaussian' in the notion GAP refers [4] to the exponential factor in (22); the word 'adjusted' refers to the factor $\|\mathbf{y}\|^2$, which is needed to fulfill condition (12) (see below); the word 'projected' refers to the delta-function in (22), guaranteeing the normalization condition (7) (see below).

In the special case of a microcanonical density operator we have $p_n = 1/N$ for all n according to (5) and our assumption at the beginning of this section, yielding with (23) the result

$$\overline{f(\mathbf{z})} = \int d\mathbf{y} \mathcal{N} e^{-\|\mathbf{y}\|^2 N} \|\mathbf{y}\|^2 f(\mathbf{y}/\|\mathbf{y}\|). \tag{24}$$

It follows that $\overline{f(U\mathbf{z})} = \overline{f(\mathbf{z})}$ for arbitrary unitary $N \times N$ matrices U. Since also $f(\mathbf{z})$ is arbitrary, we recover the fact [4] that two arguments $\mathbf{z}$ with equal length $\|\mathbf{z}\|$ are realized with equal probability.

Returning to the general case, it is often convenient to change from a Cartesian representation of the complex numbers y_n in terms of real and imaginary parts (cf. (11)) to a polar representation in terms of $r_n \geq 0$ and $\varphi_n \in [0, 2\pi)$ via the usual relation $y_n = r_n e^{i\varphi_n}$. Then, the ensemble average of an arbitrary function $f(\mathbf{z})$ from (23) can be rewritten as

$$\overline{f(\mathbf{z})} = \left[\prod_{n=1}^{N} \int_0^\infty dr_n \int_0^{2\pi} d\varphi_n \frac{r_n}{\pi p_n} e^{-r_n^2/p_n}\right] \|\mathbf{r}\|^2 f(\mathbf{c}) \tag{25}$$

$$c_n := r_n e^{i\varphi_n}/\|r\|. \tag{26}$$

We first consider the special choice $f(\mathbf{z}) := 1$, implying with (10) and (25) that

$$\int d\mathbf{z} p(\mathbf{z}) = \left[\prod_{l=1}^{N} \int_0^\infty dr_l \int_0^{2\pi} d\varphi_l \frac{r_l}{\pi p_l} e^{-r_l^2/p_l}\right] \sum_{n=1}^{N} r_n^2. \tag{27}$$

The N integrals over φ_l are trivial, each yielding a factor 2π. Hence, we can infer that

$$\int d\mathbf{z} p(\mathbf{z}) = \sum_{n=1}^{N} \prod_{l=1}^{N} \int_0^\infty dr_l \frac{2r_l^{1+2\delta_{ln}}}{p_l} e^{-r_l^2/p_l}. \tag{28}$$

The integrals over r_l for $l \neq n$ are of the form

$$\int_0^\infty dr_l \frac{2r_l}{p_l} e^{-r_l^2/p_l} = \int_0^\infty dr_l \left(-\frac{d}{dr_l}\right) e^{-r_l^2/p_l} = 1. \tag{29}$$

Likewise, the integral over r_l for $l = n$ is of the form

$$\int_0^\infty dr_l \frac{2r_l^3}{p_l} e^{-r_l^2/p_l} = \int_0^\infty dr_l r_l^2 \left(-\frac{d}{dr_l}\right) e^{-r_l^2/p_l} = \int_0^\infty dr_l 2r_l e^{-r_l^2/p_l} = p_l \tag{30}$$

where the second identity follows by a partial integration and the last identity by means of (29). All in all, the right hand side of (27) thus amounts to $\sum_{n=1}^{N} p_n$ and with (3) we see that $p(\mathbf{z})$ is normalized to unity. Observing that the right hand side in (22) is non-negative for any $\mathbf{z}$, we can conclude that $p(\mathbf{z})$ is indeed a well-defined probability density.

Next, we consider the special choice $f(\mathbf{z}) := \delta(X - \sum_{n=1}^{N} |z_n|^2)$ for an arbitrary real number X. According to (26), the argument $f(\mathbf{c})$ in (25) takes the form $\delta(X-1)$ and thus can be brought in front of all the integrals. The remaining integral is identical to the one evaluated in the preceding paragraph, i.e. its value is unity, and hence $\overline{f(\mathbf{z})} = \delta(X-1)$. It follows that $p(\mathbf{z})$ indeed takes non-zero values only for arguments $\mathbf{z}$ respecting the normalization condition (7).

Finally, we consider the special choice $f(\mathbf{z}) := z_m^* z_n$. Exploiting (25) we obtain

$$\overline{z_m^* z_n} = \left[\prod_{l=1}^{N} \int_0^\infty dr_l \int_0^{2\pi} d\varphi_l \frac{r_l}{\pi p_l} e^{-r_l^2/p_l}\right] r_m r_n e^{i(-\varphi_m+\varphi_n)}. \tag{31}$$

If $m \neq n$, the integral over φ_n can be carried out first, being proportional to $\int_0^{2\pi} d\varphi_n e^{i\varphi_n} = 0$. Hence $\overline{z_m^* z_n} = 0$ if $m \neq n$. In the case $m = n$ we have $e^{i(-\varphi_m+\varphi_n)} = 1$ and we are left with N independent integrals of the form $\int_0^{2\pi} d\varphi = 2\pi$. The remaining integrals over r_l are of the same type as those already encountered in (29) and (30) yielding the final result $\overline{z_m^* z_n} = p_n$. All together we thus find that the GAP measure (22) indeed fulfills the condition (13) and hence reproduces the correct statistical ensemble (12) encoded by the preset density operator ρ from (1). Without the 'adjusting factor' $\|\mathbf{y}\|^2$ in (22) this property could not be maintained [4].

5 Evaluation of the Variance

As pointed out in Sect. 2, our main goal is to determine the probability distribution of the expectation values $\langle\psi|A|\psi\rangle$ induced by the distribution $p(\mathbf{z})$ of pure states according to the GAP measure (22). For the first moment, $\overline{\langle\psi|A|\psi\rangle}$, the expected result $\langle A\rangle$ is readily recovered by means of (12) and (15):

$$\overline{\langle\psi|A|\psi\rangle} = \overline{\mathrm{tr}(|\psi\rangle\langle\psi|A)} = \mathrm{tr}(\overline{|\psi\rangle\langle\psi|}A) = \mathrm{tr}(\rho A) = \sum_{n=1}^{N} p_n A_{nn} = \langle A\rangle. \tag{32}$$

In the present Section, our focus is on the variance

$$\sigma_A^2 := \overline{[\langle\psi|A|\psi\rangle - \langle A\rangle]^2} = \overline{\langle\psi|A|\psi\rangle^2} - \overline{\langle\psi|A|\psi\rangle}^2. \tag{33}$$

We emphasize, that this variance characterizes the dispersion of the expectation value of A for different pure states $|\psi\rangle$, and not the "quantum fluctuations" associated with individual measurements of A of a fixed pure state $|\psi\rangle$.

Observing that the two observables A and $A - \langle A\rangle$ have the same variance and the same range Δ_A according to (21), we can and will restrict ourselves in this section without loss of generality to observables A with the property

$$\sum_{n=1}^{N} p_n A_{nn} = \langle A\rangle = 0. \tag{34}$$

Furthermore, we maintain the assumptions $N < \infty$ and $p_n > 0$ for all n, as introduced at the beginning of the previous Section. In particular, we thus have $A = A_+$ and both the eigenvalues a_ν (cf. (14)) and the diagonal matrix elements A_{nn} (cf. (16)) are bounded from above

by $a_{max} := \max_{|\psi\rangle\in\mathcal{H}} \langle\psi|A|\psi\rangle = \max_\nu a_\nu$ and from below by $a_{min} := \min_{|\psi\rangle\in\mathcal{H}} \langle\psi|A|\psi\rangle = \min_\nu a_\nu$. In view of (34) it follows that $a_{max} \geq 0$ and $a_{min} \leq 0$ and hence with (21) that

$$|A_{nn}|, |a_\nu| \leq \Delta_A \quad \text{for all } n, \nu. \tag{35}$$

With (6) and (34), the variance (33) takes the form

$$\sigma_A^2 = \overline{\left[\sum_{m,n} z_m^* z_n A_{mn}\right]^2} = \sum_{j,k=1}^N \sum_{m,n=1}^N A_{jk} A_{mn} \overline{z_j^* z_k z_m^* z_n} \tag{36}$$

The average in the last term can be rewritten by means of (25) as

$$\overline{z_j^* z_k z_m^* z_n} = \left[\prod_{l=1}^N \int_0^\infty dr_l \int_0^{2\pi} d\varphi_l \frac{r_l}{\pi p_l} e^{-r_l^2/p_l}\right] \frac{r_j r_k r_m r_n e^{i(-\varphi_j+\varphi_k-\varphi_m+\varphi_n)}}{\|\mathbf{r}\|^2}. \tag{37}$$

The evaluation of these integrals is analogous but somewhat more involved than those from the preceding Section: The integrals over the angles φ_l can be readily performed, yielding a factor of $(2\pi)^N$ in the two cases (i) $j = k$ and $m = n$, (ii) $j = n$ and $k = m$, and zero in any other case. Taking care not to count the case $j = k = m = n$ twice and after a convenient renaming of the summation indices we thus obtain

$$\sigma_A^2 = \sum_{m\neq n} [A_{mm} A_{nn} + A_{mn} A_{nm}] I_{mn} + \sum_n [A_{nn}]^2 I_{nn} \tag{38}$$

$$I_{mn} := \left[\prod_{l=1}^N \int_0^\infty dr_l \frac{2r_l}{p_l} e^{-r_l^2/p_l}\right] \frac{r_m^2 r_n^2}{\|\mathbf{r}\|^2}. \tag{39}$$

In order to evaluate the integral I_{mn}, we consider the auxiliary function

$$h(x, \mathbf{y}) := \left[\prod_{l=1}^N \int_0^\infty dr_l \frac{2r_l}{p_l}\right] \exp\left\{-\sum_{l=1}^N (x + y_l) r_l^2\right\}, \tag{40}$$

where $x \geq 0$, $\mathbf{y} := (y_1, \ldots, y_N)$, and $y_n > 0$ for all n. Observing that the right hand side in (40) factorizes into N independent integrals of the form $\int_0^\infty dr_l r_l e^{-br_l^2} = 1/2b$ with $b := x + y_l > 0$ (see also (29)), we obtain

$$h(x, \mathbf{y}) = \prod_{l=1}^N \frac{1}{p_l} \frac{1}{x + y_l}. \tag{41}$$

Next, we note that the integral over the x-dependent terms in (40) is of the form $\int_0^\infty dx e^{-x\|\mathbf{r}\|^2} = 1/\|\mathbf{r}\|^2$ (see (9)), implying that

$$H(\mathbf{y}) := \int_0^\infty dx\, h(x, \mathbf{y}) = \left[\prod_{l=1}^N \int_0^\infty dr_l \frac{2r_l}{p_l}\right] \exp\left\{-\sum_{l=1}^N y_l r_l^2\right\} / \|\mathbf{r}\|^2. \tag{42}$$

By comparison with (39) we can conclude that

$$I_{mn} = \left.\frac{\partial^2 H(\mathbf{y})}{\partial y_m \partial y_n}\right|_{y_l = 1/p_l}. \tag{43}$$

By combining (41)–(43) it follows that

$$I_{mn} = \int_0^\infty dx \frac{\partial^2}{\partial y_m \partial y_n} \prod_{l=1}^N \frac{1}{p_l} \frac{1}{x+y_l} \Bigg|_{y_l=1/p_l} = p_m p_n (1+\delta_{mn}) K_{mn} \tag{44}$$

$$K_{mn} := \int_0^\infty dx\, g_{mn}(x) G(x) \tag{45}$$

$$g_{nm}(x) := (1+xp_m)^{-1}(1+xp_n)^{-1} \tag{46}$$

$$G(x) := \prod_{l=1}^N (1+xp_l)^{-1}. \tag{47}$$

Finally, this yields for the variance (38) the result

$$\sigma_A^2 = \sum_{m,n=1}^N [A_{mm}A_{nn} + A_{mn}A_{nm}] p_m p_n K_{mn}. \tag{48}$$

Next we turn to a more detailed discussion of the integrals K_{mn} in (45). Clearly, the integrand is a positive function of x, bounded from above by unity, and decaying like $1/x^{N+2}$ for large x due to our assumption that $p_n > 0$ for all n, see below (34). Hence the integrals K_{mn} are finite and positive. Specifically, for the microcanonical density operator we have $p_n = 1/N$ for all n according to (5) and our assumption below (34), yielding with (45) the exact result

$$K_{mn} = \int_0^\infty dx (1+x/N)^{-N-2} = \frac{N}{N+1}. \tag{49}$$

To further evaluate K_{mn} in the general case, we rewrite $g_{mn}(x)$ from (46) by means of Taylor's theorem [19] as

$$\begin{aligned} g_{mn}(x) &= g_{mn}(0) + x g'_{mn}(0) + \frac{x^2}{2} g''_{mn}(x\theta_{mn}(x)) \\ &= 1 - x(p_m+p_n) + x^2(p_m^2 + p_m p_n + p_n^2)\chi_{mn}(x) \end{aligned} \tag{50}$$

for certain functions $\theta_{mn}(x)$ and $\chi_{mn}(x)$, satisfying $\theta_{mn}(x)$, $\chi_{mn}(x) \in [0,1]$ for all $x \geq 0$, but for the rest depending in a non-trivial manner on x, m, and n. Note that while an infinite power series expansion would not converge for arbitrary $x \geq 0$, the above finite order Taylor expansion is an exact identity [19] for all $x \geq 0$. As a consequence, (45) can be rewritten as

$$K_{mn} = K^{(0)} - (p_m+p_n)K^{(1)} + 2(p_m^2 + p_m p_n + p_n^2)\kappa_{mn} K^{(2)} \tag{51}$$

$$K^{(k)} := \frac{1}{k!} \int_0^\infty dx\, x^k G(x), \quad k = 0,1,2 \tag{52}$$

$$\kappa_{mn} \in [0,1]. \tag{53}$$

From (47) and $p_n > 0$ for all n we can infer that the integrals in (52) are finite (and positive) if and only if

$$N \geq 4. \tag{54}$$

The latter condition is tacitly taken for granted henceforth.

Next we return to the variance in (48). In view of (34) we see that sums of the form

$$\sum_{m,n=1}^{N} A_{mm}A_{nn}p_m p_n Q_{mn} = \sum_{m=1}^{N} A_{mm}p_m \sum_{n=1}^{N} A_{nn}p_n Q_{mn} \tag{55}$$

are zero if the coefficients Q_{mn} are either independent of m or independent of n. Hence the first sum on the right hand side of (48) vanishes in the special case (49) corresponding to the microcanonical density operator. Likewise, in the general case we can conclude with (51) that the first sum on the right hand side of (48) takes the form

$$\sum_{m,n=1}^{N} A_{mm}A_{nn}p_m p_n K_{mn} = 2K^{(2)} \sum_{m,n=1}^{N} A_{mm}A_{nn}p_m p_n (p_m^2 + p_m p_n + p_n^2)\kappa_{mn}, \tag{56}$$

yielding with (53) and $K^{(2)} \geq 0$ the estimate

$$\left| \sum_{m,n=1}^{N} A_{mm}A_{nn}p_m p_n K_{mn} \right| \leq 2K^{(2)} \sum_{m,n=1}^{N} |A_{mm}||A_{nn}|(2p_m p_n^3 + p_m^2 p_n^2). \tag{57}$$

With (3) and (35) we obtain

$$\sum_{m,n=1}^{N} A_{mm}A_{nn}p_m p_n K_{mn} \leq 2K^{(2)}\Delta_A^2 \left(2\sum_{n=1}^{N} p_n^3 + \left[\sum_{n=1}^{N} p_n^2 \right]^2 \right). \tag{58}$$

Turning to the second sum on the right hand side of (48), we note that

$$0 \leq \sum_{m,n=1}^{N} A_{mn}A_{nm}p_m p_n K_{mn} \leq K^{(0)} \sum_{m,n=1}^{N} A_{mn}A_{nm}p_m p_n. \tag{59}$$

The first inequality follows from the fact that $A_{mn}A_{nm} = |A_{mn}|^2 \geq 0$, $p_m p_n \geq 0$, and $K_{mn} \geq 0$ for all m, n. The second inequality follows from $g_{mn}(x) \leq 1$ according to (46), hence $K_{mn} \leq K^{(0)}$ according to (45) and (52). By means of (1) and (16) one readily finds that

$$\begin{aligned} \mathrm{tr}(\rho A)^2 &= \sum_{m=1}^{N} \langle m|\rho A \sum_{n=1}^{N} |n\rangle\langle n|\rho A|m\rangle = \sum_{m,n=1}^{N} p_m \langle m|A|n\rangle p_n \langle n|A|m\rangle \\ &= \sum_{m,n=1}^{N} p_m p_n A_{mn}A_{nm}. \end{aligned} \tag{60}$$

Likewise, by using the eigenvectors $|\nu\rangle$ and eigenvalues a_ν of A from (14) to evaluate the trace one obtains

$$\mathrm{tr}(\rho A)^2 = \sum_{\mu=1}^{N} \langle \mu|\rho A \sum_{\nu=1}^{N} |\nu\rangle\langle \nu|\rho A|\mu\rangle = \sum_{\mu,\nu=1}^{N} \langle \mu|\rho|\nu\rangle a_\nu \langle \nu|\rho|\mu\rangle a_\mu = \sum_{\mu,\nu=1}^{N} a_\mu a_\nu \rho_{\mu\nu}\rho_{\nu\mu}. \tag{61}$$

Combining (59)–(61) and $\rho_{\mu\nu}\rho_{\nu\mu} = |\rho_{\mu\nu}|^2$ yields

$$\sum_{m,n=1}^{N} A_{mn}A_{mn}p_m p_n K_{mn} \leq K^{(0)} \sum_{\mu,\nu=1}^{N} a_\nu a_\mu |\rho_{\mu\nu}|^2 \leq K^{(0)}\Delta_A^2 \sum_{\mu,\nu=1}^{N} |\rho_{\mu\nu}|^2 = K^{(0)}\Delta_A^2 \operatorname{tr}\rho^2, \quad (62)$$

where the second inequality follows from (35) and the last equality from (61) with $A = 1$.

In the special of a microcanonical density operator (4) we have seen below (55) that the first sum on the right hand side of (48) vanishes. Exploiting (49), (60), and the fact that $p_n = 1/N$ for all n we obtain

$$\sigma_A^2 = \frac{N}{N+1}\operatorname{tr}(\rho_{mic}A)^2 = \frac{\operatorname{tr}A^2}{N(N+1)} = \frac{\sum_{\nu=1}^{N} a_\nu^2}{N^2}\left[1 + \mathcal{O}\left(\frac{1}{N}\right)\right], \quad (63)$$

where a_ν are the eigenvalues of A, see (14).

Returning to the general case, the variance (48) can be estimated from above by means of (58), (62), and the relations $\sum_{n=1}^{N} p_n^2 = \operatorname{tr}\rho^2$, $\sum_{n=1}^{N} p_n^3 \leq (\operatorname{tr}\rho^2)^{3/2}$, derived in the Appendix, as follows

$$\sigma_A^2 \leq K^{(0)}\Delta_A^2 \operatorname{tr}\rho^2 + 2K^{(2)}\Delta_A^2\left(2[\operatorname{tr}\rho^2]^{3/2} + [\operatorname{tr}\rho^2]^2\right). \quad (64)$$

Our next goal is to find upper and lower bounds for $G(x)$ for $x > 0$ ($x = 0$ is trivial) in order to estimate $K^{(k)}$ from (52). To this end, we consider $x > 0$ as arbitrary but fixed, and consider the right hand side in (47) as a function of $\mathbf{p} := (p_1, \ldots, p_N)$,

$$Q(\mathbf{p}) := \prod_{n=1}^{N}(1 + xp_n)^{-1}. \quad (65)$$

The basic idea is to determine its maximum and the minimum under the three constraints (2), (3), and $p_n \leq p_{max}$ for all n, where

$$p_{max} := \max_n p_n. \quad (66)$$

The differential/variation of (65) reads

$$\delta Q(\mathbf{p}) := -xQ(\mathbf{p}) \sum_{n=1}^{N} \frac{\delta p_n}{1 + xp_n}, \quad (67)$$

complemented by the constraints $\sum \delta p_n = 0$, $\delta p_n \geq 0$ if $p_n = 0$, and $\delta p_n \leq 0$ if $p_n = p_{max}$. Observing that $xQ(\mathbf{p}) > 0$ on the right hand side of (67) and that the factors $1/(1 + xp_n)$ are smaller (but still positive) for large p_n than for small p_n implies that $Q(\mathbf{p})$ can always be made larger ($\delta Q(\mathbf{p}) > 0$) by making the already large p_n still larger ($\delta p_n > 0$) and the already small p_n still smaller ($\delta p_n < 0$). As a consequence, $Q(\mathbf{p})$ is minimal if all p_n are equal, implying that

$$G(x) \geq (1 + x/N)^{-N}. \quad (68)$$

On the other hand, $Q(\mathbf{p})$ cannot be increased any more if and only if the small p_n have reached the lower limit $p_n = 0$ and the large p_n the upper limit $p_n = p_{max}$. Denoting by N_{max} the number of those p_n equal to p_{max}, their total weight $N_{max}p_{max}$ is generically still

not exactly equal to unity for any integer N_{max}. Hence there must remain one weight p_n with a value $1 - N_{max}p_{max} =: p_0 \in [0, p_{max}]$ in order to fulfill the constraint (3). All in all, this implies the upper bound

$$G(x) \leq (1 + xp_{max})^{-N_{max}}(1 + xp_0)^{-1}. \tag{69}$$

Next we note that for any $a > 0$ the auxiliary function $f(y) := \ln(1 + ay) - y\ln(1 + a)$ is zero for $y = 0$ and $y = 1$ and has a negative second derivative for $y \geq 0$, implying that $f(y) \geq 0$ for all $y \in [0, 1]$. Setting $a = xp_{max}$ and $y = p_0/p_{max}$ it follows that $(1+xp_0)^{p_{max}} \geq (1 + xp_{max})^{p_0}$ and due to $p_0 := 1 - N_{max}p_{max}$ that $1 + xp_0 \geq (1 + xp_{max})^{1/p_{max}-N_{max}}$. With (69) we thus can infer that

$$G(x) \leq (1 + xp_{max})^{-1/p_{max}}. \tag{70}$$

From (3) and (66) we see that

$$p_{max} \geq 1/N. \tag{71}$$

Further, the upper bound (70) yields finite integrals (52) only if

$$p_{max} < 1/3. \tag{72}$$

Note that this condition implies $N > 3$ and hence condition (54) is automatically satisfied. Taking for granted (72) we can infer by exploiting the bounds (68) and (70) in (52) and after performing k partial integrations that

$$K^{(k)} = \prod_{j=1}^{k+1} \frac{1}{1 - jp^{(k)}}, \quad k = 0, 1, 2 \tag{73}$$

$$p^{(k)} \in [1/N, p_{max}] \tag{74}$$

With (64) we thus obtain for the variance the upper bound

$$\sigma_A^2 \leq \Delta_A^2 \left(\frac{\mathrm{tr}\,\rho^2}{1 - p_{max}} + \frac{4[\mathrm{tr}\,\rho^2]^{3/2} + 2[\mathrm{tr}\,\rho^2]^2}{(1 - p_{max})(1 - 2p_{max})(1 - 3p_{max})} \right). \tag{75}$$

6 Discussion of the Main Results

The upper bound (75) for the variance from (33) is the first main result of our paper. In our derivation we have assumed that $\langle A \rangle = 0$ (see (34)), but since the variance from (33) and also all the other quantities appearing in (75) remain unchanged upon replacing A by $A - \langle A \rangle$ we can conclude that (75) remains valid for arbitrary A. Moreover, we made the assumption that $N < \infty$ and $p_n > 0$ for all n in deriving (75). Since neither of the quantities appearing in the final result (75) give rise to any problem in the limit $p_n \to 0$, the assumptions $p_n > 0$ can be given up as well. Finally, the limit $N \to \infty$ depends on the meaning and existence of this limit for the quantities appearing on the right hand side of (75). In particular after dividing both sides by Δ_A^2 (see below), we expect that in many important cases this limit will not give rise to any problems. The only remaining condition for (75) to be applicable is thus $p_{max} < 1/3$ (see (72)).

In the special case of a microcanonical density operator (4) we have obtained as a second main result the exact relation (63) for the variance under the same assumptions as above, namely $\langle A\rangle = 0$, $N < \infty$, and $p_n > 0$ for all n. Accordingly, for more general observables A with $\langle A\rangle \neq 0$ we have to replace A by $A - \langle A\rangle$ in (63). Next, if not all p_n are positive and thus equal to $1/N$, then we have to replace A by A_+ and N by N_+, as discussed at the beginning of Sect. 3. All in all, we thus obtain for a microcanonical density operator (4) the general exact result

$$\sigma_A^2 = \frac{N_+}{N_+ + 1}\,\mathrm{tr}[\rho_{mic}(A_+ - \langle A_+\rangle)]^2 = \frac{1}{N_+ + 1}\left(\frac{\mathrm{tr}\,A_+^2}{N_+} - \left[\frac{\mathrm{tr}\,A_+}{N_+}\right]^2\right), \tag{76}$$

where A_+ is the projection of the original operator A onto the subspace spanned by the basis vectors $|n\rangle$ with non-trivial weights $p_n > 0$, see (20). As before, the meaning and existence of the limit $N \to \infty$ depends on the behavior of N_+, $\mathrm{tr}\,A_+/N_+$, and $\mathrm{tr}\,A_+^2/N_+$ in this limit, but is expected not to give rise to any problems in many important cases.

Results similar to (76) have been previously derived in [10], in [11] (see formula (C.17) therein), and in [18] (see Lemma 3 therein, whose proof is very close in spirit to [1]). The main difference is that these results only apply to the special case that $S = \{1, \ldots, N\}$ and hence $N_+ = N$ in (5), implying that ρ_{mic} in (4) is proportional to the identity operator, cf. (19). At first glance, a further difference appears to be that the above mentioned results do not refer to the GAP measure (22) associated with the above ρ_{mic} but rather are derived under the assumption that all (normalized) pure states $|\psi\rangle \in \mathcal{H}$ are realized with equal probability. However, by noting that the latter assumption uniquely determines the probability density $p(\mathbf{z})$ for the coefficients z_n in (6) and that the GAP measure does fulfill the assumption (see below (24)) we can conclude that there is in fact no difference in this respect. As a by product we can infer that (76) in particular applies to the case that all (normalized) pure states $|\psi\rangle$ within the subspace $\mathcal{H}_+ \subset \mathcal{H}$ are realized with equal probability and all other $|\psi\rangle \in \mathcal{H}$ are excluded. After submission of this paper, A. Sugita pointed out that the same finding is also contained in his recent work [9].

Of particular interest in (75) are situations for which the bracket on the right hand side becomes a small quantity. Therefore, we now focus on the case that the so-called purity $\mathrm{tr}\,\rho^2$ of the mixed state ρ is low, i.e.

$$\mathrm{tr}\,\rho^2 \ll 1. \tag{77}$$

We recall the well-known facts that the purity is one if and only if ρ corresponds to a pure state ($\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle \in \mathcal{H}$), is smaller than unity in any other case, and takes the minimal possible value $1/N$ if $p_n = 1/N$ for all n in (1). Roughly speaking, a low purity $\mathrm{tr}\,\rho^2$ thus means that the mixed state ρ is very 'far' from any pure state $|\psi\rangle \in \mathcal{H}$.

According to the Appendix, the quantity p_{max} from (66) can be estimated from above and from below as follows

$$\mathrm{tr}\,\rho^2 \leq p_{max} \leq \sqrt{\mathrm{tr}\,\rho^2}. \tag{78}$$

Hence, assumption (77) is fulfilled if and only if all p_n in (1) are small and is tantamount to the condition

$$p_{max} \ll 1. \tag{79}$$

Accordingly, there cannot be just a few dominating p_n in (1) in the sense that their sum would already be of the order of unity. In particular, the dimensionality N of the Hilbert space $\mathcal{H}$ must be large according to (71).

Exploiting (77) and (78) in (75) yields our final main result

$$\sigma_A^2 \leq \Delta_A^2 \operatorname{tr}\rho^2 (1 + \mathcal{O}(\sqrt{\operatorname{tr}\rho^2})). \tag{80}$$

By rewriting (76) in a form analogous to the last expression in (63) one readily sees that the upper bound (80) is expected to be quite tight in typical cases.

7 Summary and Conclusions

Given any mixed state ρ of low purity (77) there exists at least one probability distribution $p(\mathbf{z})$ of pure states (6), namely the GAP measure (22), with the following properties: (i) By randomly sampling pure states $|\psi\rangle$ according to this probability distribution, the preset statistical ensemble ρ is reproduced. (ii) Given any observable A, for the overwhelming majority of pure states $|\psi\rangle$ sampled according to $p(\mathbf{z})$ the expectation value $\langle\psi|A|\psi\rangle$ deviates extremely little from the ensemble averaged expectation value $\operatorname{tr}(\rho A)$ compared to the full range Δ_A of a priori possible outcomes of a measurement corresponding to A. The latter statement can be expressed more rigorously [5] by means of (80) in combination with Chebyshev's inequality [20], and it is tacitly assumed that this range Δ_A from (21) is nonzero and remains bounded even in the case of an infinite dimensional Hilbert space $\mathcal{H}$, as is the case for any experimentally realistic observable A (see Sect. 3).

On the one hand, in general there are other measures $p(\mathbf{z})$ besides the GAP measure which also satisfy property (i) above but for which property (ii) may not necessarily remain true. On the other hand, ρ fixes all observable properties of the system via (15), so that under typical circumstances any further information regarding $p(\mathbf{z})$ is neither necessary nor available. Hence, in order to uniquely specify $p(\mathbf{z})$ for a given ρ, one has either to introduce and justify additional postulates regarding $p(\mathbf{z})$ [1, 2], or to show that many or all of the $p(\mathbf{z})$ compatible with ρ lead to essentially the same final conclusions (ii) [5], or one has to include the preparation and equilibration process of the system into the consideration [12, 13]. In our present work we have focused on the first among those three options.

The justification for selecting the GAP measure has been discussed in detail in [4]. In particular, it is argued in [4] that this measure arises naturally when considering macroscopic systems in thermal equilibrium and hence is the most appropriate choice, at least in cases when ρ is known to be the canonical density matrix. Furthermore, as shown in [4] and again in Sect. 6, this measure is the unique solution in the case of a microcanonical density operator (4) under the additional assumption that all (normalized) pure states $|\psi\rangle \in \mathcal{H}_+$ are equally likely and all other $|\psi\rangle \in \mathcal{H}$ are excluded, where $\mathcal{H}_+$ is the sub Hilbert space spanned by all the eigenvectors $|n\rangle$ with $p_n > 0$ in (1), i.e. the quantum mechanical analogue of the classical energy shell within the standard microcanonical formalism.

In other words, whenever the Hilbert space $\mathcal{H}$ of the system contains a subspace $\mathcal{H}_+ \subset \mathcal{H}$ with the property that all $|\psi\rangle \in \mathcal{H}_+$ are realized with equal probability and all other $|\psi\rangle \in \mathcal{H}$ are excluded then the variance Δ_A, characterizing the dispersion of the random variable $\langle\psi|A|\psi\rangle$ (see (33)), is given by the exact relation (76), where N_+ is the dimension of $\mathcal{H}_+$ and A_+ the restriction/projection of A to $\mathcal{H}_+$ (see (20)).

Acknowledgements Special thanks is due Chris Van den Broeck and Jochen Gemmer for inspiring discussions and to David Speer for carefully reading the manuscript.

Appendix

By means of (1), the definition of p_{max} in (66) and the normalization (3) we can conclude that

$$\operatorname{tr}\rho^{k+1} = \sum_{n=1}^{N} p_n^{k+1} \leq \sum_{n=1}^{N} p_{max}^k p_n = p_{max}^k \tag{81}$$

for any integer $k \geq 0$. For the microcanonical density operator (4), the above inequality becomes an equality, i.e. the lower bound for p_{max} following from (81) cannot be improved in general. Likewise, one readily sees that

$$\operatorname{tr}\rho^k = \sum_{n=1}^{N} p_n^k \geq \max_n p_n^k = p_{max}^k. \tag{82}$$

Here the inequality becomes an equality if $p_n \to 1$ for one index n and $p_m \to 0$ for all $m \neq n$, and hence again no general improvement of the corresponding upper bound for p_{max} is possible. In particular, for $k = 2$ we have $p_{max}^2 \leq \operatorname{tr}\rho^2$ and hence we can conclude that

$$0 \leq p_n^k \leq p_{max}^k = (p_{max}^2)^{k/2} \leq (\operatorname{tr}\rho^2)^{k/2} \tag{83}$$

for any integer $k \geq 1$. Finally, this result yields

$$\operatorname{tr}\rho^k = \sum_{n=1}^{N} p_n^k \leq \sum_{n=1}^{N} p_{max}^{k-2} p_n^2 = p_{max}^{k-2} \operatorname{tr}\rho^2 \leq (\operatorname{tr}\rho^2)^{(k-2)/2} \operatorname{tr}\rho^2 = (\operatorname{tr}\rho^2)^{k/2} \tag{84}$$

for any integer $k \geq 2$.

References

1. Goldstein, S., Lebowitz, J.L., Tumulka, R., Zanghi, N.: Phys. Rev. Lett. **96**, 050403 (2006)
2. Popescu, S., Short, A.J., Winter, A.: Nature Phys. **2**, 754 (2006); quant-ph/0511225
3. Lloyd, S.: Nature Phys. **2**, 727 (2006)
4. Goldstein, S., Lebowitz, J.L., Tumulka, R., Zanghi, N.: J. Stat. Phys. **125**, 1197 (2006)
5. Reimann, P.: Phys. Rev. Lett. **99**, 160404 (2007)
6. Reimann, P.: Phys. Rev. Lett. **100**, 119901 (2008) [publishers note]
7. Brody, D.C.: Phys. Rev. Lett. **100**, 148901 (2008) [comment]
8. Reimann, P.: Phys. Rev. Lett. **100**, 148902 (2008) [reply]
9. Sugita, A.: Nonlinear Phenom. Complex Syst. **10**, 192 (2007)
10. Lloyd, S.: Black holes, demons and the loss of coherence. Thesis, Rockefeller Univ. (1988)
11. Gemmer, J., Michel, M., Mahler, G.: Quantum Thermodynamics. Springer, Berlin (2004)
12. Jensen, R.V., Shankar, R.: Phys. Rev. Lett. **54**, 1879 (1985)
13. Rigol, M., Dunjko, V., Olshanii, M.: arXiv:0708.1324v1 (2007)
14. Bocchieri, P., Loinger, A.: Phys. Rev. **114**, 948 (1959)
15. Tasaki, H.: Phys. Rev. Lett. **80**, 1373 (1998)
16. Page, D.N.: Phys. Rev. Lett. **71**, 1291 (1993)
17. Lubkin, E.: J. Math. Phys. **19**, 1028 (1978)
18. Hayden, P., Shor, P.W., Winter, A.: arXiv:0712.0975 (2007)
19. Wikipedia, the free encyclopedia: Taylor's theorem. en.wikipedia.org (2008)
20. Wikipedia, the free encyclopedia: Chebyshev's inequality. en.wikipedia.org (2008)