Detecting entanglement of random states with an entanglement witness

2007 J. Phys. A: Math. Theor. 40 13787

(http://iopscience.iop.org/1751-8121/40/45/017)

J. Phys. A: Math. Theor. **40** (2007) 13787–13798 doi:10.1088/1751-8113/40/45/017

Detecting entanglement of random states with an entanglement witness

Marko Žnidarič[1], Tomaž Prosen[1], Giuliano Benenti[2,3] and Giulio Casati[2,3,4]

[1] Physics Department, Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia
[2] CNISM, CNR-INFM and Center for Nonlinear and Complex Systems, Università degli Studi dell'Insubria, via Valleggio 11, I-22100 Como, Italy
[3] Instituto Nazionale di Fisica Nucleare, Sezione di Milano, via Celoria 16, I-20133 Milano, Italy
[4] Department of Physics and Centre for Computational Science and Engineering, National University of Singapore, Singapore 117542, Republic of Singapore

Received 18 June 2007, in final form 31 August 2007
Published 23 October 2007
Online at stacks.iop.org/JPhysA/40/13787

Abstract

The entanglement content of high-dimensional random pure states is almost maximal; nevertheless, we show that, due to the complexity of such states, the detection of their entanglement using witness operators is rather difficult. We discuss the case of *unknown* random states, and the case of *known* random states for which we can optimize the entanglement witness. Moreover, we show that coarse graining, modelled by considering mixtures of m random states instead of pure ones, leads to a decay in the entanglement detection probability exponential with m. Our results also allow us to explain the emergence of classicality in coarse grained quantum chaotic dynamics.

PACS numbers: 03.65.Ud, 03.67.−a, 05.45.Mt

1. Introduction

Random pure quantum states, i.e., states whose distribution is invariant under an arbitrary unitary transformation, are almost maximally entangled. Let us consider, for instance, a random state $|\psi\rangle$ from $N \times N$ dimensional Hilbert space $\mathcal{H} = \mathbb{C}^{N^2}$. We consider a bipartition to subsystems A and B, namely $|\psi\rangle$ as an element of a tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ where $\mathcal{H}_A = \mathbb{C}^N, \mathcal{H}_B = \mathbb{C}^N$ are Hilbert spaces of subsystems A and B, respectively. The entanglement E of the state $|\psi\rangle$ is then measured by the von Neumann entropy $S(\rho) = -\mathrm{tr}(\rho \log_2 \rho)$ of the reduced density matrices $\rho_A = \mathrm{tr}_B|\psi\rangle\langle\psi|$ or $\rho_B = \mathrm{tr}_A|\psi\rangle\langle\psi|$. It turns out [1–3] that

$$E(|\psi\rangle) = S(\rho_A) = S(\rho_B) \approx \log_2 N - \frac{1}{\log_e 4}, \tag{1}$$

which is close to the maximal value $\log_2 N$ attained for the maximally entangled state. Since random states carry a lot of entanglement and entanglement [4, 5] has no analogue in classical mechanics, one can conclude that random states are highly non-classical.

On the other hand, pseudo-random states with properties close to those of true random states can be efficiently generated by dynamical systems (maps) in the regime of quantum chaos [6–11]. In such chaotic maps the classical limit is recovered when $N \to \infty$. Therefore, one can argue that for high-dimensional random states, i.e., in the limit $N \to \infty$, the quantum expectation value of an operator with a well-defined classical limit will be close to its classical microcanonical average. According to this picture random states in a way 'mimic' classical microcanonical density. Expectation values are therefore close to the classical ones.

At first sight this is in striking contrast with the almost maximal entanglement of such states. How can we reconcile this apparent contradiction? In the present paper, we are going to tackle this question by considering how can we detect entanglement of random states. By studying the detection of entanglement with decomposable entanglement witnesses we are going to argue that in the limit of large systems the detection of entanglement in a random state becomes increasingly difficult as it would demand the control of very finely interwoven degrees of freedom and a measurement resolution inversely proportional to the size of the Hilbert space, which seems hardly feasible experimentally. Therefore, as far as the detection of entanglement is concerned, high dimensional random states are effectively classical.

Moreover, coarse graining naturally appears. For instance, one could repeat several times the measurement of the entanglement witness for a random state and the prepared random state would be different from time to time due to unavoidable experimental imperfections. We model this problem by considering mixtures of m pure random states, namely

$$\rho = \sum_{i=1}^{m} \frac{1}{m} |\psi_i\rangle\langle\psi_i|, \tag{2}$$

where $|\psi_i\rangle$ are mutually independent random pure states, but in general they are not orthogonal. We are going to show that the detection of entanglement is even more difficult for these mixed states, as it requires a number of measurements growing exponentially with m.

There are other physical contexts in which formally the same kind of coarse graining (mixing of the state) naturally appears: (i) time averaging. For example, if a state $|\psi\rangle$ undergoes a time evolution $|\psi(t)\rangle = U(t)|\psi\rangle$ given in terms of some unitary dynamics $U(t)$, then the time average of a physical observable A over an interval T is essentially determined by the expectation value $\mathrm{tr}(A\rho)$ in the mixed state

$$\rho = \frac{1}{T} \int_0^T \mathrm{d}t\, |\psi(t)\rangle\langle\psi(t)| \tag{3}$$

which has an effective rank $m \approx T/t_{\mathrm{corr}}$ where t_{corr} is a dynamical correlation time of the dynamics $U(t)$. For a quantum chaotic evolution $U(t)$, the state $|\psi(t)\rangle$ can be, after some time, arguably well described by a random state and the correlation time t_{corr} is expected to be short, so ρ (3) may be considered as a mixture (2) of m uncorrelated random states. (ii) *Phase space averaging.* Sometimes it is useful to represent quantum states in terms of distribution functions in the classical phase space, like the Husimi function (see, e.g., [7]), which can be understood as a convolution of the Wigner function or its coarse graining over a phase-space volume $(2\pi\hbar)^d$ in d degrees of freedom. In fact, the Husimi function of a pure state can be understood as a Wigner function of the following mixed state,

$$\rho = (2\pi\hbar)^{-d} \int \mathrm{d}\vec{q}\, \mathrm{d}\vec{p} \exp\left(-\frac{1}{2\hbar}(\alpha q^2 + \alpha^{-1} p^2)\right) T(\vec{q}, \vec{p})|\psi\rangle\langle\psi|T^\dagger(\vec{q}, \vec{p}), \tag{4}$$

where $T(\vec{q}, \vec{p})$ are unitary phase space translation operators, and α is an arbitrary squeezing parameter. A random pure state $|\psi\rangle$ has a Wigner function with random sub-Planck structures with the phase space correlation length $l_{\mathrm{corr}} \sim \hbar$ which is semi-classically smaller than the coarse-graining width $\sim\hbar^{1/2}$, so ρ (4) can be again considered as a mixture (2) of m random pure states with $m \sim \hbar^{-1/2}$.

The paper is organized as follows. In section 2, we review known results about entanglement witnesses and pure states entanglement. Depending on our *a priori* knowledge of the random state, two relevant cases can be distinguished: (i) we do not know in advance of which random state we are going to detect the entanglement, i.e., the random state is unknown. In such a case the best one can do is to choose some generic entanglement witness in advance, independently of the state we measure. Such a situation can also be thought to arise in the case when we are not able to prepare an arbitrary entanglement witness but just some subset of witnesses. (ii) We know the random state in advance and are able to prepare an arbitrary entanglement witness. In such a case, we can use the optimal entanglement witness for each random state separately. These two cases are discussed in sections 3 and 4, respectively. Finally, in section 5, we provide a brief discussion of our results.

2. Entanglement witnesses

First, let us introduce some known facts about entanglement witnesses and entanglement of pure states that we will need for the derivation of our results in subsequent sections. Given a pure state $|\psi\rangle$, its bipartite entanglement content is completely specified by the Schmidt decomposition:

$$|\psi\rangle = \sum_i \mu_i |a_i\rangle \otimes |b_i\rangle, \tag{5}$$

where the Schmidt coefficients μ_i are positive real numbers satisfying $\sum_i \mu_i^2 = 1$ and $|a_i\rangle$ and $|b_i\rangle$ are orthonormal states on subspaces $\mathcal{H}_{\mathrm{A}}$ and $\mathcal{H}_{\mathrm{B}}$. The squares of the Schmidt coefficients are equal to the eigenvalues of the reduced density matrices ρ_{A} and ρ_{B}. For such bipartition there are at most N nonzero Schmidt coefficients μ_i. For a state having r roughly equal nonzero Schmidt coefficients ($\mu_i \sim 1/\sqrt{r}$), the reduced von Neumann entropy is $S(\rho_A) \sim \log_2 r$. For instance, the simple (GHZ) state $|\psi\rangle = \frac{1}{\sqrt{2}}|0, \ldots, 0\rangle + \frac{1}{\sqrt{2}}|1, \ldots, 1\rangle$ leads to two nonzero eigenvalues of ρ_{A} ($\mu_1 = \mu_2 = 1/\sqrt{2}$), and simple Schmidt vectors $|a_1\rangle \otimes |b_1\rangle = |0, \ldots, 0\rangle$ and $|a_2\rangle \otimes |b_2\rangle = |1, \ldots, 1\rangle$, resulting in $S = 1$. On the other hand, a random pure state has all N eigenvalues nonzero, very complicated (random) Schmidt vectors $|a_i\rangle$ and $|b_i\rangle$, and almost maximal entropy $S \approx \log_2 N - 1/\log_e 4$ [1–3]. Note, however, that all the eigenvalues of the reduced density matrix for a bipartite random state are very small. In fact, they decrease with N, the largest being, on average, $4/N$ while the smallest is $1/N^3$ [12]. The guiding idea of this paper is that this smallness of eigenvalues and the complexity of eigenvectors should be somehow reflected in the difficulty of detecting entanglement in such a state, despite the fact that the entanglement content of a random state is large.

Besides pure states, we will also be interested in the entanglement of a mixture of pure states (2). A quantum state ρ of a bipartite system is called separable if it can be written as

$$\rho = \sum_k p_k \rho_{\mathrm{A}k} \otimes \rho_{\mathrm{B}k}, \qquad p_k \geqslant 0, \qquad \sum_k p_k = 1, \tag{6}$$

where $\rho_{\mathrm{A}k}$ and $\rho_{\mathrm{B}k}$ are density matrices for the two subsystems. A state is entangled if it is not separable. To decide whether a given mixed state is entangled or not is a difficult problem [4, 5]. Fortunately, there is an operational criterion which is able to detect the

most useful entangled states. This is the famous positive partial transposition (PPT) criterion [13]: since separable states have a positive semidefinite partial transpose[5], all non-PPT states are entangled. Note, however, that for dimensions higher than 2×2 and 2×3 there exist PPT-entangled states, known as bound entangled states [14].

A convenient way to detect entanglement is to use the so-called entanglement witnesses [14, 15]. By definition, an entanglement witness is a Hermitian operator W such that $\mathrm{tr}(W\rho_{\mathrm{sep}}) \geqslant 0$ for all separable states ρ_{sep} while there exists at least one state ρ_{ent} such that $\mathrm{tr}(W\rho_{\mathrm{ent}}) < 0$. Therefore, the negative expectation value of W is a signature of entanglement and the state ρ_{ent} is said to be detected as entangled by the witness W. The concept of entanglement witness is close to experimental implementations and detection of entanglement by means of entanglement witnesses has been realized in several experiments [16–18]. Because it is easier to measure a larger negative expectation value of an entanglement witness, one can argue that the detection of entanglement is easier the larger this negative value is. The expectation value of W also provides lower bounds to various entanglement measures [19, 20]. Estimation of entanglement entropy by measurement of observables has been considered in [21]. Finally, it is interesting to note that violation of Bell inequalities can be rewritten in terms of nonoptimal entanglement witnesses [15, 22].

In general, classification of entanglement witnesses is a hard problem. However, much simpler is the issue with the so-called decomposable entanglement witnesses (D-EW) [23]. D-EW is a witness which can be decomposed as

$$W = P + Q^{\mathrm{T_B}}, \qquad P, Q \geqslant 0, \tag{8}$$

that is, with the positive semidefinite operators P, Q. D-EW can detect only non-PPT entangled states, i.e., those with negative eigenvalues of $\rho^{\mathrm{T_B}}$. They are therefore equivalent to the PPT criterion but closer to experimental implementations, as full tomographic knowledge about the state is not needed. In the present paper, we are going to limit ourselves only to D-EW. General non-decomposable entanglement witness (ND-EW) can be written in a canonical form as $W = P + Q^{\mathrm{T_B}} - \epsilon \mathbb{1}$ [23, 24] and can also detect entangled states with PPT. Finding an optimal ND-EW, for which violation of positivity is maximal, is in general hard [23].

3. Unknown random state

In this section, we assume that the random state $|\psi\rangle$ whose entanglement we would like to detect is unknown so that we are not able to use an optimal W for a particular $|\psi\rangle$. The best one can do is to choose some fixed entanglement witness W in advance, independently of the state. Since we will be interested in the average behaviour over a unitary invariant ensemble of pure random states, W can be chosen to be random as well. That is, in the present section we are going to study detection of entanglement with a random entanglement witness, whose precise definition will be given later. What we want to calculate is the distribution of the expectation values $\langle\psi|W|\psi\rangle$ for a fixed W and an ensemble of random pure states $|\psi\rangle$. Averaging over random states $|\psi\rangle$ we see that the average expectation value $\overline{\langle\psi|W|\psi\rangle}$ is

$$\overline{\langle\psi|W|\psi\rangle} = \int \mathrm{d}\mathcal{P}\langle\psi|W|\psi\rangle = \mathrm{tr}\, W/N^2, \tag{9}$$

[5] Introducing an orthonormal basis $\{|i\alpha\rangle = |i\rangle_{\mathrm{A}} \otimes |\alpha\rangle_{\mathrm{B}}\}$ in the Hilbert space associated with the bipartite system, the density matrix ρ has matrix elements $\rho_{i\alpha,j\beta} = \langle i\alpha|\rho|j\beta\rangle$. The partial transpose is constructed by taking the transpose of only Latin or only Greek indices (here Latin indices refer to subsystem A and Greek indices to subsystem B). For instance, the partial transpose with respect to subsystem B is given by

$$\rho^{\mathrm{T_B}}_{i\alpha,j\beta} = \rho_{i\beta,j\alpha}. \tag{7}$$

where $\overline{\bullet} = \int \mathrm{d}\mathcal{P}\bullet$ denotes an integration over a unique $U(N^2)$-invariant distribution of pure states $|\psi\rangle$, and we used the fact that for a random state $|\psi\rangle = \sum_i c_i|i\rangle$ we have $\overline{c_i c_j^*} = \delta_{ij}/N^2$. We fix normalization of the entanglement witness W such that $\mathrm{tr}\, W = 1$. Therefore, the average expectation value $\overline{\langle\psi|W|\psi\rangle}$ scales $\propto 1/N^2$. We therefore define the rescaled quantity $w = N^2\langle\psi|W|\psi\rangle$ such that $\overline{w} = 1$, independently of the dimension N. From now on we will focus on the random variable w, and its distribution with density $p(w) = \mathrm{d}\mathcal{P}/\mathrm{d}w$.

Because operator P in D-EW (8) just shifts the expectation value towards positive values, we limit ourselves to D-EW of the form $W = Q^{\mathrm{T_B}}$. Any positive semidefinite operator Q can be written in its eigenbasis in terms of positive eigenvalues d_i, satisfying $\sum_i d_i = 1$, and eigenvectors $|\phi_i\rangle$, hence $W = \sum_i d_i(|\phi_i\rangle\langle\phi_i|)^{\mathrm{T_B}}$. We will first study the case when Q is a simple rank-one projector (subsection 3.1), that is W is given by $W = (|\phi\rangle\langle\phi|)^{\mathrm{T_B}}$. There are two reasons why this is the most important case. First, as we will see below, optimal D-EW is always of such a 'projector' form. Second, the expectation value for a general Q can be written (subsection 3.2) as a sum of expectation values for individual eigenvectors of Q and therefore the probability distribution of w will be a simple convolution of distributions for the case of Q being rank one.

Most of our theoretical results are derived for one mixing component only, $m = 1$ (i.e. for a pure random state), since the general result for arbitrary number m of independent mixing components is obtained by simple convolutions as discussed in subsection 3.3.

3.1. Q is a projector

The entanglement witness is of the form $W = (|\phi\rangle\langle\phi|)^{\mathrm{T_B}}$, with a fixed $|\phi\rangle$, and we would like to calculate the density of probability distribution $p(w)$ of its normalized expectation values w for random pure states. This distribution can depend on the chosen $|\phi\rangle$. First few moments of the density $p(w)$ can be calculated explicitly. We have already seen that $\overline{w} = 1$. For higher moments we can in the leading order (in Hilbert space dimension $1/N^2$) use Gaussian averages and Wick contractions in order to approximate integrals over the unitary group [25]. Using

$$\overline{c_i^* c_j c_k^* c_l} = (\delta_{ij}\delta_{kl} + \delta_{il}\delta_{jk})/N^4 + \mathcal{O}(N^{-6}), \tag{10}$$

and similarly for higher order products, we arrive, up to corrections $\mathcal{O}(N^{-2})$, at

$$\begin{aligned} \kappa_2 &= \overline{(w-\overline{w})^2} = \mathrm{tr}\, W^2 = 1, \\ \kappa_3 &= \overline{(w-\overline{w})^3} = 2\,\mathrm{tr}\, W^3, \\ \kappa_4 &= \overline{(w-\overline{w})^4} - 3\kappa_2^2 = 6\,\mathrm{tr}\, W^4. \end{aligned} \tag{11}$$

Here we denote by κ_n the nth cumulant. We see that the average value of w as well as the width of the distribution is 1, independently of the state $|\phi\rangle$ we use for W. While $\mathrm{tr}\, W = 1$ is a simple normalization, second moment is 1 due to Q being of rank 1. If only the first two cumulants were nonzero our probability density $p(w)$ would be a simple Gaussian.

Let us first see what happens if we choose for $|\phi\rangle$ a state with a large Schmidt rank r (number of nonzero eigenvalues of the reduced density matrix $\sigma_{\mathrm{A}} = \mathrm{tr_B}|\phi\rangle\langle\phi|$) of order $r \sim N$. Using Schmidt decomposition (5) for the state $|\phi\rangle = \sum_i \mu_i|a_i\rangle \otimes |b_i\rangle$ we can immediately write eigenvalues and eigenvectors of the Hermitian operator $W = (|\phi\rangle\langle\phi|)^{\mathrm{T_B}}$. There are $N(N-1)$ eigenvalues $\pm\mu_i\mu_j, i < j$, with the corresponding eigenvectors $(|a_i b_j^*\rangle \pm |a_j b_i^*\rangle)/\sqrt{2}$ and N eigenvalues μ_i^2 with the corresponding eigenvectors $|a_i b_i^*\rangle$. In our notation, $|a_i b_j^*\rangle = |a_i\rangle \otimes |b_j^*\rangle$, where $|b_i^*\rangle = \sum_\alpha b_{i\alpha}^*|\alpha\rangle$ if $|b_i\rangle = \sum_\alpha b_{i\alpha}|\alpha\rangle$. Using these eigenvalues one can see that the traces of powers of W are

$$\mathrm{tr}\, W^{2k} = \left(\mathrm{tr}\,\sigma_{\mathrm{A}}^k\right)^2, \qquad \mathrm{tr}\, W^{2k+1} = \mathrm{tr}\,\sigma_{\mathrm{A}}^{2k+1}. \tag{12}$$

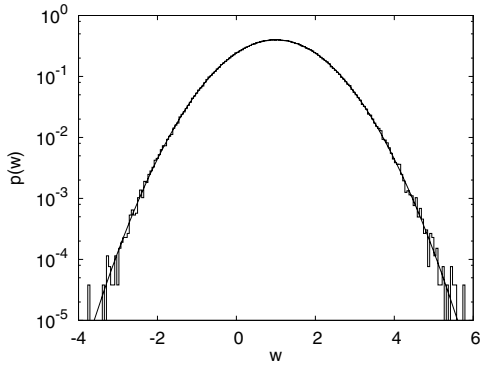

Figure 1. Density of probability distribution of $w = N^2\langle\psi|W|\psi\rangle$ for random vectors $|\psi\rangle$ and a single $W = (|\phi\rangle\langle\phi|)^{T_B}$ with a random $|\phi\rangle$. Total Hilbert space dimension $N^2 = 2^{10}$. The full curve is the theoretical Gaussian prediction (13). We use an ensemble of 5×10^5 random states.

Because eigenvalues of σ_A are of order $\sim 1/r$, we have $\mathrm{tr}\, W^3 \sim 1/r^2 \sim \mathcal{O}(1/N^2)$, $\mathrm{tr}\, W^4 \sim 1/r^2 \sim \mathcal{O}(1/N^2)$ and so on. Note that formulae (11) are exact up to order $1/N^2$ and therefore cannot be used for $\kappa_{n\geqslant 3}$ and states $|\phi\rangle$ with the full Schmidt rank. Nevertheless, we can conclude that the higher cumulants are at most $\sim 1/N^2$ and therefore vanish in the limit $N \to \infty$. Therefore, for W and $|\phi\rangle$ with an increasing rank, the probability density $p(w)$ converges to a Gaussian in the limit $N \to \infty$,

$$p(w) = \frac{1}{\sqrt{2\pi}} \exp(-(w-1)^2/2). \tag{13}$$

The theoretical prediction (13) is compared with the results of numerical simulation in figure 1. The probability of measuring negative w, i.e., of detecting entanglement, is $\int_{-\infty}^{0} p(w)\,\mathrm{d}w$ and therefore

$$\mathcal{P}(w < 0) = (1 - \mathrm{erf}(1/\sqrt{2}))/2 \approx 0.159. \tag{14}$$

Note that this entanglement detection probability is independent of the details of $|\phi\rangle$, provided that its Schmidt rank r is large, more precisely $r \propto N$.

The above Gaussian form of $p(w)$ can be understood also as a consequence of the central limit theorem. Indeed, let us write W in its eigenbasis, then we have, writing $\lambda_i = \mu_i^2$,

$$w = \sum_{i=1}^{r} \lambda_i |\langle\psi|a_i b_i^*\rangle|^2 + \sum_{i<j} \sqrt{\lambda_i\lambda_j}(\langle\psi|a_i b_j^*\rangle\langle a_j b_i^*|\psi\rangle + \langle\psi|a_j b_i^*\rangle\langle a_i b_j^*|\psi\rangle). \tag{15}$$

Denoting overlaps by $\langle\psi|a_i b_j^*\rangle = \sqrt{y_{ij}}\,\mathrm{e}^{\mathrm{i}\varphi_{ij}}$, where y_{ij} and φ_{ij} are two real numbers, we have

$$w = \sum_{i,j=1}^{r} \sqrt{\lambda_i\lambda_j}\sqrt{y_{ij}y_{ji}}\cos(\varphi_{ij} - \varphi_{ji}). \tag{16}$$

For the overlap of two random states we know that the angle φ is distributed uniformly, while the amplitude (scaled by N^2) has an exponential distribution, $p(y_{ij}) = \exp(-y_{ij})$. Therefore, for a given set of eigenvalues λ_i of σ_A, and assuming y_{ij} to be independent, we can calculate the distribution of w using the above formula. It is a convolution of distributions of individual terms. In the limit $r \to \infty$ the central limit theorem can be used, resulting in a Gaussian distribution (13).

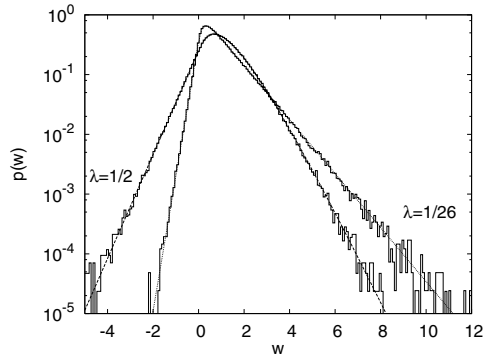

Figure 2. Density of probability distribution of $w = N^2\langle\psi|W|\psi\rangle$ for random vectors $|\psi\rangle$ and a single $W = |\phi\rangle\langle\phi|^{T_B}$ with $|\phi\rangle$ having two nonzero Schmidt coefficients $\sqrt{\lambda}$ and $\sqrt{1-\lambda}$. We show two cases, $\lambda = 1/2$ and $\lambda = 1/26$, both for $N^2 = 2^{10}$. The dotted curve is the theoretical prediction (17) for $\lambda = 1/26$ while the dashed one is for $\lambda = 1/2$, see equation (18). We use an ensemble of 5×10^5 random states.

For $|\phi\rangle$ with a small finite Schmidt rank there will be deviations from Gaussian because higher cumulants of $p(w)$ (11) will be in general nonzero also in the limit $N \to \infty$. The above form for w (16) is actually very handy for an explicit calculation of $p(w)$ for $|\phi\rangle$ with a small rank r. Let us take the extreme case of rank $r = 2$, where we expect strongest deviations from a Gaussian. We therefore have only two nonzero eigenvalues of σ_A, $\lambda_1 = \lambda$ and $\lambda_2 = 1 - \lambda$. Assuming the overlaps y_{ij} to be independent and exponentially distributed and the angles φ_{ij} to be uniform, we arrive after evaluating few convolutions at

$$p(w) = \begin{cases} \dfrac{1}{(1-2\lambda)^2}\left\{\lambda\, \mathrm{e}^{-\frac{w}{\lambda}} + (1-\lambda)\, \mathrm{e}^{-\frac{w}{1-\lambda}}\right\} + \dfrac{1}{4\sqrt{\lambda(1-\lambda)} - 2}\, \mathrm{e}^{-\frac{w}{\sqrt{\lambda(1-\lambda)}}} & w > 0, \\ \dfrac{1}{4\sqrt{\lambda(1-\lambda)} + 2}\, \mathrm{e}^{\frac{w}{\sqrt{\lambda(1-\lambda)}}} & w < 0. \end{cases} \tag{17}$$

The distribution is a sum of exponentials. The probability of detecting entanglement, i.e., of measuring negative value of w is $\mathcal{P}(w < 0) = 1/(4 + 2/\sqrt{\lambda(1-\lambda)})$. As a function of λ it reaches a maximum for $\lambda = 1/2$, i.e., both eigenvalues of σ_A are equal, where it is equal to $1/8$ (which is less than for Gaussian distribution (13)). In figure 2 we compare results of numerical simulation for $p(w)$ for two cases: one with $\lambda = 1/2$, for which the appropriate limit of equation (17) gives

$$p(w)_{\lambda=1/2} = \begin{cases} (1 + 4w + 8w^2)\dfrac{\mathrm{e}^{-2w}}{4} & w > 0, \\ \dfrac{1}{4}\, \mathrm{e}^{2w} & w < 0, \end{cases} \tag{18}$$

and the case with an almost pure σ_A, $\lambda = 1/26$, for which the probability of detecting entanglement is $5/72 \approx 0.07$. Note that in the limit $\lambda \to 0$, i.e., of a pure product state for $|\phi\rangle$, w will always be positive with an exponential distribution.

3.2. Q of higher rank

So far we have discussed only the case when Q is a one-dimensional projector, $Q = |\phi\rangle\langle\phi|$. What happens if the rank of Q is larger? If Q is of rank 2, W can be written as

$W = d_1 W_1 + d_2 W_2$ with positive $d_{1,2}$ and $W_{1,2} = (|\phi_{1,2}\rangle\langle\phi_{1,2}|)^{\mathrm{T_B}}$. If we assume W_1 and W_2 are statistically independent, so that $w_1 = N^2\langle\psi|W_1|\psi\rangle$ and $w_2 = N^2\langle\psi|W_2|\psi\rangle$ are uncorrelated, then the distribution of $w = d_1 w_1 + d_2 w_2$ is given by a convolution of distributions for $w_{1,2}$. Let us calculate the second moment of w given by the above sum. Using

$$N^2 \int \mathrm{d}\mathcal{P}\langle\psi|W_1|\psi\rangle\langle\psi|W_2|\psi\rangle = \mathrm{tr}(W_1 W_2) + \mathrm{tr}\, W_1\, \mathrm{tr}\, W_2 + \mathcal{O}(1/N), \tag{19}$$

and

$$\mathrm{tr}(W_1 W_2) = |\langle\phi_1|\phi_2\rangle|^2, \tag{20}$$

we get

$$\overline{(w - \overline{w})^2} = d_1^2 + d_2^2 + 2d_1 d_2 |\langle\phi_1|\phi_2\rangle|^2. \tag{21}$$

We see that the width of the distribution of w is the same as in the case of convolution of two independent distributions, leading to a width $d_1^2 + d_2^2$, provided the two vectors $|\phi_1\rangle$ and $|\phi_2\rangle$ are orthogonal. Using similar considerations for higher moments we can conclude that in the case of $Q = \sum_i^k d_i |\phi_i\rangle\langle\phi_i|$ having rank k, and because eigenvectors of Q are orthogonal, the distribution of w is Gaussian of width $\sigma^2 = \sum_i^k d_i^2$, at least for sufficiently large k. Assuming for simplicity that all d_i are the same, $d_i = 1/k$, we obtain

$$p(w) = \sqrt{\frac{k}{2\pi}}\, \mathrm{e}^{-k(w-1)^2/2}. \tag{22}$$

3.3. Mixtures of random states

It is also interesting to consider the case in which Q is of rank 1 but we wish to detect the entanglement of mixed states, for instance of a mixture of m pure random states, as given in equation (2). In such a case the expectation value is $w = \frac{1}{m}\sum_i \mathrm{tr}(W|\psi_i\rangle\langle\psi_i|)$, with $W = (|\phi\rangle\langle\phi|)^{\mathrm{T_B}}$. Repeating the same derivation as for Q of higher rank and pure states, just replacing k by m, we see that the distribution of w will be a convolution of distributions for individual $|\psi_i\rangle$ and, provided these are statistically independent, the width of the resulting Gaussian (22) will be $1/m$. Because the probability density $p(w)$ becomes narrowly peaked about its mean $\overline{w} = 1$ with increasing m, the probability of measuring negative values decreases with m; namely for the Gaussian form of equation (22) we have

$$\mathcal{P}(w < 0) = \frac{1 - \mathrm{erf}(\sqrt{m/2})}{2} \asymp \frac{1}{\sqrt{2\pi m}}\, \mathrm{e}^{-m/2}. \tag{23}$$

This probability decays to zero exponentially with m. Therefore, the detection of entanglement for a mixture of random states is very hard.

4. Known random state

In this section we assume that the random state $|\psi\rangle$ whose entanglement we want to measure is known in advance and furthermore, that we are able to prepare an arbitrary D-EW. In addition, we have to assume that our state $|\psi\rangle$ is neither separable, nor bound entangled, which is true with probability which converges to one exponentially in N. Therefore, for each $|\psi\rangle$ we can prepare an optimal entanglement witness, such that its expectation value will be minimal. As far as D-EW are concerned, the optimal choice of $W = W_{\mathrm{opt}}$ is to take for Q a projector to the eigenspace corresponding to the minimal (negative) eigenvalue λ_{min} of $\rho^{\mathrm{T_B}}$, $W_{\mathrm{opt}} = (|\phi_{\mathrm{min}}\rangle\langle\phi_{\mathrm{min}}|)^{\mathrm{T_B}}$. The maximal violation of positivity is therefore

$$\mathrm{tr}(W_{\mathrm{opt}}\rho) = -|\lambda_{\mathrm{min}}(\rho^{\mathrm{T_B}})|. \tag{24}$$

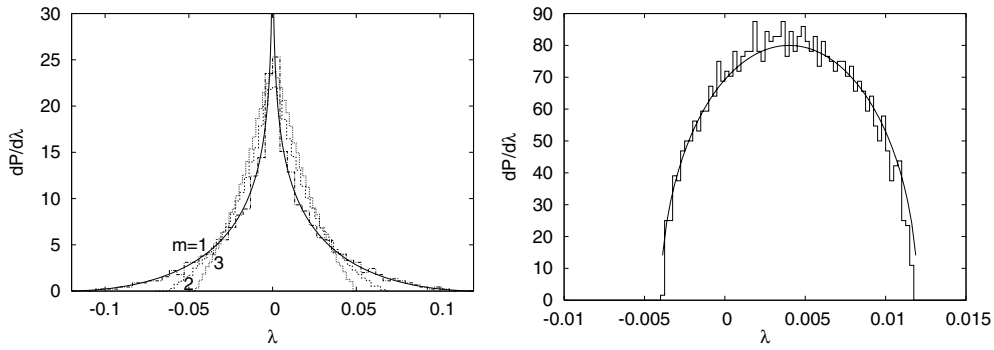

Figure 3. Distribution of eigenvalues of ρ^{T_B}, equation (2), for mixing few states, $m = 1, 2, 3$ (left plot), all for $N^2 = 2^{10}$. The theoretical formula for $m = 1$ (27) is shown with the full curve. The right frame shows the distribution after mixing many states, $m = 256$, $N^2 = 2^8$. The full curve is a semicircle with the centre at $1/N^2$. All histograms are averages over an ensemble of 10 mixtures ρ (2).

If we are able to measure entanglement witness with a given precision[6] it is the size of $\lambda_{\min}$ which determines the difficulty of detecting entanglement in $|\psi\rangle$. Note that the optimal entanglement witness W_{opt} depends on the state $|\psi\rangle$. For each state $|\psi\rangle$ we have to pick a different W_{opt}.

4.1. Distribution of eigenvalues of ρ^{T_B}

First, let us look at the distribution of eigenvalues after PT operation for a single random pure state. The eigenvalues can be written in terms of Schmidt coefficients μ_i as $\lambda = \pm\mu_i\mu_j$. Formally, we can write the distribution of λ, for $\lambda > 0$, as

$$\frac{d\mathcal{P}}{d\lambda} = \int d\mu_i \int d\mu_j \delta(\lambda - \mu_i\mu_j) \frac{d\mathcal{P}}{d\mu_i}\frac{d\mathcal{P}}{d\mu_j} \tag{25}$$

where distribution of μ_i can be, for large N, obtained from the Marčenko–Pastur law [26] for the distribution of $\tau = N\mu_i^2$, namely

$$d\mathcal{P}/d\tau = \sqrt{\tau(4-\tau)}/(2\pi\tau), \tag{26}$$

by a simple change of variables $\tau \to \mu_i$. The result for the distribution of scaled eigenvalues $y = N\lambda$ reads

$$\frac{d\mathcal{P}}{dy} = \frac{1}{8\pi^2}[(16+y^2)K(1-y^2/16) - 32E(1-y^2/16)], \qquad y \in [-4, 4], \tag{27}$$

where $E(x)$ and $K(x)$ are elliptic integrals. Note that this distribution takes into account only the $N(N-1)$ 'off-diagonal' eigenvalues $\pm\mu_i\mu_j$ with $i < j$. 'Diagonal' eigenvalues μ_i^2 have the same distribution as the eigenvalues of ρ_A and are only N in number. The expectation value of the minimal eigenvalue equals $\overline{\lambda}_{\min} = -4/N$. In fact, the distribution of $\lambda_{\min}$ becomes strongly peaked around $-4/N$ with diminishing relative fluctuations as $N \to \infty$.

When we mix several independent (in general non-orthogonal) random vectors, $\rho = \sum_i^m |\psi_i\rangle\langle\psi_i|/m$, the minimal eigenvalue $\lambda_{\min}$ increases and the distribution becomes increasingly sharply peaked (for $m \to \infty$ we get $\rho \to \mathbb{1}/N^2$ with all eigenvalues being equal to $1/N^2$). We numerically verified the above theoretical prediction for $d\mathcal{P}/dy$ (27) in figure 3. We can also see that after mixing many random states ($m \sim N$) the distribution becomes a semicircle, which is a numerical result for which we have yet no analytical explanation.

[6] Different normalization of W, e.g., $\operatorname{tr} W = f(N)$, would result in the maximal violation $-f(N)|\lambda_{\min}|$.

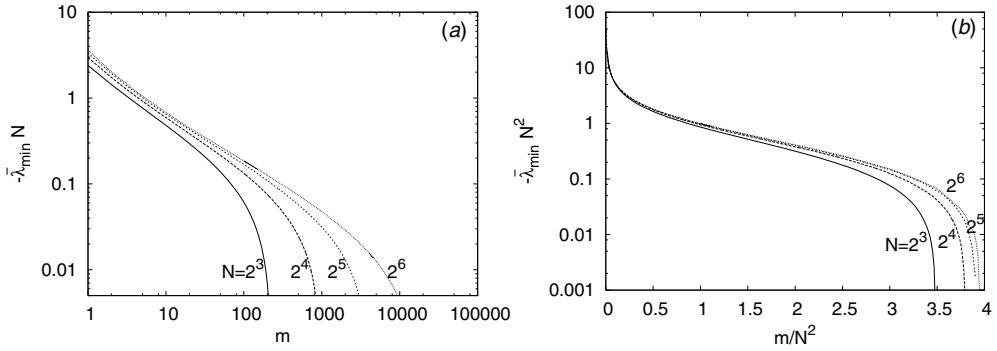

Figure 4. Dependence of $\bar{\lambda}_{\min}$ on m. Overlapping of curves for different dimension N signals the scaling $\lambda_{\min} \sim -1/N$ for small m in the left frame (a) and $\lambda_{\min} \sim -1/N^2$ for large m (but smaller than m^* where $\bar{\lambda}_{\min}$ changes sign) in the right frame (b). We average over 1000, 100, 10 and 6 mixed random states (2), for $N^2 = 2^6, 2^8, 2^{10}$ and 2^{12}, respectively.

4.2. Minimal eigenvalue of ρ^{T_B}

Because $\lambda_{\min}$ determines the maximal violation for an optimal witness W (24) we are going to look in more detail at the dependence of $\lambda_{\min}$ on m and N. For a single pure state ($m = 1$) we know that the average minimal eigenvalue is $\bar{\lambda}_{\min} = -4/N$. On the other hand, we also know that for large m, as we approach a completely mixed state the minimal eigenvalue must scale as $\bar{\lambda}_{\min} \sim 1/N^2$. Therefore, the scaling of $\bar{\lambda}_{\min}$ has to change as we increase the number of mixed states m. To confirm this expectation, we performed numerical simulation, calculating the average $\bar{\lambda}_{\min}$ for different m. The results are shown in figure 4. Note that the average minimal eigenvalue $\bar{\lambda}_{\min}$ is positive for $m > m^*$, with $m^* \approx 4N^2$.

Although von Neumann entropy of a random state is large all eigenvalues of ρ^{T_B} are very small and will therefore be hard to detect. If we assume that we are able to measure values of $\mathrm{tr}(W\rho)$ and ϵ with accuracy then we can, depending on the scaling of ϵ with N, tell for which values of m the detection of entanglement is possible. If ϵ does not depend on N, i.e., precision does not increase with N, then for sufficiently large N, such that $4/N < \epsilon$, detection of entanglement will be impossible. Already a single random state becomes from the viewpoint of entanglement detection 'classical', since measuring a negative expectation value of its optimal entanglement witness is below the detection limit. If on the other hand we are able to measure ϵ which decreases as $1/N$, the critical m_{crit}, beyond which the entanglement detection is impossible, will be independent of N, i.e., in the limit $N \to \infty$ the ratio $m_{\mathrm{crit}}/N \to 0$ (see figure 4(a)). If however we are able to detect very small expectation values of order $1/N^2$, then m_{crit} will be proportional to N^2 (see figure 4(b)). Furthermore, even with arbitrary accuracy, detection of entanglement with D-EW is impossible beyond $m = m^* \propto N^2$.

Worth mentioning is that for $m \geqslant N$ the mixed state ρ of equation (2) cannot be used for dense coding. Indeed, a quantum state is useful for dense coding [27] only if $S(\rho_A) - S(\rho) > 0$ (see, e.g., [28]). For mixtures of m random states $S(\rho)$ is roughly equal to $\sim \log m$. On the other hand $S(\rho_A)$ is at most $\log N$. Therefore, $S(\rho_A)$ will be smaller than $S(\rho)$ for $m \sim N$.

5. Conclusions

In this paper we have considered random states and have shown that, while their entanglement content is almost maximal, the detection of such entanglement is very difficult. This is a

consequence of the complexity of a random state, which leads to a large number of small coefficients in the Schmidt decomposition of the state. Nevertheless, for random pure states, a finite success probability in the detection of entanglement exists also in the limit in which the Hilbert space dimension $N^2 \to \infty$. This implies that quantum chaos alone is not sufficient to erase any trace of entanglement in the classical limit, provided that ideal measurements are possible. On the other hand such erasure becomes very efficient when coarse graining is taken into account, for instance when mixtures instead of pure states are considered.

We note that all our results can be straightforwardly generalized to the case of unbalanced bipartition $N = \dim \mathcal{H}_A \neq N' = \dim \mathcal{H}_B$; for instance, distribution of $w = NN'\langle\psi|W|\psi\rangle$ is Gaussian (13) provided both dimensions N, N' are large.

We would like to stress once more that the detection difficulties are a consequence of the complexity of random states. If instead one considers 'regular states' such as the GHZ state $|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}(|0,\ldots,0\rangle + |1,\ldots,1\rangle)$, then the optimal witness is $W_{\mathrm{opt}} = (|\phi_{\mathrm{min}}\rangle\langle\phi_{\mathrm{min}}|)^{\mathrm{T_B}}$, with $|\phi_{\mathrm{min}}\rangle = \frac{1}{\sqrt{2}}(|0,\ldots,0\rangle|1,\ldots,1\rangle - |1,\ldots,1\rangle|0,\ldots,0\rangle)$ which corresponds to the minimal eigenvalue $\lambda_{\mathrm{min}} = -1/2$ of $(|\mathrm{GHZ}\rangle\langle\mathrm{GHZ}|)^{\mathrm{T_B}}$. Since the value of λ_{min} is $-1/2$ instead of $-4/N$ as for a random state, it turns out that it will be much easier to detect entanglement in a 'regular' rather than in a random state. This happens in spite of the fact that the entanglement content is larger in a random than in such a regular state.

Acknowledgments

The authors would like to thank Slovenian Research Agency, programme P1-0044, grant J1-7437, and the MIUR-PRIN 2005 (2005025204) for support.

References

[1] Lubkin E 1978 Entropy of an n-system from its correlation with a k-reservoir *J. Math. Phys.* **19** 1028–31
[2] Lloyd S and Pagels H 1988 Complexity as thermodynamic depth *Ann. Phys., NY* **188** 186–213
[3] Page D 1993 Average entropy of a subsystem *Phys. Rev. Lett.* **71** 1291
[4] Plenio M B and Virmani S 2007 An introduction to entanglement measures *Quantum Inf. Comput.* **7** 1
[5] Horodecki R, Horodecki P, Horodecki M and Horodecki K 2007 Quantum entanglement *Preprint* quant-ph/0702225
[6] Stöckmann H-J 1999 *Quantum Chaos: An Introduction* (Cambridge: Cambridge University Press)
[7] Saraceno M 1990 Classical structures in the quantized baker transformation *Ann. Phys., NY* **199** 37
[8] Schack R 1998 Using a quantum computer to investigate quantum chaos *Phys. Rev.* A **57** 1634
[9] Benenti G, Casati G, Montangero S and Shepelyansky D L 2002 Eigenstates of an operating quantum computer: hypersensitivity to static imperfections *Eur. Phys. J.* D **20** 293
[10] Scott A J and Caves C M 2003 Entangling power of the quantum baker's map *J. Phys. A: Math. Gen.* **36** 9553
[11] Weinstein Y S and Hellberg C S 2005 Entanglement generation of nearly random operators *Phys. Rev. Lett.* **95** 030501
[12] Žnidarič M 2007 Entanglement of random vectors *J. Phys. A: Math. Theor.* **40** F105
[13] Peres A 1996 Separability criterion for density matrices *Phys. Rev. Lett.* **77** 1413
[14] Horodecki M, Horodecki P and Horodecki R 1996 Separability of mixed states: necessary and sufficient conditions *Phys. Lett.* A **223** 1
[15] Terhal B 2000 Bell inequalities and the separability criterion *Phys. Lett.* A **271** 319
[16] Bourennane M, Eibl M, Kurtsiefer C, Gaertner S, Weinfurter H, Gühne O, Hyllus P, Bruß D, Lewenstein M and Sanpera A 2004 Experimental detection of multipartite entanglement using witness operators *Phys. Rev. Lett.* **92** 087902
[17] Leibfried D *et al* 2005 Creation of a six-atom 'Schrödinger cat' state *Nature* **438** 639
[18] Häffner H *et al* 2005 Scalable multi-particle entanglement of trapped ions *Nature* **438** 643
[19] Eisert J, Brandao F G S L and Audenaert K M R 2007 Quantitative entanglement witnesses *New J. Phys.* **9** 46

[20] Gühne O, Reimpell M and Werner R F 2007 Estimating entanglement measures in experiments *Phys. Rev. Lett.* **98** 110502
[21] Klich I, Refael G and Silva A 2006 Measuring entanglement entropes in many-body systems *Phys. Rev.* A **74** 032306
[22] Hyllus P, Gühne O, Bruß D and Lewenstein M 2005 Relations between entanglement witnesses and Bell inequalities *Phys. Rev.* A **72** 012321
[23] Lewenstein M, Kraus B, Cirac J I and Horodecki P 2000 Optimization of entanglement witnesses *Phys. Rev.* A **62** 052310
[24] Lewenstein M, Kraus B, Horodecki P and Cirac J I 2001 Characterization of separable states and entanglement witnesses *Phys. Rev.* A **63** 044304
[25] Prosen T, Seligman T H and Weidenmüller H A 2002 Integration over matrix spaces with unique invariant measures *J. Math. Phys.* **43** 5135
[26] Marčenko V A and Pastur L A 1967 Distribution of eigenvalues of some sets of random matrices *Math. USSR-Sb* **1** 457–83
[27] Bennett C H and Wiesner S J 1992 Communication via one- and two-particle operators on Einstein–Podolsky–Rosen states *Phys. Rev. Lett.* **69** 2881
[28] Sen A, Sen U, Lewstein M and Sanpera A 2005 The separability versus entanglement problem *Preprint* quant-ph/0508032