Measuring $\mathrm{Tr}\rho^n$ on Single Copies of ρ Using Random Measurements

S. J. van Enk[1,2] and C. W. J. Beenakker[3]

[1]Department of Physics and Oregon Center for Optics University of Oregon, Eugene, Oregon 97403, USA
[2]Institute for Quantum Information, California Institute of Technology, Pasadena, California 91125, USA
[3]Instituut-Lorentz, Universiteit Leiden, P.O. Box 9506, 2300 RA Leiden, The Netherlands

(Received 5 December 2011; published 16 March 2012)

While it is known that $\mathrm{Tr}\rho^n$ can be measured directly (i.e., without first reconstructing the density matrix) by performing joint measurements on n copies of the same state ρ, it is shown here that random measurements on single copies suffice, too. Averaging over the random measurements directly yields estimates of $\mathrm{Tr}\rho^n$, even when it is not known what measurements were actually performed (so that ρ cannot be reconstructed).

DOI: 10.1103/PhysRevLett.108.110503 PACS numbers: 03.65.Ta, 03.65.Wj, 03.67.−a

The standard textbook quantum measurement of an observable $\hat{O}$ on a given quantum system produces an estimate of the expectation value $\mathrm{Tr}(\rho\hat{O})$, where ρ is the density matrix of the system. This expectation value is linear in ρ. As is well known by now [1–6], nonlinear functions of the density matrix ρ, such as the purity $p_2 = \mathrm{Tr}\rho^2$ and its cousins $p_n = \mathrm{Tr}\rho^n$ for $n > 2$, can be measured directly, too, without first having to reconstruct the whole density matrix. For this direct measurement method to work, one needs n quantum systems that are all in the same state ρ, plus the ability to perform the appropriate joint measurement(s) on those multiple copies.

Here we point out that estimates of the same nonlinear quantities can be obtained from *random* measurements on *single* copies as well. A random measurement can be assumed to be implemented by performing a random unitary rotation on the single copy (possibly including an ancilla which starts off in a standard state), followed by a fixed measurement on the single copy (and possibly on the ancilla). By averaging the measurement results over the random unitaries, one can directly infer estimates of $\mathrm{Tr}\rho^n$ [with $n = 2, \ldots, M$, with M the Hilbert space dimension of the system of interest], without having to reconstruct the density matrix. One point of the averaging procedure is that one does not have to know which random unitaries were in fact applied, and, as a consequence, one cannot reconstruct the density matrix in that case. An example of a random measurement is furnished by intensity measurements of speckle patterns resulting from light (be it two photons, or a single photon, or a coherent laser beam) propagating through a disordered medium [7,8], and in that case the purity p_2 can (and was indeed) inferred directly from those measurements (see also [9]).

There is an important difference between the known direct method and the current random method in what quantity exactly is estimated. Suppose one's source does not produce the same state every single time but instead a state ρ_j at try j. In this case, standard quantum measurements of a given observable on J instances $j = 1, \ldots, J$ can still be described by a single density matrix, namely, the mean $\bar{\rho} = \sum_j \rho_j / J$. Since the random method involves only measurements on single copies, it produces, likewise, an estimate of $\mathrm{Tr}(\bar{\rho}^n)$. This requires no assumption about the quantum systems being uncorrelated or unentangled with each other, since ρ_j is obtained by tracing out all degrees of freedom except those of system j.

On the other hand, a direct measurement would yield an estimate of $\mathrm{Tr}(\hat{S}\,\overline{\rho_{j,j+1,\ldots,j+n-1}})$ instead, where $\rho_{j,j+1,\ldots,j+n-1}$ is the joint density matrix of n systems j, $j+1, \ldots, j+n-1$ and $\hat{S}$ is the cyclical shift operator, which acts on the basis states of the n quantum systems as $\hat{S}|\psi_j\rangle|\psi_{j+1}\rangle\cdots|\psi_{j+n-1}\rangle = |\psi_{j+1}\rangle|\psi_{j+2}\rangle\cdots|\psi_j\rangle$. It is only under the assumption that the states of the n systems are identical and independent (i.i.) that the direct measurement yields $\mathrm{Tr}\rho^n$. In fact, the direct measurement is eminently suited for detecting that the states are *not* identical [10]. Although the assumption of i.i. states is standard, it is only recently that precise conditions have been stated under which the approximate i.i. character can be inferred [11]. The required permutation invariance is easily enforced when performing measurements on single copies but not when performing joint measurements on multiple copies [12]. Avoiding this difficulty is the main advantage of the random method.

An $N \times N$ random unitary matrix, distributed according to the Haar measure, can be easily constructed by the method presented in Ref. [13]. One first constructs a matrix whose elements are independent complex Gaussian variables, and one then performs an orthogonalization of the resulting random matrix (where one small pitfall needs to be avoided [13]). We first consider approximate results for random unitaries, because the resulting expressions are quite simple, and subsequently we will give the more involved exact results.

If we consider an arbitrary submatrix V (of size M) of U (of size N), with $M \ll N$ [14], then the real and imaginary parts of its matrix elements can still be very well approximated by independent and normally distributed numbers if

N is large. With this Gaussian approximation, we can compute the following *averages* (we indicate averages over the distribution of random unitaries by $\langle\cdot\rangle$): First, we have $\langle V_{kl}V^*_{mn}\rangle = \delta_{km}\delta_{ln}/N$. Here and in all of the following, we assume we have picked some basis $\{|k\rangle\}$, and we write all matrix elements with respect to that basis. The normalization factor $1/N$ follows immediately from the fact that U, of which V is a submatrix, is unitary, so that $\sum_{l=1}^{N} U_{kl}U^*_{ml} = \delta_{km}$. Higher-order averages follow from the Isserlis (''Gaussian-moments'') theorem [15]. In particular, the only nonzero averages arise from products of $2K$ factors of the form

$$\langle V_{k_1 l_1}\cdots V_{k_K l_K} V^*_{m_1 n_1}\cdots V^*_{m_K n_K}\rangle = \frac{\sum_{\text{all pairs}(i,j)} \delta_{k_i m_j}\delta_{l_i n_j}}{N^K}. \quad (1)$$

We now apply the preceding approximate results to the following scenario. Consider an ''input'' density matrix ρ^{in} of size $M \times M$. Embed the system in a larger Hilbert space of size N, by constructing a new $N \times N$ density matrix by adding zero matrix elements. Then apply a random unitary U to the larger matrix. Finally, consider measurements in a fixed M-dimensional (sub)basis $\{|k\rangle\}$. The probability $\text{Prob}(k)$ of finding measurement outcome k is given by

$$\text{Prob}\,(k) = \sum_{m,n} \rho^{\text{in}}_{mn} V_{mk} V^*_{nk}. \quad (2)$$

This expectation value depends on what V is, of course, but its average is given simply by

$$\langle \text{Prob}(k)\rangle = \sum_{m,n} \rho^{\text{in}}_{mn}\langle V_{mk}V^*_{nk}\rangle = 1/N, \quad (3)$$

where we used the fact that $\text{Tr}(\rho^{\text{in}}) = \sum_m \rho^{\text{in}}_{mm} = 1$. Defining $P_n(k) = \langle \text{Prob}(k)^n\rangle$, the following averages are obtained by using the Isserlis theorem (up to order $n = 4$; subsequent orders can be easily obtained, too, but for our purposes this will do):

$$P_2(k) = [1 + p_2]/N^2, \quad (4a)$$
$$P_3(k) = [1 + 3p_2 + 2p_3]/N^3, \quad (4b)$$
$$P_4(k) = [1 + 3p_2^2 + 6p_2 + 8p_3 + 6p_4]/N^4, \quad (4c)$$

where we defined $p_n = \text{Tr}[(\rho^{\text{in}})^n]$. Inverting these equations gives estimates of p_n in terms of the measurable quantities on the left-hand sides. We denote those estimates by an overbar, e.g., $\bar{p}_2 = N^2 P_2(k) - 1$. We refrain from giving the other inverse relations now, as we will give the *exact* relations below in (8).

We can also compute standard deviations in the (mean) estimates. For example, assuming we average the results for one value of k over N_{rand} random unitaries, then the statistical error in the estimate of the purity is

$$\sqrt{N_{\text{rand}} - 1}\,\Delta(\bar{p}_2) = \sqrt{4p_2 + 2p_2^2 + 8p_3 + 6p_4}. \quad (5)$$

This is an increasing function of p_2, p_3, and p_4, so that the variance is largest for a pure state and smallest for the totally mixed state $\rho^{\text{in}} = \mathbb{1}/M$.

In an actual experiment, one may not know exactly what the values of N and/or M are (for instance, this is the case in the speckle experiments of Refs. [7,8]). In such a case, N can be directly estimated from $P_1(k)$ through $N = 1/P_1(k)$. So, we would use

$$\tilde{p}_2 = P_2(k)/P_1(k)^2 - 1, \quad (6)$$

instead (such estimates we indicate by a tilde). Now this estimate $\tilde{p}_2$ has a *smaller* variance than $\bar{p}_2$ has, simply because the errors in $P_1(k)$ and $P_2(k)$ are positively correlated. It is, therefore, better to use $\tilde{p}_2$ as estimate for p_2, even when N is in principle known. The numerical results given below will confirm this, also for the exact result for $\tilde{p}_2$. For the estimates $\tilde{p}_3$ and $\tilde{p}_4$, however, there is not much difference between the two methods.

When N is not very large, Eqs. (1) and hence (4) are not correct. The exact results, which can be extracted from Refs. [16,17], are still given by (4) upon multiplication of $P_n(k)$ by the correction factor C_n, where

$$C_n = (1 + 1/N)(1 + 2/N)\cdots[1 + (n-1)/N]. \quad (7)$$

Note that these factors depend only on N, not on M, and the results are valid even when $M = N$. This then leads to the inverse formulas:

$$\bar{p}_2 = D_2 P_2(k) - 1, \quad (8a)$$
$$\bar{p}_3 = \tfrac{1}{2}[D_3 P_3(k) - 1 - 3p_2], \quad (8b)$$
$$\bar{p}_4 = \tfrac{1}{6}[D_4 P_4(k) - 1 - 3p_2^2 - 6p_2 - 8p_3], \quad (8c)$$

with $D_n = (N + n - 1)!/(N - 1)!$. Taking into account the correction factors (7) leads to different values for the statistical errors in estimates. It is still true that pure states lead to the largest errors; for those we get

$$\sqrt{N_{\text{rand}} - 1}\,\Delta(\bar{p}_2) = \sqrt{\frac{24(1 + 1/N)}{(1 + 2/N)(1 + 3/N)} - 4}. \quad (9)$$

The right-hand side (slowly) increases with increasing N, from $\sqrt{52/7}$ for $N = 4$ to $\sqrt{20}$ for $N \to \infty$.

In order to illustrate the method and the meanings of N and M, we consider the following examples here. (i) Suppose we have a single photon occupying one of M input modes. We then apply a random linear optics transformation that involves $N - M$ ancilla modes. The photon now ends up being coherently distributed over N output modes. We then estimate the probability $\text{Prob}(k)$ with which the photon ends up in one of a fixed set of M output modes $k = 1, \ldots, M$. This is an example akin to that considered in Refs. [7,8].

(ii) Suppose our system of interest consists of 2 qubits, so that $M = 4$. Suppose we have an ancilla qubit in a fixed state $|0\rangle$, and we apply a random unitary operation to the 3 qubits. In this case, $N = 8$. We then perform measurements

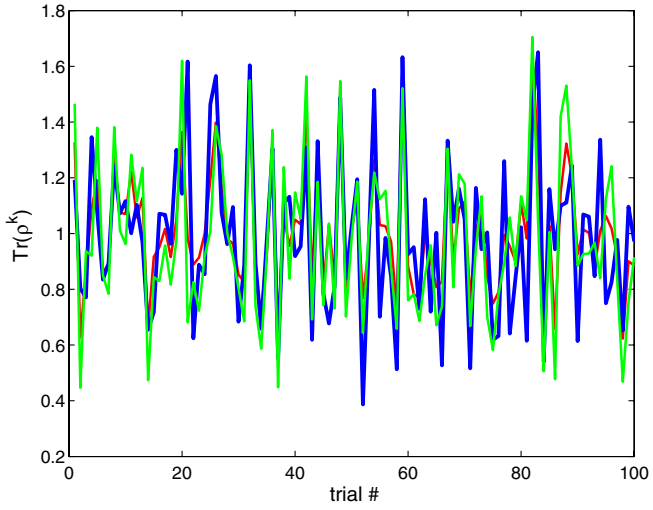

FIG. 1 (color online). This plot shows, for a pure two-qubit state, the estimated values of p_2, p_3, and p_4 (blue, p_2; red, p_3; green, p_4) for 100 trials, each trial using just one value of k, containing an average over $N_{\text{rand}} = 100$ random unitaries, and using $N = 4$ in (8). The mean standard deviations (over 100 trials) were $\Delta\bar{p}_2 = 0.282$ [note that this agrees with the result (9), since $\sqrt{(52/7)/99} \approx 0.274$], $\Delta\bar{p}_3 = 0.21$, $\Delta\bar{p}_2 = 0.29$. The mean estimates obtained by pooling all data from the 100 trials for $\bar{p}_n$ are $\bar{p}_2 = 0.990$, $\bar{p}_3 = 1.01$, and $\bar{p}_4 = 1.02$, which are all consistent with their mean standard deviations (10 times smaller than the $\Delta\bar{p}_n$ given above).

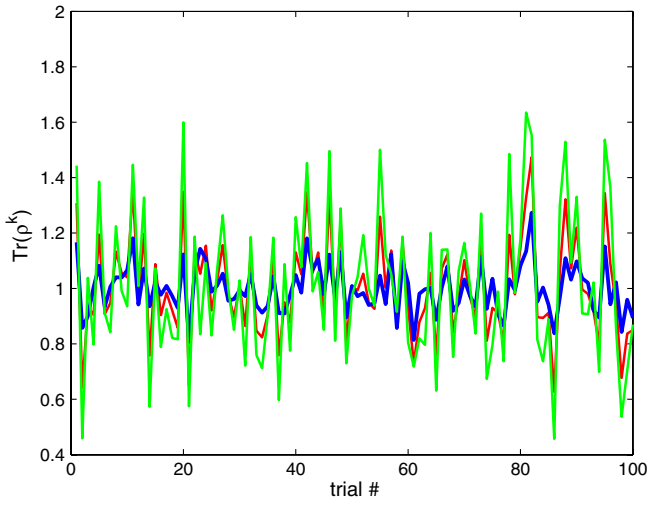

FIG. 2 (color online). The same as the previous figure but using the estimate $N \approx 1/P_1(k)$ in (8). Here we have $\Delta\tilde{p}_2 = 0.09$, $\Delta\tilde{p}_3 = 0.17$, and $\Delta\tilde{p}_4 = 0.27$. The mean estimates obtained from pooling all data (which are the *same* "raw" data as in Fig. 1) from the 100 trials for $\tilde{p}_n$ are $\tilde{p}_2 = 1.005$ [which is indeed better than $\bar{p}_2$], $\tilde{p}_3 = 1.01$, and $\tilde{p}_4 = 1.02$, all consistent with the statistical errors in the mean (which are 10 times smaller than $\Delta\tilde{p}_n$).

on each of the three qubits separately in the standard basis. We measure the probability Prob(k) of the two qubits ending up in one of the $M = 4$ combinations $k = 00, 01, 10, 11$ *and* the ancilla ending up in $|0\rangle$ (thus measuring only an M-dimensional subspace).

(ii′) There is no need for any ancillas if dealing with a fixed and known number of qubits, say, Q. In that case, we simply have $N = M = 2^Q$. We consider only case (ii′) in the following numerical results.

We assume that we run an experiment with a fixed random ("unknown") unitary of size N sufficiently many times that we get a very good estimate of Prob(k) for each k for the given unitary and the given input state (of size M). Subsequently, we average over N_{rand} random unitaries to obtain $P_n(k) = \langle \text{Prob}(k)^n \rangle$. From those results we estimate the values of p_2, p_3, and p_4.

The first example we consider corresponds to case (ii′) mentioned above, where we have two qubits. In Figs. 1 and 2, we plot results for pure input states, where we use the results for just 1 value of k to estimate p_n, in two different ways: using the exact value $N = 4$ (Fig. 1) or using the estimate $N \approx 1/\langle \text{Prob}(k) \rangle$ (Fig. 2). The results show how the latter method is more accurate for estimating purity. The *same* data are used in the two figures, so that all differences between them are entirely due to the different analysis of those data. This different analysis reduces the statistical variation in $\tilde{p}_2$ but not in $\tilde{p}_3$ and $\tilde{p}_4$. In addition, the plots show that the statistical errors in $\tilde{p}_2$, $\tilde{p}_3$, and $\tilde{p}_4$ are strongly correlated in the latter case.

In the remaining figures, we perform an additional average over the M different values of k, leading to smaller (by a factor of about $\sqrt{M}$) error bars.

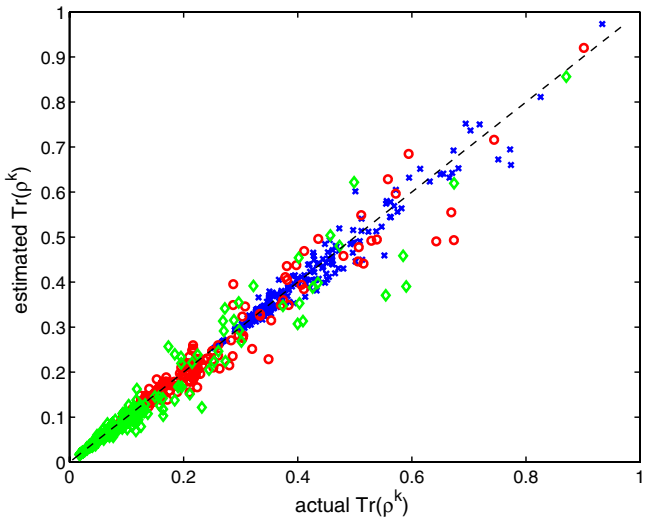

FIG. 3 (color online). Scatter plot of estimated values of p_2 (blue crosses), p_3 (red circles), and p_4 (green diamonds) versus their actual values for 200 randomly picked two-qubit input states. Here an average is taken over $N_{\text{rand}} = 30$ random unitaries, as well as over 4 measurement outcomes. For convenience, the dashed line gives the diagonal on which estimated and actual values agree.

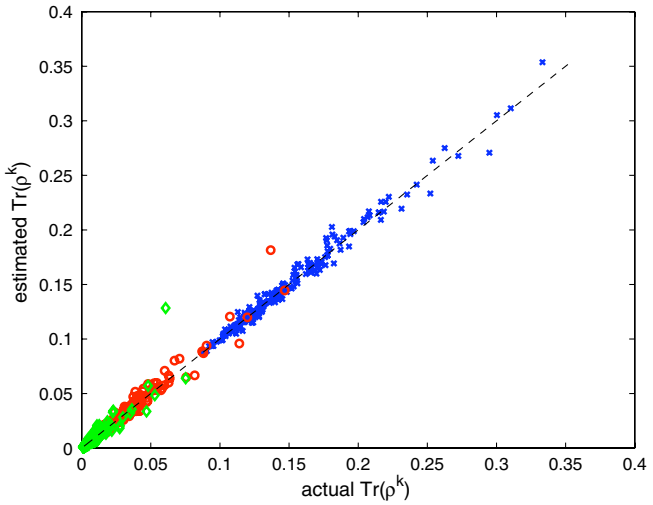

FIG. 4 (color online). The same as Fig. 3 but for randomly picked *five*-qubit states ($M = N = 32$). Averaging over the same number of random unitaries (here $N_{\text{rand}} = 30$) produces a smaller statistical error for larger systems.

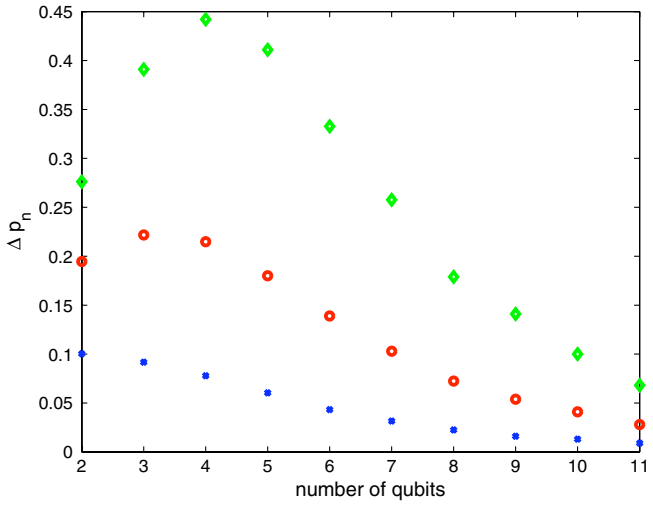

FIG. 5 (color online). The standard deviation in the mean estimates for $\tilde{p}_2$ (blue crosses), $\tilde{p}_3$ (red circles), and $\tilde{p}_4$ (green diamonds) after averaging over $N_{\text{rand}} = 30$ random unitaries, for pure multiqubit states, as a function of the number of qubits.

Performing tomography on two qubits would require 15 independent (and known) measurements. Here we show that with just a moderate overhead one can obtain good estimates of p_2, p_3, and p_4 for *generic* (i.e., randomly picked [18]) states.

In Fig. 3, results are displayed for 200 generic two-qubit states, using $N_{\text{rand}} = 30$.

In Fig. 4, we show (for five qubits) that the number of random unitaries needed to obtain a fixed-size error bar does not increase with the number of qubits. For $N_{\text{rand}} = 30$, one still obtains good estimates: In fact, the error bars *decrease* (roughly as $1/\sqrt{M}$) when going to more and more qubits, just because the number M of measurement results one can average over increases exponentially with the number of qubits, while the variance (9) increases only very slowly. This is illustrated for pure multiqubit states in Fig. 5. It shows that the statistical error in the estimate of $\text{Tr}\rho^n$ for $n = 2, 3, 4$ first increases with the number of qubits before (at $\geq n$ qubits) it starts to decrease monotonically.

In conclusion then, using the ideas of random matrix theory, we showed that nonlinear functions of the density matrix such as $\text{Tr}\rho^n$ can be directly obtained from appropriately averaged random measurements on single copies. No assumptions are needed on the independence of the copies nor on their states being identical. This contrasts the random method with so-called direct measurements on n identical copies [1–6].

Moreover, one does not need to know which random measurements were actually performed, because the averaging procedure keeps all information about the eigenvalues of ρ, which is all that is needed to estimate $\text{Tr}\rho^n$. One does need to verify that the random unitaries have been drawn from the appropriate ensemble. There are two tests one could perform: First of all, the definition of the ensemble is that it is unitarily invariant. This means in our context that all averages $\langle \text{Prob}(k)^n \rangle$ should be independent of k. This is a statistically testable property. In addition, one can apply the random measurements to *known* input states, so that the values of those k-independent averages are known.

Importantly, the number of unitaries over which one has to average in order to obtain a fixed error bar in the estimates of $\text{Tr}\rho^n$ scales very favorably with the Hilbert space dimension of one's system: In fact, this number even tends to *decrease*. For two qubits this amounts to needing a small overhead as compared to full quantum-state tomography, but for larger systems (more than, say, four qubits) the random method requires (far) fewer resources than does full quantum-state tomography.

One can implement our method on multiple (say, ≥ 4) ions in an ion trap [19,20], for instance, by applying fixed two-qubit gates to randomly picked pairs of ions, interspersed with random single-qubit gates, or on the transverse spatial degrees of freedom of either single photons or photon pairs, as in Ref. [8].

[1] R. Filip, Phys. Rev. A **65**, 062320 (2002).
[2] P. Horodecki and A. Ekert, Phys. Rev. Lett. **89**, 127902 (2002).
[3] P. Horodecki, Phys. Rev. Lett. **90**, 167901 (2003).
[4] H. A. Carteret, arXiv:quant-ph/0309212.
[5] T. A. Brun, Quantum Inf. Comput. **4**, 401 (2004).
[6] F. Mintert and A. Buchleitner, Phys. Rev. Lett. **98**, 140505 (2007).
[7] C. W. J. Beenakker, J. W. F. Venderbos, and M. P. van Exter, Phys. Rev. Lett. **102**, 193601 (2009).

[8] W. H. Peeters, J. J. D. Moerman, and M. P. van Exter, Phys. Rev. Lett. **104**, 173601 (2010).
[9] A different case is the experiment of M. Munroe *et al.*, Phys. Rev. A **52**, R924 (1995), where diagonal density matrix elements in the photon-number basis (hard to measure directly) were obtained by phase-averaging more straightforward quadrature measurements.
[10] L. Schwarz and S. J. van Enk, Phys. Rev. Lett. **106**, 180501 (2011).
[11] R. Renner, Nature Phys. **3**, 645 (2007).
[12] S. J. van Enk, Phys. Rev. Lett. **102**, 190503 (2009).
[13] F. Mezzadri, Not. Am. Math. Soc. **54**, 592 (2007).
[14] K. Zyczkowski and H.-J. Sommers, J. Phys. A **33**, 2045 (2000).
[15] L. Isserlis, Biometrika **12**, 134 (1918).
[16] B. Collins, Int. Math. Res. Not. **2003**, 953 (2003).
[17] Z. Puchala and J. A. Miszczak, arXiv:1109.4244.
[18] As only the eigenvalues of ρ matter, the states were chosen according to a simple distribution, without significance otherwise: First, M uniformly distributed random numbers (z_i) between 0 and 1 are picked; then ρ is chosen as the diagonal matrix $\mathrm{diag}(z_i^E)/(\sum_i z_i^E)$, with $E = 2$ in Fig. 3 and $E = 8$ in Fig. 4. These choices produce a wider spread of values for $\mathrm{Tr}\rho^n$ than do standard ensembles.
[19] H. Haeffner *et al.*, Nature (London) **438**, 643 (2005).
[20] D. Leibfried *et al.*, Nature (London) **438**, 639 (2005).