

Permutationally Invariant Quantum Tomography - Supplementary Material

Géza Tóth,^{1,2,3} Witłef Wieczorek,^{4,5,*} David Gross,⁶ Roland Krischek,^{4,5} Christian Schwemmer,^{4,5} and Harald Weinfurter^{4,5}

¹*Department of Theoretical Physics, The University of the Basque Country, P.O. Box 644, E-48080 Bilbao, Spain*

²*IKERBASQUE, Basque Foundation for Science, E-48011 Bilbao, Spain*

³*Research Institute for Solid State Physics and Optics,*

Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary

⁴*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany*

⁵*Fakultät für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany*

⁶*Institute for Theoretical Physics, Leibniz University Hannover, D-30167 Hannover, Germany*

(Dated: December 23, 2010)

The supplement contains some derivations to help to understand the details of the proofs of the main text. It also contains some additional experimental results.

PACS numbers: 03.65.Wj, 42.50.Dv, 03.65.Ud

Proof of that we have to measure the same operator on all qubits. From the proof of Observation 1, we know that at least \mathcal{D}_N measurements are needed to get the expectation values of all the \mathcal{D}_N independent symmetric full N -particle correlations. What if we measure \mathcal{D}_N settings, but several of them are not $\{A_j, A_j, \dots, A_j\}$ -type, but $\{A_j^{(1)}, A_j^{(2)}, \dots, A_j^{(N)}\}$ -type, i.e., we do not measure the same operator on all qubits? Each setting makes it possible to get a single operator containing full N -qubit correlations. Let us denote this operator by M_k for $k = 1, 2, \dots, \mathcal{D}_N$. Then, we know the expectation value of any operator of the space defined by the M_k operators. However, not all M_k 's are permutationally invariant. Thus, the size of the PI subspace of the space of the M_k operators is less than \mathcal{D}_N . We do not have \mathcal{D}_N linearly independent symmetric operators in this space. Thus, \mathcal{D}_N measurement settings are sufficient to measure ϱ_{PI} only if we have settings of the type $\{A_j, A_j, \dots, A_j\}$.

Derivation of Eq. (7). The eigen-decomposition of the correlation term is

$$(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}} = \sum_k \Lambda_{j,n,k} |\Phi_{j,k}\rangle \langle \Phi_{j,k}|. \quad (\text{S1})$$

The individual counts $N_C(A_j)_k$ follow a Poissonian distribution $f(n_c, \lambda_{j,k})$, where $\lambda_{j,k}$ are the parameters of the Poissonian distributions and $\sum_k \lambda_{j,k} = \lambda_j$. The conditional variance, knowing that the total count is $N_C(A_j)$, is

$$\mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}} | N_C(A_j)] = \frac{[\Delta(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}]^2}{N_C(A_j)}. \quad (\text{S2})$$

After straightforward algebra, the variance is obtained as

$$\begin{aligned} & \mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}] \\ &= \sum_m f(m, \lambda_j) \mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}} | N_C(A_j) = m] \\ &= \frac{[\Delta(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}]^2}{\lambda_j - 1}. \end{aligned} \quad (\text{S3})$$

Similar results can be obtained through assuming Poissonian measurement statistics and Gaussian error propagation

[S1, S2]. If $\varrho_0 = \mathbb{1}/2^N$, then $\Delta(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}$ is independent from the choice of A_j . By substituting $A_j = Z$, straightforward calculations give

$$\mathcal{E}^2[(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}] = \frac{\binom{N}{n}^{-1}}{\lambda_j - 1}. \quad (\text{S4})$$

Obtaining the formula for $c_j^{(k,l,m)}$ for the smallest error.

We look for $c_j^{(k,l,m)}$ for which the squared uncertainty given in Eq. (6) is the smallest. In the following, we use the definition given in the main text for \vec{c} , \vec{v} , V and E . Thus, V is matrix mapping a large space \mathbb{R}^l to a small space \mathbb{R}^s . Let E be a non-singular diagonal matrix in the small space. We have to solve

$$\min_{\vec{c}} \|E\vec{c}\|^2 \quad \text{s.t.} \quad V\vec{c} = \vec{v}, \quad (\text{S5})$$

where $\|\vec{a}\|$ is the Euclidean norm of \vec{a} . Using Lagrangian multipliers, we write down the condition for a minimum fulfilling the constraints $V\vec{c} = \vec{v}$

$$\nabla_{\vec{c}} \{ \vec{c}^T E^2 \vec{c} + \sum_{i=1}^s \lambda_i [(V\vec{c})_i - v_i] \} = 0. \quad (\text{S6})$$

Hence, the condition for a local (and, due to convexity, global) minimum is

$$\vec{c} = \frac{1}{2} E^{-2} V^T \vec{\lambda}, \quad (\text{S7})$$

where $\lambda \in \mathbb{R}^s$ is the vector of multipliers. In other words, we have a minimum if and only if $\vec{c} \in \text{range } E^{-2} V^T$. Because the range of V^T is an s -dimensional subspace in \mathbb{R}^l , there is a *unique* \vec{c} in that range such that $V\vec{c} = \vec{v}$. A solution in a closed form can be obtained as

$$\vec{c} = E^{-2} V^T (V E^{-2} V^T)^{-1} \vec{v}. \quad (\text{S8})$$

Simple calculation shows that the $V\vec{c} = \vec{v}$ condition holds

$$V \vec{c} = V E^{-2} V^T (V E^{-2} V^T)^{-1} \vec{v} = \vec{v}. \quad (\text{S9})$$

Table S1: Fidelities to the 4-qubit Dicke states.

measurement	$ D_4^{(0)}\rangle$	$ D_4^{(1)}\rangle$	$ D_4^{(2)}\rangle$	$ D_4^{(3)}\rangle$	$ D_4^{(4)}\rangle$	Σ
full tomography	-0.001 ± 0.002	0.023 ± 0.004	0.873 ± 0.005	0.026 ± 0.004	0.002 ± 0.002	0.922
full tomography (max-like)	0.001	0.021	0.869	0.023	0	0.914
PI tomography	-0.001 ± 0.002	0.040 ± 0.007	0.852 ± 0.009	0.036 ± 0.007	-0.002 ± 0.002	0.925
PI tomography (max-like)	0.003	0.038	0.850	0.037	0	0.928
PI tomography (ran)	0.000 ± 0.002	0.055 ± 0.027	0.814 ± 0.059	0.023 ± 0.027	0.001 ± 0.002	0.893
PI tomography (ran,max-like)	0.004	0.050	0.816	0.020	0.007	0.897

Proof of Observation 2. The eigenstates of $\vec{J}^2 = J_x^2 + J_y^2 + J_z^2$ are usually labelled by $|j, m, \alpha\rangle$, where $\vec{J}^2|j, m, \alpha\rangle = j(j+1)|j, m, \alpha\rangle$, $J_z|j, m, \alpha\rangle = m|j, m, \alpha\rangle$, and α is used to label the different eigenstates having the same j and m [S3]. Let $P_{j,\alpha}$ denote the projector to the subspace of a given j and α . The number of subspaces is denoted by N_{SS} , and, for a given N , it can be calculated from group theory. Moreover, $P_s \equiv P_{N/2,1}$. Using this notation, $\varrho_{PI} = \sum_{j,\alpha} P_{j,\alpha} \varrho P_{j,\alpha} = (P_s \varrho P_s) + \sum_{j < N/2, \alpha} (P_{j,\alpha} \varrho P_{j,\alpha})$. In the basis of \vec{J}^2 eigenstates, ϱ_{PI} can be written as a block diagonal matrix

$$\varrho_{PI} = \bigoplus_{j,\alpha} (\langle P_{j,\alpha} \rangle_{\varrho} \hat{\varrho}_{j,\alpha}), \quad (S10)$$

where $\hat{\varrho}_{j,\alpha}$ are density matrices of size $(2j+1) \times (2j+1)$. In another context,

$$\varrho_{PI} = \sum_{j,\alpha} \langle P_{j,\alpha} \rangle_{\varrho} \varrho_{j,\alpha}, \quad (S11)$$

where $\varrho_{j,\alpha} = P_{j,\alpha} \varrho P_{j,\alpha} / \text{Tr}(P_{j,\alpha} \varrho P_{j,\alpha})$. Based on that, we obtain

$$F(\varrho, \varrho_{PI}) = \langle P_s \rangle_{\varrho}. \quad (S12)$$

Then, due to the separate concavity of the fidelity, i.e., $F(\varrho, p_1 \varrho_1 + p_2 \varrho_2) \geq p_1 F(\varrho, \varrho_1) + p_2 F(\varrho, \varrho_2)$, we obtain $F(\varrho, \varrho_{PI}) \geq \langle P_s \rangle_{\varrho} F(\varrho, \varrho_s) + \sum_{j < N/2, \alpha} \langle P_{j,\alpha} \rangle_{\varrho} F(\varrho, \varrho_{j,\alpha})$. Substituting Eq. (S12) into this inequality, we obtain $F(\varrho, \varrho_{PI}) \geq \langle P_s \rangle_{\varrho}^2 + \sum_{j < N/2, \alpha} \langle P_{j,\alpha} \rangle_{\varrho}^2$. Using the fact that $\langle P_s \rangle_{\varrho} + \sum_{j < N/2, \alpha} \langle P_{j,\alpha} \rangle_{\varrho} = 1$, we obtain

$$F(\varrho, \varrho_{PI}) \geq \langle P_s \rangle_{\varrho}^2 + \frac{(1 - \langle P_s \rangle_{\varrho})^2}{N_{SS} - 1}. \quad (S13)$$

In many practical situations, the state ϱ is almost symmetric and N is large. In such cases the second term in Eq. (S13) is negligible. Thus, a somewhat weaker bound presented in Observation 2 can be used.

Numerical optimization used to minimize $\mathcal{E}_{\text{total}}$. The measurement directions minimizing $\mathcal{E}_{\text{total}}$ can be obtained as follows. Let us represent the measurement directions by three-dimensional vectors $\{\vec{a}_j\}_{j=1}^N$. The operators can be obtained as $A_j = a_{j,x} X + a_{j,y} Y + a_{j,z} Z$.

First, we need an initial guess. This can come from a set of randomly chosen vectors representing the measurement directions. One can also use the result of a minimization for

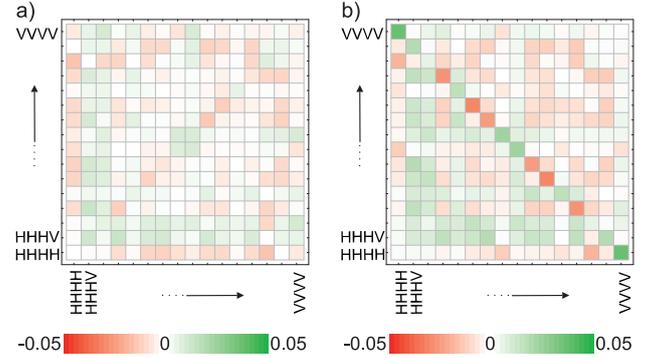


Figure S1: (a) The difference of the real part of the density matrices from optimized settings and the one of full tomography. (b) The difference of the density matrices from random settings and the one of full tomography. For the former, no clear structure is observed, whereas for the latter the largest difference is observed for the anti-diagonal elements.

some measure that characterizes how equally the vectors are distributed. Such a measure is defined by

$$\mathcal{F}(\{v_j\}) = \sum_{k,l} (\vec{v}_k \cdot \vec{v}_l)^{2m}, \quad (S14)$$

where \vec{v}_k represent the measurement directions and \cdot is the scalar product and m is an integer. Such cost functions, called frame potentials, appear in the theory of t -designs essentially for the same purpose.

After we obtain the initial guess from such a procedure, we start an optimization for decreasing $\mathcal{E}_{\text{total}}$. At each iteration of the method, we change the measurement directions by rotating them with a small random angle around a randomly chosen axis. If the change decreases $\mathcal{E}_{\text{total}}$, then we keep the new measurement directions, while if it does not then we discard it. We repeat this procedure until $\mathcal{E}_{\text{total}}$ does not change significantly.

Three-setting witness for estimating the fidelity The three-setting witness for detecting genuine multipartite entanglement in the vicinity of the Dicke state is [S4]

$$\mathcal{W}_{D(4,2)}^{(P3)} = 2 \cdot \mathbb{1} + \frac{1}{6} (J_x^2 + J_y^2 - J_x^4 - J_y^4) + \frac{31}{12} J_z^2 - \frac{7}{12} J_z^4. \quad (S15)$$

For this witness we have [S4]

$$\mathcal{W}_{D(4,2)}^{(P3)} - 3\mathcal{W}_{D(4,2)}^{(P)} \geq 0, \quad (S16)$$

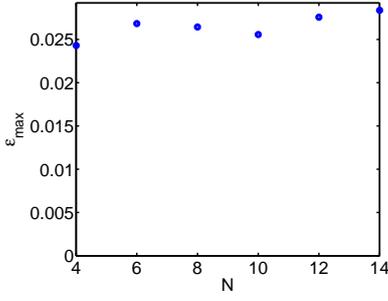


Figure S2: The maximum uncertainty of the Bloch vector elements defined in Eq. (S21) for the optimal measurement settings as a function of the number of qubits, N , for $N = 4, 6, 8, 10, 12$ and 14 .

where the projector witness is defined as

$$\mathcal{W}_{D(4,2)}^{(P)} = \frac{2}{3} \cdot \mathbb{1} - |D_4^{(2)}\rangle\langle D_4^{(2)}|. \quad (\text{S17})$$

Hence, the fidelity with respect to the state $|D_4^{(2)}\rangle$ is bounded from below as [S4]

$$F_{D(4,2)} \geq \frac{2}{3} - \frac{1}{3} \langle \mathcal{W}_{D(4,2)}^{(P3)} \rangle. \quad (\text{S18})$$

Fidelities with respect to the four-qubit Dicke states. In Table S1 we summarize the results for full tomography (full) and for permutationally invariant tomography (pi) for random (ran) and optimized (opt) directions. To obtain a physical density matrix with non-negative eigenvalues we perform a maximum-likelihood fit (max-like) of the measured data. In Fig. S1, the differences between the density matrix obtained from full tomography and the ones obtained from permutationally invariant tomography can be seen.

Efficient representation of permutationally invariant operators on a digital computer. Every PI operator O can be decomposed as

$$O = \sum_{k+l+m+n=N} c_{k,l,m,n}^{(O)} (X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}. \quad (\text{S19})$$

Such a decomposition for operators of the form $(A^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}$ with $A = a_x X + a_y Y + a_z Z$ is given by

$$\sum_{k,l,m} a_x^k a_y^l a_z^m \frac{(k+l+m)!}{k!l!m!} (X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}, \quad (\text{S20})$$

where the summation is carried out such that $k+l+m+n=N$.

Results for larger systems. We determined the optimal A_j for PI tomography for $N = 4, 6, \dots, 14$. In Fig. S2, we plot the maximal uncertainty of the Bloch vector elements

$$\epsilon_{\max} = \max_{k,l,m,n} \mathcal{E}[(X^{\otimes k} \otimes Y^{\otimes l} \otimes Z^{\otimes m} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}] \quad (\text{S21})$$

for the total count realized in the experiment $\lambda_j = \lambda = 2050$ as a function of N , when the state of the system is $\varrho_0 = \mathbb{1}/2^N$. It increases slowly with N . Thus, for large N the number of

counts per measurement setting does not have to increase very much in order to keep the maximal uncertainty of the Bloch vector elements the same as for the $N = 4$ case. In particular, for $N = 14$, a total count of 2797 per setting yields the same maximal uncertainty as we had for the $N = 4$ case.

An upper bound on the uncertainty of PI tomography for ϱ_0 different from the white noise can be obtained by using $[\Delta(A_j^{\otimes(N-n)} \otimes \mathbb{1}^{\otimes n})_{\text{PI}}]_{\varrho_0}^2 = 1$ for error calculations. According to numerics, for optimal A_j for $N = 4, 6, \dots, 14$, ϵ_{\max} remains the same as in the case of white noise, since for the full correlation terms with $n = 0$ the upper bound equals the value for white noise, and the full correlations terms contribute to the noise of the Bloch vector elements with the largest uncertainty. Thus, the total count per setting will not increase more with the number of qubits even for states different from the completely mixed state.

The operators that give a bound on $\langle P_s \rangle$ with three settings for $N = 6$ and 8 are the following

$$\begin{aligned} P_s^{(6)} &\geq \frac{2}{225} (Q_2 + J_z^2) - \frac{1}{90} (Q_4 + J_z^4) + \frac{1}{450} (Q_6 + J_z^6), \\ P_s^{(8)} &\geq -0.001616Q_2 + 0.002200Q_4 - 0.0006286Q_6 \\ &\quad + 0.00004490Q_8 + 0.003265J_z^2 - 0.004444J_z^4 \\ &\quad + 0.001270J_z^6 - 0.00009070J_z^8, \end{aligned} \quad (\text{S22})$$

where $Q_n = J_x^n + J_y^n$. They were determined using semi-definite programming, with a method similar to one used for obtaining three-setting witnesses in Ref. [S4]. They have an expectation value $+1$ for the Dicke states $|D_6^{(3)}\rangle$ and $|D_8^{(4)}\rangle$, respectively. Moreover, their expectation value give the highest possible lower bound on $\langle P_s \rangle$ for states of the form

$$\varrho_{\text{noisy}}(p) = p \frac{\mathbb{1}}{2^N} + (1-p) |D_N^{(N/2)}\rangle\langle D_N^{(N/2)}| \quad (\text{S23})$$

among the operators that are constructed as a linear combination of the operators J_i^n . The validity of the relations in Eq. (S22) can easily be checked by direct calculation.

Bounding the differences between elements of ϱ and ϱ_{PI} based on the fidelity. For any pure state $|\Psi\rangle$, it is possible to bound the difference between $|\langle \Psi | \varrho_{\text{PI}} | \Psi \rangle|$ and $|\langle \Psi | \varrho | \Psi \rangle|$ as

$$|\langle \Psi | \varrho | \Psi \rangle - \langle \Psi | \varrho_{\text{PI}} | \Psi \rangle| \leq \sqrt{1 - F(\varrho, \varrho_{\text{PI}})}. \quad (\text{S24})$$

Thus, if the fidelity is close to 1, then $\langle \Psi | \varrho | \Psi \rangle \approx \langle \Psi | \varrho_{\text{PI}} | \Psi \rangle$, even if $|\Psi\rangle$ is non-symmetric. If $|\Psi\rangle$ is an element of the product basis, e.g., $|0011\rangle$, then Eq. (S24) is a bound on the difference between the corresponding diagonal elements of ϱ and ϱ_{PI} .

Eq. (S24) can be proved as follows: There is a well-known relation between the trace norm and the fidelity [S5]

$$\frac{1}{2} \|\varrho - \varrho_{\text{PI}}\|_{\text{tr}} \leq \sqrt{1 - F(\varrho, \varrho_{\text{PI}})}. \quad (\text{S25})$$

Moreover, for a projector P and density matrices ϱ_k we have [S6]

$$|\text{Tr}(P\varrho_1) - \text{Tr}(P\varrho_2)| \leq \frac{1}{2} \|\varrho_1 - \varrho_2\|_{\text{tr}}. \quad (\text{S26})$$

Combining Eq. (S25) and Eq. (S26), leads to Eq. (S24).

* Present address: Faculty of Physics, University of Vienna, Boltzmannngasse 5, A-1090 Wien, Austria

[S1] C. Schmid, Ph.D. Thesis, Ludwig-Maximilian-Universität, Munich, Germany, 2008; D.F.V. James *et. al.*, Phys. Rev. A **64**,

052312 (2001).

[S2] B. Jungnitsch *et. al.*, Phys. Rev. Lett. **104**, 210401 (2010).

[S3] See, for example, J.I. Cirac *et. al.*, Phys. Rev. Lett. **82**, 4344 (1999).

[S4] G. Tóth *et al.*, New J. Phys. **11**, 083002 (2009).

[S5] J.A. Miszczak *et al.*, J. Quant. Inf. Comp. **9**, 0103 (2009).

[S6] See Eqs. (9.18) and (9.22) of M.A. Nielsen and I.L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, Cambridge, 2000).