Machine-Learning-Assisted Many-Body Entanglement Measurement

Johnnie Gray,[1,*] Leonardo Banchi,[1] Abolfazl Bayat,[2,1] and Sougato Bose[1]

[1]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom*
[2]*Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 610051, China*

Entanglement not only plays a crucial role in quantum technologies, but is key to our understanding of quantum correlations in many-body systems. However, in an experiment, the only way of measuring entanglement in a generic mixed state is through reconstructive quantum tomography, requiring an exponential number of measurements in the system size. Here, we propose a machine-learning-assisted scheme to measure the entanglement between arbitrary subsystems of size N_A and N_B, with $\mathcal{O}(N_A + N_B)$ measurements, and without any prior knowledge of the state. The method exploits a neural network to learn the unknown, nonlinear function relating certain measurable moments and the logarithmic negativity. Our procedure will allow entanglement measurements in a wide variety of systems, including strongly interacting many-body systems in both equilibrium and nonequilibrium regimes.

Introduction.—Entanglement is a key property for many emerging quantum technologies [1–6], but also is essential for understanding the structure of strongly correlated many-body systems [7,8]. Despite its paramount importance, only for the very limited case of a bipartition of a pure state can the entanglement, quantified by subsystem entropy, be measured in an efficient and state-independent way [9]. There are multiple proposals to carry out such a scheme in various physical systems, such as optical lattices [10,11], quantum dot arrays [12] and Gaussian systems [13]. Recently, some of these have also been experimentally realized in simulated spin chains, e.g., in cold atoms [14] and photonic chips [15]. Nonetheless, pure states are very rare: they are not only difficult to prepare in realistic situations, but also difficult to maintain in the presence of an environment. For example, just consider the entanglement between (i) optical modes traversing fibers, crucial for quantum communication, (ii) spatially separated parts of an extended many-body pure state, important for characterizing long range entanglement [16–21], and (iii) two systems in a thermal state—in none of the above cases, ironically, can the entanglement entropy quantify the entanglement. Witnesses do exist for specific forms of entanglement, but these are state dependent and provide only a simple yes or no answer [22–24] or bounds on the quantity of entanglement [25–31]. However, the crucial task of being able to accurately measure entanglement for mixed states in an experimental setting remains open.

While for pure states bipartite entanglement is uniquely defined by the entropy of the subsystems, for mixed states the landscape is far more complex [32,33]. Aside from isolated special cases such as two qubit states [34] and bosonic Gaussian states [13,35], only the (logarithmic) negativity [36–39] is a computationally tractable quantity

[40]. It bounds the distillable entanglement and teleportation capacity [38], and is a pivotally important quantity to estimate for both quantum technologies [32,33,41] and condensed matter systems [42]. Nonetheless, there is no state-independent observable that can measure the logarithmic negativity, and thus its experimental measurement requires full state tomography [43]—demanding, in general, an exponential number of measurements in the system size. Recently, polynomial tomography schemes have emerged, such as tensor networks for lowly entangled states [41,44], or breakthroughs in neural network state reconstruction [45,46]. However, these may be insufficient for estimating entanglement, since many entanglement measures, such as the logarithmic negativity, are not continuous [47]. Namely, even if reconstructed state ρ_r approximates actual state ρ closely, the two may have significantly different negativities [48].

Here, we put forward a machine-learning-assisted scheme for accurately estimating the logarithmic negativity in a completely general and realistic setting, using an efficient number of measurements—scaling polynomially with system size. Our estimator works for a wide range of states, and is remarkably accurate for highly entangled states. Our method is based on measuring a finite number of moments of a partially transposed density matrix [49–51] from which we extract the entanglement negativity using machine learning. This *direct* estimation of negativity avoids approximate state reconstruction [41,44–46], and represents a new front in applying classical machine learning to quantum problems [8,52–55]. Moreover, we propose a new method for measuring those moments, beyond Refs. [49–51], which is experimentally feasible in the many-body setting, since the individual building blocks have already been demonstrated in solid state [56] and cold atoms [14].

Logarithmic Negativity.—Logarithmic negativity [36–39] for a generic mixed state ρ_{AB} quantifies the entanglement between subsystems A and B. It is defined as

$$\mathcal{E} = \log_2|\rho_{AB}^{T_A}| = \log_2|\rho_{AB}^{T_B}| = \log_2\sum_k|\lambda_k|, \qquad (1)$$

with $|\cdot|$ the trace norm, $\rho_{AB}^{T_X}$ the partial transpose with respect to subsystem X, and $\{\lambda_k\}$ the eigenvalues of $\rho_{AB}^{T_X}$. Because of the nontrivial dependence of $\mathcal{E}$ on ρ_{AB}, there is no state-independent observable that can measure it— generally demanding full state tomography. The $\{\lambda_k\}$ are the roots of the characteristic polynomial, $P(\lambda) = \det(\lambda - \rho_{AB}^{T_B}) = \sum_n c_n\lambda^n$, where each c_n is a polynomial function of the partially transposed moments:

$$\mu_m = \text{Tr}[(\rho_{AB}^{T_B})^m] = \sum_k\lambda_k^m. \qquad (2)$$

In this way, full information about the spectrum $\{\lambda_k\}$ is contained in $\{\mu_m\}$. It is known that measuring these moments is technically possible using m copies of the state and controlled swap operations [49]. However, even if these experimentally challenging operations were available, the problem of extracting $\{\lambda_k\}$ from the moments is notoriously ill conditioned [57], with a closely related problem being described as numerically catastrophic. Alongside this, an exponential number of moments respective to the size of AB are needed to exactly solve the equations. On the other hand, to estimate the logarithmic negativity, a precise knowledge of all λ_k is not required. Since $-\frac{1}{2} \le \lambda_k \le 1$ for all k [58] and $\sum_k\lambda_k = 1$, generically, the magnitude of the moments quickly decreases with m, with the first few carrying the most information. Backing up this intuition, we will show that the moments required, $\{\mu_m : m \le M\}$, to accurately estimate the entanglement can number as few as $M = 3$. We do this by employing machine learning to *directly* map moments to logarithmic negativity, avoiding reconstruction of the spectrum or state. Note that μ_0 is simply the dimension of the systems Hilbert space, while $\mu_1 = 1$ in all cases. Additionally, it can be easily shown that μ_2 is equal to the purity of the state $= \text{Tr}[\rho_{AB}^2]$, and as such, $M \ge 3$ is needed to extract any information about $\mathcal{E}$. In this sense our method is optimal in terms of number of copies.

Measuring the moments of $\rho_{AB}^{T_B}$—.The method for measuring the moments proposed in Ref. [49] based on three-body controlled swaps is practically challenging in a many-body setup where natural interactions are two-body. A simpler protocol, for 4 moments only, was provided in Ref. [51]. Here, we show that any moment in Eq. (2) can be measured using only SWAP operators between the *individual* constituents of the m copies of the state ρ_{AB}, namely, $\rho_{AB}^{\otimes m} = \otimes_{c=1}^m \rho_{A_cB_c}$. This general setup is shown in Fig. 1(a), where the mixedness of ρ_{AB} arises from possible entanglement with a third system C, such that $\rho_{AB} = \text{Tr}_C|\Psi_{ABC}\rangle\langle\Psi_{ABC}|$ with $|\Psi_{ABC}\rangle$ being a pure tripartite state. The first step is to write the matrix power as an expectation of a permutation operator, similar to Refs. [9,59], but here on the partially transposed copies:

$$\mu_m = \text{Tr}\left[\left(\overset{m}{\underset{c=1}{\otimes}}\rho_{A_cB_c}^{T_{B_c}}\right)\mathbb{P}^m\right]$$
$$= \text{Tr}\left[\left(\overset{m}{\underset{c=1}{\otimes}}\rho_{A_cB_c}\right)(\mathbb{P}^m)^{T_B}\right], \qquad (3)$$

where $\mathbb{P}^m$ is any linear combination of cyclic permutation operators of order m and the second line makes use of the identity $\text{Tr}(\rho_{AB}^{T_B}O) = \text{Tr}(\rho_{AB}O^{T_B})$, valid for any operator O. A schematic of the equality in Eq. (3) for $m = 3$ is shown in Fig. 1(b). In the Supplemental Material [48] we provide a
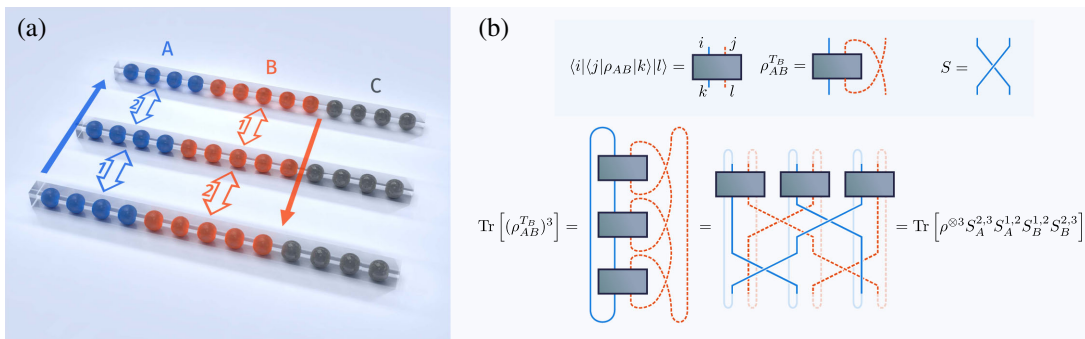

FIG. 1. Schematics. (a) Example measurement setup for the moments, $\mu_m = \text{Tr}[(\rho_{AB}^{T_B})^m]$, here for $m = 3$, from which one can extract the logarithmic negativity $\mathcal{E}$ between A and B. The generic mixedness of ρ_{AB} could arise from entanglement with environment C. Here the subsystems contain N_A, N_B, and N_C particles, respectively. The scheme involves three copies of the original system, and two counterpropagating sets of measurements on A and B, ordered by the shown numbers, with direction depicted by the filled arrows. (b) Diagrammatic proof (for $m = 3$) of the equivalence between the moments μ_m and expectation of two opposite permutations (decomposed as swaps) on A and B—from which a measurement scheme can be derived.

choice of $\mathbb{P}^m$ with a neat operational meaning, both for spin and bosonic systems. For spin lattices, our choice of $\mathbb{P}^m$ to measure the moments μ_m results to the following steps in practice: (i) prepare m copies of the state ρ_{AB}; (ii) sequentially measure a "forward" sequence of adjacent swaps, $S_A^{c,c+1}$ between neighboring copies of system A from $c = 1$ to $m - 1$; (iii) sequentially measure a "backward" sequence of adjacent swaps, $S_B^{c,c-1}$ between neighboring copies of system B from $c = m$ to 2; (iv) repeat these steps in order to yield an expectation value. This procedure is also depicted for $m = 3$ in Fig. 1(a). For bosonic lattices, our procedure corresponds to the following steps: (i) prepare m copies of the state ρ_{AB}; (ii) perform forward Fourier transforms between modes in different copies for each site in A—this can be achieved using a series of beam splitters [71]; (iii) perform backwards (reverse) Fourier transform between modes in different copies for each site in B, via reverse beam splitter transformations; (iv) measure the boson occupation numbers $n_{j,c}$ on all sites $j \in \{A, B\}$ and all copies c to compute $\phi = e^{i \sum_{j \in \{A,B\},c} 2\pi c n_{j,c}/m}$. (v) Repeat these steps to obtain the expectation value μ_m as an average of ϕ. Both procedures require $\mathcal{O}(N_A + N_B)$ measurements for each m between 2 and M, and are explained in detail in the Supplemental Material [48]. This is in stark contrast to tomography, which generically for qubit systems requires $2^{2(N_A+N_B)}$ measurement settings.

It is worth emphasizing the difference between our procedure and recently proposed operational methods for measuring Renyi entropies [10,12,72]. First of all, Renyi entropies only quantify entanglement for pure states, and cannot be used in the more general mixed state scenario. Second, while for entropies the operations are only performed on a single subsystem, here, one performs both forward and backward operations on two subsystems at once, as explained above. Remarkably, even though partially transposed density matrices are generically unphysical, measurement of their moments is possible.

Machine learning entanglement.—We focus now on estimating the logarithmic negativity from the information contained in the moments, μ_m. One approach using only the even moments has been proposed in the quantum field theory literature [42,73] by exploiting numerical extrapolation. However, this method neglects the odd moments and generally requires a large number of moments and thus copies. We have developed an alternative analytical method based on Chebyshev functional approximation, detailed in the Supplemental Material [48], which takes into account these odd moments. Indeed with the same number of copies we find it produces more accurate estimates, and thus serves as a reference quantity. The Chebyshev expansion is analytically tractable, and becomes accurate for large enough M, as is shown in the Supplemental Material [48]. Nonetheless, this expansion is based on a linear mapping between the moments and the negativity, despite

this relationship being inherently nonlinear. Therefore it is natural to think that a nonlinear transformation could be more optimal, and thus more efficient for smaller M—namely, fewer copies.

Machine learning has recently emerged as a key tool for modeling an unknown nonlinear relationship between sets of data. In the supervised learning paradigm, one trains a model with a set of known inputs and their corresponding outputs. Once trained, the model can then be used to predict the unknown output of new input data. Here, we take the moments μ_m as the input and the logarithmic negativity $\mathcal{E}$ as the output. Training is performed by taking a large set of states for which μ_m and $\mathcal{E}$ can be computed on a classical computer. This model can then be used to predict $\mathcal{E}$ from a set of experimentally measured moments. The experimental system under study motivates the choice of which training states to use, so that they share, e.g., similar entanglement features. Among the most successful machine learning algorithms for nonlinear regression are supervised vector machines [74], random decision forests [75], and deep neural networks [76,77]. However, we have found that using the same training set for each, neural networks are superior when it comes to predicting logarithmic negativity for a wide range of states beyond the training set. As we show with our numerical results, neural networks provide a very accurate method for extracting the logarithmic negativity with as few as $M = 3$ copies. The details of our neural network construction can be found in the Supplemental Material [48].

Training with random states.—In order to train a neural network, a set of suitable training states are required for which both the moments and logarithmic negativities are known. From an entanglement perspective, relevant states in condensed matter physics can be classified as either area law, or volume law. In the first case, the entanglement of a subsystem A with the rest is proportional to the number of qubits along their boundary. In the second, this entanglement is instead proportional to N_A, the number of qubits in A. Area law states arise as low energy eigenstates of local gapped Hamiltonians, with logarithmic corrections in critical systems. Volume law states, however, are associated with the eigenstates found in the midspectrum, and as such arise in nonequilibrium dynamics, e.g., quantum quenches [78,79].

Rather than concentrate on a training with a specific model system, we initially consider the very general case of random states. To encompass both area- and volume-law states, we consider two classes of states $|\Psi_{ABS}\rangle$: (i) random generic pure states (R-GPS), e.g., sampled from the Haar measure, which typically have volume-law entanglement [80,81]; (ii) random matrix product states (R-MPS) with fixed bond dimension, which satisfy an area law by construction [8]. In order to generate a training set with a wide range of entanglement features, subsystem sizes, and mixedness, we perform the following procedure: (i) For a fixed number of qubits N, take either a R-GPS, or R-MPS with bond dimension D. (ii) Take different tripartitions

such that $N = N_A + N_B + N_C$, and for each calculate μ_m and $\mathcal{E}$ for ρ_{AB}. (iii) Repeat for different random instances, while separately varying N and D. Further generation and training are provided in the Supplemental Material [48].

Numerical results for random Sstates.—To check the performance of our neural network estimator, we take the set of random states described in the previous section and split this data in two, one-half for training the neural network model, and the other as "unseen" test data. In Fig. 2(a) we plot the machine learning model's predictions, $\mathcal{E}_M^{\mathrm{ML}}$, for the test data, using only $M = 3$ copies, in which a high degree of accuracy is achieved. In the inset of Fig. 2(a), we plot a histogram of the errors $\mathcal{E}_M^{\mathrm{ML}} - \mathcal{E}$, which displays a very sharp peak at zero error with standard deviation ~ 0.09. A further improvement, particularly in outliers, is achieved by increasing the number of copies M to 10, see Fig. 2(b), where the error standard deviation decreases to ~ 0.07 Regardless, the machine learning method is already very accurate for extracting entanglement using only three copies. The machine learning approach works particularly well for large bond dimension and volume-law-like states—an important fact given that these are the exact cases where efficient tomography fails. A more detailed discussion about sensitivity and ascribing errors to machine learning predictions can be found in the Supplemental Material [48].

Numerical results for physical states.—We now consider the more realistic setting of quench dynamics in a many-body system. We take a system of N spin-1/2 particles with nearest neighbour Heisenberg Hamiltonian $H = J \sum_{i=1}^{N-1} \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_{i+1}$ with J the interaction strength and $\boldsymbol{\sigma}_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ the vector of Pauli matrices acting on site i. The system is initialized in the (separable) Néel-state $|\Psi(0)\rangle = |\uparrow\downarrow\uparrow\ldots\rangle$. As the chain unitarily evolves in time as $|\Psi(t)\rangle = e^{-iHt}|\Psi(0)\rangle$, it becomes entangled, with an effective MPS
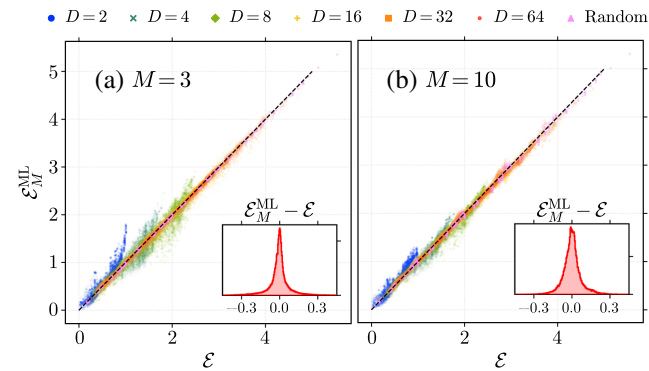
description whose bond dimension increases until the state is essentially volume law [78,79].

In Fig. 3 we plot the evolution of $\mathcal{E}$ and three approximation methods, as functions of time for three different choices of subsystems. The three methods are the Chebyshev approximation with $M = 10$ and $M = 20$, discussed in the Supplemental Material [48], and machine learning with $M = 3$, with respective approximate entanglements $\mathcal{E}_{M=10}^{\mathrm{Cheb}}$, $\mathcal{E}_{M=20}^{\mathrm{Cheb}}$ and $\mathcal{E}_{M=3}^{\mathrm{ML}}$. In Fig. 3(a) we consider a specific partition with $N_A = 2$, $N_B = 2$, and $N_C = 4$. Here, $\mathcal{E}_{M=3}^{\mathrm{ML}}$ and $\mathcal{E}_{M=20}^{\mathrm{Cheb}}$ are comparably accurate. For larger subsystems, as shown in Figs. 3(b) and 3(c), the machine learning approximation, using only $M = 3$ copies, significantly outperforms the Chebyshev approximations, using either $M = 10$ or $M = 20$ copies. It is remarkable that despite being trained on a arbitrary set of random states with no knowledge of the underlying physical system, the evolution of $\mathcal{E}$ is accurately captured by the neural network estimator for all partitions and times, with as few as $M = 3$ copies.

In the Supplemental Material [48], we explore various other physical situations, including the ground state of an *XX* chain across its phase transition, the fully symmetric *W* state, and a quench across the critical point of a transverse Ising chain.

Conclusions.—The measurement of logarithmic negativity in generic multiparticle mixed states (where Renyi entropies are insufficient to quantify entanglement) has so far relied on the complete reconstruction of a quantum state, which in general requires an exponential number of measurements, and is thus limited to small system sizes. In this work, we have devised an alternative strategy, based on machine learning, by which we can extract the entanglement from very few measurements. These measurements
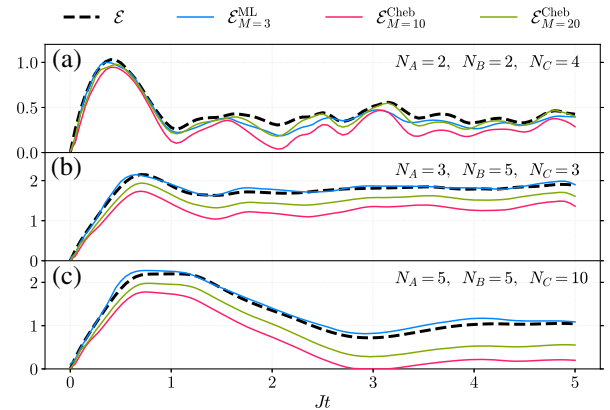

FIG. 2. Machine learning entanglement. Estimated logarithmic negativity $\mathcal{E}_M^{\mathrm{ML}}$, using a machine learning vs actual logarithmic negativity $\mathcal{E}$, for the same set of random states described in the main text. Training and prediction is performed using the moments μ_m generated from (a) $M = 3$ copies, (b) $M = 10$ copies. The respective insets show the distribution of error, $\mathcal{E}_M^{\mathrm{ML}} - \mathcal{E}$.

FIG. 3. Estimating entanglement for physical states. Logarithmic negativity $\mathcal{E}$ and its approximations, using machine learning ($M = 3$) and a Chebyshev expansion ($M = 10$, $M = 20$), as a function of time Jt for the quench dynamics of a Heisenberg spin chain initialized in a Néel state. A variety of system sizes with different partitions is shown here: (a) $N = 8$, $N_A = N_B = 2$; (b) $N = 11$, $N_A = 3$, $N_B = 5$; (c) $N = 20$, $N_A = N_B = 5$.

are based on two counterpropagating series of swap operators on copies of the state—techniques for achieving this have already been demonstrated in a number of physical setups ranging from quantum dot arrays [56,82] to cold atoms in optical lattices [14,83]. Our method is based on learning the functional relationship between these measurement outcomes—the first few moments of the partially transposed density matrix—and the logarithmic negativity using a neural network. Remarkably, our method is already very accurate for as few as three copies—making it very resource efficient and desirable for practical applications—even for estimating the entanglement of highly entangled physical states, such as those arising in quantum quenches.

---

[*]john.gray.14@ucl.ac.uk

[1] C. R. Rao, Bull. Calcutta Math. Soc. **37**, 81 (1945).
[2] P. W. Shor, SIAM Rev. **41**, 303 (1999).
[3] A. W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. **103**, 150502 (2009).
[4] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
[5] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).
[6] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[7] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Rev. Mod. Phys. **80**, 517 (2008).
[8] U. Schollwöck, Ann. Phys. (Amsterdam) **326**, 96 (2011).
[9] P. Horodecki and A. Ekert, Phys. Rev. Lett. **89**, 127902 (2002).
[10] A. J. Daley, H. Pichler, J. Schachenmayer, and P. Zoller, Phys. Rev. Lett. **109**, 020505 (2012).
[11] C. M. Alves and D. Jaksch, Phys. Rev. Lett. **93**, 110501 (2004).
[12] L. Banchi, A. Bayat, and S. Bose, Phys. Rev. B **94**, 241117 (2016).
[13] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Rev. Mod. Phys. **84**, 621 (2012).
[14] R. Islam, R. Ma, P. M. Preiss, M. E. Tai, A. Lukin, M. Rispoli, and M. Greiner, Nature (London) **528**, 77 (2015).
[15] I. Pitsios, L. Banchi, A. S. Rab, M. Bentivegna, D. Caprara, A. Crespi, N. Spagnolo, S. Bose, P. Mataloni, R. Osellame et al., Nat. Commun. **8**, 1569 (2017).
[16] B. Reznik, Found. Phys. **33**, 167 (2003).
[17] A. Bayat, P. Sodano, and S. Bose, Phys. Rev. B **81**, 064429 (2010).
[18] B. Alkurtass, A. Bayat, I. Affleck, S. Bose, H. Johannesson, P. Sodano, E. S. Sørensen, and K. Le Hur, Phys. Rev. B **93**, 081106 (2016).
[19] J. Gray, S. Bose, and A. Bayat, Phys. Rev. B **97**, 201105 (2018).
[20] L. Banchi, F. Colomo, and P. Verrucchi, Phys. Rev. A **80**, 022341 (2009).
[21] T. Roscilde, P. Verrucchi, A. Fubini, S. Haas, and V. Tognetti, Phys. Rev. Lett. **93**, 167203 (2004).
[22] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
[23] B. M. Terhal, Phys. Lett. A **271**, 319 (2000).
[24] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, Phys. Rev. A **62**, 052310 (2000).
[25] F. G. S. L. Brandao, Phys. Rev. A **72**, 022310 (2005).
[26] K. Audenaert and M. Plenio, New J. Phys. **8**, 266 (2006).
[27] D. Cavalcanti and M. O. Terra Cunha, Appl. Phys. Lett. **89**, 084102 (2006).
[28] J. Eisert, F. G. Brandao, and K. M. Audenaert, New J. Phys. **9**, 46 (2007).
[29] O. Gühne, M. Reimpell, and R. F. Werner, Phys. Rev. Lett. **98**, 110502 (2007).
[30] O. Gühne, M. Reimpell, and R. F. Werner, Phys. Rev. A **77**, 052317 (2008).
[31] H. A. Carteret, arXiv:1605.08751.
[32] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[33] M. B. Plenio and S. Virmani, Quantum Inf. Comput. **7**, 1 (2007).
[34] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).
[35] M. M. Wolf, G. Giedke, O. Krüger, R. F. Werner, and J. I. Cirac, Phys. Rev. A **69**, 052320 (2004).
[36] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Phys. Rev. A **58**, 883 (1998).
[37] J. Lee, M. Kim, Y. Park, and S. Lee, J. Mod. Opt. **47**, 2151 (2000).
[38] G. Vidal and R. F. Werner, Phys. Rev. A **65**, 032314 (2002).
[39] M. B. Plenio, Phys. Rev. Lett. **95**, 090503 (2005).
[40] Y. Huang, New J. Phys. **16**, 033027 (2014).
[41] B. Lanyon, C. Maier, M. Holzäpfel, T. Baumgratz, C. Hempel, P. Jurcevic, I. Dhand, A. Buyskikh, A. Daley, M. Cramer et al., Nat. Phys. **13**, 1158 (2017).
[42] P. Calabrese, J. Cardy, and E. Tonni, Phys. Rev. Lett. **109**, 130502 (2012).
[43] G. M. D'Ariano, M. G. Paris, and M. F. Sacchi, Adv. Imaging Electron Phys. **128**, 206 (2003).
[44] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, Nat. Commun. **1**, 149 (2010).
[45] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. **14**, 447 (2018).
[46] G. Torlai and R. G. Melko, Phys. Rev. Lett. **120**, 240503 (2018).
[47] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, England, 2006).
[48] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.121.150503 for details regarding measuring moments, the Chebyshev method,

error analysis, implementation details and extra numerical results. This material includes Refs. [8,10,12,14,41, 56,57,60–70].

[49] H. A. Carteret, Phys. Rev. Lett. **94**, 040502 (2005).

[50] J. Cai and W. Song, Phys. Rev. Lett. **101**, 190503 (2008).

[51] K. Bartkiewicz, P. Horodecki, K. Lemr, A. Miranowicz, and K. Życzkowski, Phys. Rev. A **91**, 032315 (2015).

[52] J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).

[53] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[54] A. Hentschel and B. C. Sanders, Phys. Rev. Lett. **104**, 063603 (2010).

[55] L. Banchi, N. Pancotti, and S. Bose, npj Quant. Inf. **2**, 16019 (2016).

[56] J. R. Petta, A. C. Johnson, J. M. Taylor, E. A. Laird, A. Yacoby, M. D. Lukin, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Science **309**, 2180 (2005).

[57] G. Viano, J. Math. Anal. Appl. **156**, 410 (1991).

[58] S. Rana, Phys. Rev. A **87**, 054301 (2013).

[59] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Phys. Rev. Lett. **88**, 217901 (2002).

[60] J.-P. Blaizot and G. Ripka, *Quantum Theory of Finite Systems* (MIT Press, Cambridge, MA, 1986), Vol. 3.

[61] R. K. Puddy, L. W. Smith, H. Al-Taie, C. H. Chong, I. Farrer, J. P. Griffiths, D. A. Ritchie, M. J. Kelly, M. Pepper, and C. G. Smith, Appl. Phys. Lett. **107**, 143501 (2015).

[62] M. Veldhorst, H. Eenink, C. Yang, and A. Dzurak, Nat. Commun. **8**, 1766 (2017).

[63] L. R. Mead and N. Papanicolaou, J. Math. Phys. (N.Y.) **25**, 2404 (1984).

[64] I. Han, D. Malioutov, H. Avron, and J. Shin, arXiv:1606.00942.

[65] V. Eisler and Z. Zimborás, New J. Phys. **17**, 053048 (2015).

[66] A. Coser, E. Tonni, and P. Calabrese, J. Stat. Mech. (2016) 033116.

[67] G. Aubrun, Random Matrices Theory Appl. **01**, 1250001 (2012).

[68] M. Fukuda and P. Śniady, J. Math. Phys. (N.Y.) **54**, 042202 (2013).

[69] J. Bergstra, D. Yamins, and D. Cox, *Proceedings of the International Conference on Machine Learning* (JMLR: W&CP, Atlanta, Georgia, USA, 2013), p. 115.

[70] F. Chollet *et al.*, Keras, https://github.com/fchollet/keras (2015).

[71] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).

[72] D. A. Abanin and E. Demler, Phys. Rev. Lett. **109**, 020504 (2012).

[73] C. De Nobili, A. Coser, and E. Tonni, J. Stat. Mech. (2015) P06021.

[74] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and other Kernel-Based Learning Methods* (Cambridge University Press, Cambridge, England, 2000).

[75] T. K. Ho, IEEE Trans. Pattern Anal. Mach. Intell. **20**, 832 (1998).

[76] R. Rojas, *Neural Networks: A Systematic Introduction* (Springer Science & Business Media, New York, 2013).

[77] J. Schmidhuber, Neural Netw. **61**, 85 (2015).

[78] P. Barmettler, M. Punk, V. Gritsev, E. Demler, and E. Altman, Phys. Rev. Lett. **102**, 130603 (2009).

[79] A. Nanduri, H. Kim, and D. A. Huse, Phys. Rev. B **90**, 064201 (2014).

[80] S. Popescu, A. J. Short, and A. Winter, Nat. Phys. **2**, 754 (2006).

[81] A. Hamma, S. Santra, and P. Zanardi, Phys. Rev. Lett. **109**, 040502 (2012).

[82] M. Schuld, I. Sinayskiy, and F. Petruccione, Contemp. Phys. **56**, 172 (2015).

[83] S. Trotzky, P. Cheinet, S. Fölling, M. Feld, U. Schnorrberger, A. M. Rey, A. Polkovnikov, E. Demler, M. Lukin, and I. Bloch, Science **319**, 295 (2008).