HAL
open science

Optimal entanglement witnesses: a scalable data-driven approach

Irénée Frérot, Tommaso Roscilde

To cite this version:

Irénée Frérot, Tommaso Roscilde. Optimal entanglement witnesses: a scalable data-driven approach. 2021. hal-03099169

HAL Id: hal-03099169

https://hal.science/hal-03099169v1

Preprint submitted on 6 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimal entanglement witnesses: a scalable data-driven approach

Irénée Frérot[1,2,*] and Tommaso Roscilde[3,†]

[1] *ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Barcelona, Spain*

[2] *Max-Planck-Institut für Quantenoptik, D-85748 Garching, Germany*

[3] *Laboratoire de Physique, CNRS UMR 5672, Ecole Normale Supérieure de Lyon, Université de Lyon, 46 Allée d'Italie, Lyon, F-69364, France*

Multipartite entanglement is the key resource allowing quantum devices to outperform their classical counterparts, and entanglement certification is fundamental to assess any quantum advantage. The only scalable certification scheme relies on entanglement witnessing, typically effective only for special entangled states. Here we focus on finite sets of measurements on quantum states (hereafter called quantum data); and we propose an approach which, given a particular spatial partitioning of the system of interest, can effectively ascertain whether or not the data set is compatible with a separable state. When compatibility is disproven, the approach produces the optimal entanglement witness for the quantum data at hand. Our approach is based on mapping separable states onto equilibrium classical field theories on a lattice; and on mapping the compatibility problem onto an inverse statistical problem, whose solution is reached in polynomial time whenever the classical field theory does not describe a glassy system. Our results pave the way for systematic entanglement certification in quantum devices, optimized with respect to the accessible observables.

Introduction. Preparing and processing strongly entangled many-body states, in both a controlled and scalable way, is the goal of all quantum simulators and computers. Indeed, as the efficient representation of generic entangled many-body states is impossible on classical machines, entanglement represents the key computational resource of these devices [1, 2]. As a consequence, developing generic and scalable methods to certify entanglement in multipartite systems stands as a grand challenge of quantum information science. Even more fundamentally, entanglement certification is a central task to probe the resilience of quantum correlations from the microscopic world to the macroscopic one [3].

Any practical method must circumvent the tomographic reconstruction of the full density matrix [4, 5] (which implies a number of measurements scaling exponentially with system size), and it should instead infer entanglement from the partial information contained in a given data set of measurement results (hereafter referred to as *quantum data*). When one adopts this data-driven strategy, the goal of entanglement certification is to establish whether or not the quantum data are compatible with a separable state [5–7]. Given an extended quantum system composed of $N_{\rm tot}$ degrees of freedom, grouped together into $N \leq N_{\rm tot}$ clusters [see Fig. 1(a)], the state $\hat\rho$ of the system is separable [8] if it can be written in the form

$$\hat\rho_p := \int d\boldsymbol{\lambda}\; p(\boldsymbol{\lambda})\; \hat\rho_{\rm prod}(\boldsymbol{\lambda}) \qquad (1)$$

where $\hat\rho_{\rm prod}(\boldsymbol{\lambda}) = \otimes_{i=1}^N |\psi_i(\boldsymbol{\lambda}_i)\rangle\langle\psi_i(\boldsymbol{\lambda}_i)|$ is a product state of the partition, $|\psi_i(\boldsymbol{\lambda}_i)\rangle$ being the state of the i-th cluster, parametrized by parameters $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, ..., \boldsymbol{\lambda}_i, ..., \boldsymbol{\lambda}_N)$, distributed according to $p(\boldsymbol{\lambda}) \geq 0$. The distribution p fully specifies classical correlations across the partition. A multipartite entangled state $\hat\rho$, on the other hand, cannot be written in the above form. Given a set of observables $\hat A_a$ ($a = 1, ..., R$), multipartite entanglement is therefore *witnessed* by the quantum data set $\{\langle\hat A_a\rangle_{\hat\rho}\}_{a=1}^R$ [where $\langle\hat A_a\rangle_{\hat\rho} = {\rm Tr}(\hat A_a\hat\rho)$] if one proves that the latter cannot be reproduced by *any* separable state. This task is accomplished by proving that the quantum data violate an entanglement witness (EW) inequality, $\langle\hat{\mathcal W}\rangle_{\hat\rho_p} = \sum_a W_a\langle\hat A_a\rangle_{\hat\rho_p} \geq B_{\rm sep}$, valid for all separable states $\hat\rho_p$ [9]. Here W_a are suitable coefficients and $B_{\rm sep}$ is the so-called separable bound.

EW operators $\hat{\mathcal W}$ are generally defined based on the properties of special entangled states (e.g. squeezed states, total spin singlets, etc.) [9], and failure of a data set to violate a given EW inequality does not exclude the existence of a different violated inequality involving the same data, yet to be discovered. This may erroneously suggest that entanglement witnessing is limited by creativity and physical insight; and that the entanglement witnessing problem ("is a quantum data set compatible with a separable state?") [5–7] is generically undecidable. The goal of our work is to show that this is *not* the case, and that the entanglement witnessing capability of a quantum data set can be *exhaustively* tested. Our key insight is that the problem of finding the distribution $p(\boldsymbol{\lambda})$, which defines the separable state reproducing at best the quantum data, is a statistical inference problem; and remarkably it has the structure of a convex optimization problem, whose solution can be attained in a time scaling polynomially with the partition size (under mild assumptions), and with the Hilbert space dimension of the subsystems composing the partition. When the optimal separable state fails to reproduce the quantum data, the distance between the quantum data set $\{\langle\hat A_a\rangle_{\hat\rho}\}$ and the optimal separable set $\{\langle\hat A_a\rangle_{\hat\rho_p}\}$ allows one to reconstruct the optimal EW inequality violated by the quan-

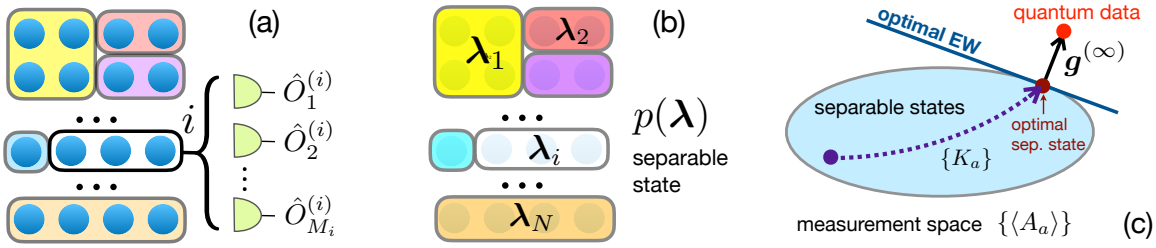

FIG. 1: (a) Partition of a quantum device into N clusters, each of which is subject to M_i measurements; (b) A separable state of the system is described as a probability distribution $p(\boldsymbol{\lambda})$ of local states defined by the $\{\boldsymbol{\lambda}_i\}$ parameters; (c) Our algorithm builds a trajectory of separable states (parametrized by couplings $\{K_a\}$ defining $p(\boldsymbol{\lambda})$) which converges to the optimal state approximating at best some target quantum data. If the state fails to reproduce the quantum data exactly, the vector joining the optimal separable data and the quantum data reconstructs the optimal EW inequality.

tum data. We benchmark our approach by establishing new EW inequalities satisfied by the low-temperature states of the Heisenberg antiferromagnetic chain and the quantum Ising chain; in the latter case, our new EW inequalities outperform all previously known EW criteria for multipartite entanglement. Our work parallels the recent mapping of the Bell-nonlocality detection problem onto an inverse statistical problem [10], and it offers an efficient scheme for entanglement detection in state-of-the-art quantum devices within a device-dependent scenario.

Quantum data set. For definiteness, we assume that on each subsystem $i = 1, ..., N$, M_i local observables $\hat{O}_m^{(i)}$ can be measured ($m = 1, \dots M_i$; *e.g.* the Pauli matrices $\hat{\sigma}_a^{(i)}$, $a \in \{x, y, z\}$ for individual qubits taken as subsystems). For convenience, we denote the local identity operator by $\hat{O}_0^{(i)} := \mathbb{1}$. In order to reveal entanglement, these local observables must be non-commuting ($[\hat{O}_m^{(i)}, \hat{O}_n^{(i)}] \neq 0$ for $1 \leq m < n \leq M_i$) [11]. From these local observables, we build p-body correlators of the form $\hat{O}_{\boldsymbol{m}} = \otimes_{i=1}^N \hat{O}_{m_i}^{(i)}$ where $m_i = 0$ for $N-p$ subsystems. Arbitrary observables can be built as linear combinations of correlators – such as *e.g.* powers of collective spin variables [12, 13] $\hat{J}_a = \sum_i \hat{\sigma}_a^{(i)}/2$ ($a = x, y, z$) for systems of qubits. Hence we shall assume that R observables of the form $\hat{A}_a = \sum_{\boldsymbol{m}} x_{\boldsymbol{m}}^{(a)} \hat{O}_{\boldsymbol{m}}$ can be measured, where the sum runs over all strings $\boldsymbol{m} = (m_1, \dots m_N)$, and $x_{\boldsymbol{m}}^{(a)}$ are arbitrary real coefficients. The quantum data $\{\langle \hat{A}_a \rangle_{\hat{\rho}}\}_{a=1}^R$ form the basis for entanglement certification in our scheme. The problem of entanglement certification based on a data set has been discussed in the past, but the proposed methods either lack scalability [6], or are scalable only under some restrictive assumptions (short-range correlations, low-dimensional geometry) [7]. Our method aims at surpassing these limitations.

Mapping onto an inverse statistical problem. The key aspect behind our approach is the limited information content of separable states. The parameters $\boldsymbol{\lambda}$ specifying the product state $\hat{\rho}_{\text{prod}}(\boldsymbol{\lambda})$ can indeed be chosen as $\sum_i (2d_i - 2) \sim \mathcal{O}(N)$ real parameters, where d_i is the dimension of the local Hilbert space of the i-th subsystem [14]. The average of the $\hat{A}_a$ observable on a separable state reads

$$\langle \hat{A}_a \rangle_{\hat{\rho}_p} = \int d\boldsymbol{\lambda}\, p(\boldsymbol{\lambda}) \mathcal{A}_a(\boldsymbol{\lambda}) =: \langle \mathcal{A}_a \rangle_p \tag{2}$$

where $\mathcal{A}_a(\boldsymbol{\lambda}) = \sum_{\boldsymbol{m}} x_{\boldsymbol{m}}^{(a)} \prod_{i=1}^N o_{m_i}^{(i)}(\boldsymbol{\lambda}_i)$ and $o_{m_i}^{(i)}(\boldsymbol{\lambda}_i) = \langle \psi_i(\boldsymbol{\lambda}_i) | \hat{O}_{m_i}^{(i)} | \psi_i(\boldsymbol{\lambda}_i) \rangle$. Given a product state, the calculation of each term in the sum defining $\mathcal{A}_a(\boldsymbol{\lambda})$ is clearly an operation scaling as $\mathcal{O}(N)$. Once the quantum nature of the state has been absorbed in $\mathcal{A}_a(\boldsymbol{\lambda})$, the calculation of $\langle \hat{A}_a \rangle_{\hat{\rho}_p}$, Eq. (2), is a classical statistical average over the distribution p which, from a statistical physics viewpoint, can be regarded as the Boltzmann distribution $p(\boldsymbol{\lambda}) =: \exp[-H(\boldsymbol{\lambda})]/\mathcal{Z}$ of a classical field theory on a lattice (normalized by the $\mathcal{Z}$ factor), with a vector field $\boldsymbol{\lambda}_i$ defined on each of the N clusters (Fig. 1(b)). The complexity of separable states is fundamentally inscribed in the effective Hamiltonian $H(\boldsymbol{\lambda})$, which is *a priori* arbitrary, namely it is specified by a number $\mathcal{O}(\exp(N))$ of parameters.

Once the classical statistical structure of the expectation values on separable states is exposed, the problem of reproducing the quantum data with a separable state takes the form of a statistical inference problem, whose solution is well known in statistical physics [15]. First of all, applying a maximum-entropy principle [16], the Hamiltonian can be parametrized without loss of generality with as many parameters as the elements of the quantum data set:

$$H(\boldsymbol{\lambda}) = -\sum_{a=1}^R K_a \mathcal{A}_a(\boldsymbol{\lambda}) \,. \tag{3}$$

The parameters $\boldsymbol{K} = \{K_a\}_{a=1}^R$ – the coupling constants of the classical field theory – are Lagrange multipliers whose optimization allows one to build the separable state $\hat{\rho}_p$ whose expectation values $\{\langle \hat{A}_a \rangle_{\hat{\rho}_p}\}$ best approximate the quantum data $\{\langle \hat{A}_a \rangle_{\hat{\rho}}\}$. The optimization of $\boldsymbol{K}$ can be efficiently achieved upon minimizing the cost function $L(\boldsymbol{K}) := \log \mathcal{Z}(\boldsymbol{K}) - \sum_a K_a \langle \hat{A}_a \rangle_{\hat{\rho}}$

[10, 15]. The a-th component of the gradient of L is $g_a := \frac{\partial L}{\partial K_a} = \langle \mathcal{A}_a \rangle_p - \langle \hat{A}_a \rangle_{\hat{\rho}}$, and its Hessian matrix is $\frac{\partial^2 L}{\partial K_a \partial K_b} = \langle \mathcal{A}_a \mathcal{A}_b \rangle_p - \langle \mathcal{A}_a \rangle_p \langle \mathcal{A}_b \rangle_p$, namely the covariance matrix of the $\mathcal{A}_a(\boldsymbol{\lambda})$ functions. Since the latter is a semi-definite positive matrix, L is a convex function. Therefore, a simple gradient-descent algorithm, which consists in iterating the update rule $K'_a = K_a - \epsilon[\langle \mathcal{A}_a \rangle_p - \langle \hat{A}_a \rangle_{\hat{\rho}}]$ with $\epsilon \ll 1$, or any improvement thereof, is guaranteed to reach the global optimum of the problem. In practice, this requires to repeatedly compute $\langle \mathcal{A}_a \rangle_p$ as in Eq. (2), a task efficiently accomplished *e.g.* by Markov-chain Monte Carlo sampling of $p(\boldsymbol{\lambda})$, whenever the Hamiltonian H does not describe a glassy system. The latter restriction to non-glassy systems is the only practical limitation of our approach. In fact, at a fundamental level, the computational intractability of statistical sums in classical field theory is only proven for Ising models, involving discrete variables [17]. On the other hand, the classical field theories that describe separable states depend on continuous variables. Furthermore, in the examples considered below, glassy behavior is excluded by considering translationally invariant systems.

Construction of an optimal entanglement witness. As illustrated on Fig. 1(c), the algorithm converges to the distribution p which minimizes $|\boldsymbol{g}|$ – the norm of the gradient of L. If the minimal distance $g^{(\min)}$ vanishes (within the error on the quantum data), i.e. if $\langle \hat{A}_a \rangle_{\hat{\rho}_p^{(\min)}} = \langle \hat{A}_a \rangle_{\hat{\rho}}$ for all $a = 1, \dots R$, then entanglement cannot be assessed from the available data. But in the opposite case, the coupling constants K_a increase indefinitely along the optimization, and the coefficients of the gradient $g_a^{(\min)} = \langle \hat{A}_a \rangle_{\hat{\rho}_p^{(\min)}} - \langle \hat{A}_a \rangle_{\hat{\rho}}$ allow us to build a violated EW inequality. First, we define the normalized coefficients $W_a := -g_a^{(\min)}/|\boldsymbol{g}^{(\min)}|$. The condition $|\boldsymbol{g}^{(\min)}|^2 > 0$ is then rewritten as:

$$-\sum_a W_a \langle \hat{A}_a \rangle_{\hat{\rho}} < \min_{\hat{\rho}_p} \left\{ -\sum_a W_a \langle \hat{A}_a \rangle_{\hat{\rho}_p} \right\} =: B_{\text{sep}} \quad (4)$$

The linear combination $\hat{\mathcal{W}} := -\sum_{a=1}^{R} W_a \hat{A}_a$ is the data-driven EW operator. The separable bound B_{sep}, namely the minimal value of $\text{Tr}(\hat{\rho}\hat{\mathcal{W}})$ over separable states, is violated by the data set, ultimately proving that entanglement is present among the subsystems. The operator $\hat{\mathcal{W}}$ is optimal, in that any other normalized linear combination $\hat{\mathcal{W}}' = -\sum_a W'_a \hat{A}_a$ defines an EW inequality whose violation cannot exceed the violation of the inequality involving $\hat{\mathcal{W}}$. This property follows from the convexity of the set of separable states.

Complexity of the algorithm. If the quantum data contain correlation functions involving up to k points, the effective Hamiltonian H contains $\mathcal{O}(N^k)$ terms; therefore the computational cost of evaluating statistical averages of the kind of Eq. (2) with a precision of ϵ (using Monte Carlo sampling) scales as $\mathcal{O}(d^2 \epsilon^{-2} N^k)$, where $\mathcal{O}(d^2)$ is the cost of evaluating the local observables $o_{m_i}^{(i)}(\boldsymbol{\lambda}_i)$ when $d_i = d$. The polynomial scaling of the computational cost with the number N of parties and with the local Hilbert space dimension is the central asset of our approach.

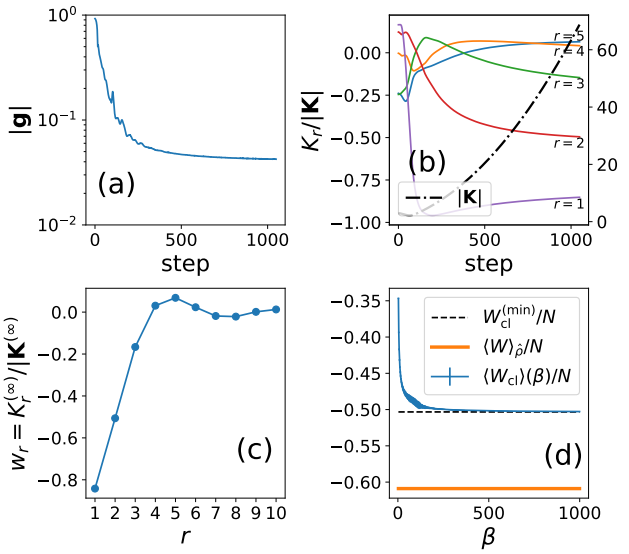

FIG. 2: Data-driven entanglement witness for the Heisenberg chain at $T/J = 1$. (a) Distance between the quantum data (all spin-spin correlators) and the optimized separable state ($\boldsymbol{g}$: gradient of the cost function), as a function of optimization steps in a Nesterov accelerated gradient descent ($\epsilon = 0.01$). Each step contains $10^5 - 10^7$ Monte Carlo steps to achieve a relative precision of 10% on the modulus of the gradient [18]. (b) Normalized coupling constants K_r in the classical Hamiltonian defining the separable state (solid lines, left axis), and overall amplitude $|\boldsymbol{K}|$ (dashed-dotted line, right axis). (c) Normalized couplings K_r at the end of the algorithm; (d) The separable bound can be obtained via simulated annealing by calculating $\langle \mathcal{W}_{\text{cl}} \rangle(\beta)$ against $\exp[-\beta \mathcal{W}_{\text{cl}}]$, ramping β from 0 to 1000. The minimum $\mathcal{W}_{\text{cl}}^{(min)}$ is actually the lowest value recorded for $\mathcal{W}_{\text{cl}}$ throughout the ramp.

Ensembles of qubits. Hereafter we shall specify our attention to the case of systems of N qubits partitioned into subsystems consisting of single qubits; and quantum data will be assumed to consist of one- and two-point correlations, $\langle \hat{\sigma}_a^{(i)} \rangle_{\hat{\rho}}$ and $\langle \hat{\sigma}_a^{(i)} \hat{\sigma}_b^{(j)} \rangle_{\hat{\rho}}$ respectively, fully specifying all one- and two-qubit reduced density matrices. Product states are parametrized by the orientations $\{\boldsymbol{\lambda}_i\} = \{\boldsymbol{n}^{(i)}\}$ of each qubit on the Bloch sphere (where $\boldsymbol{n}^{(i)}$ are unit vectors), so that the effective Hamiltonian describes *classical Heisenberg spins* (namely, rotators), coupled via bilinear interactions and immersed in an external field:

$$H(\{\boldsymbol{n}^{(i)}\}) = -\sum_{i=1}^{N} \sum_{a=x,y,z} K_a^{(i)} n_a^{(i)} - \sum_{i<j} \sum_{a,b} K_{ab}^{(ij)} n_a^{(i)} n_b^{(j)} .$$

Heisenberg antiferromagnetic chain. The first example of entangled states that we study with our approach is the thermal equilibrium state of the $S = 1/2$ Heisenberg

chain $\hat{H} = J\sum_{i=1}^{N} \hat{\boldsymbol{S}}^{(i)} \cdot \hat{\boldsymbol{S}}^{(i+1)}$, where $\hat{\boldsymbol{S}}^{(i)}$ are $S = 1/2$ spin operators, J is the exchange energy, and periodic boundary conditions (PBC) are assumed. Thermal equilibrium states $\hat{\rho}$ ($\propto \exp[-\hat{H}/k_B T]$) give $\langle \hat{\sigma}_a^{(i)} \rangle_{\hat{\rho}} = 0$ and $\langle \hat{\sigma}_a^{(i)} \hat{\sigma}_b^{(j)} \rangle_{\hat{\rho}} = \delta_{ab}\ C(|i-j|)$, due to rotational invariance of the spin-spin couplings and translational invariance. These elementary symmetries of the quantum data are directly inherited by the classical Hamiltonian defining separable states aimed at reproducing them. The Hamiltonian takes the form of a classical long-range Heisenberg model $H(\{\boldsymbol{n}^{(i)}\}) = -\sum_{i<j} K_{|i-j|} \boldsymbol{n}^{(i)} \cdot \boldsymbol{n}^{(j)}$ with $K_r = K_{N-r}$. The most effective existing multipartite entanglement criterion for this quantum data is based on the collective spin, namely $\langle \hat{\boldsymbol{J}}^2 \rangle = \sum_{ij} \langle \hat{\boldsymbol{S}}^{(i)} \cdot \hat{\boldsymbol{S}}^{(j)} \rangle < N/2$ [19, 20], which is verified for $t = T/J \lesssim 1.4$. This criterion is a permutationally invariant EW (PIEW), treating correlations at all distances on the same footing, and it cannot be optimal at sufficiently high temperatures, namely when the correlation length ξ becomes of the order of a few lattice spacings.

As a first validation of our approach, we search for the optimal EW based on two-body correlations $\langle \hat{\sigma}_a^{(i)} \hat{\sigma}_a^{(j)} \rangle$ by using as input quantum data the correlations (obtained via quantum Monte Carlo - QMC [18]) at $t = 1$ for $N = 64$ spins, at which $\xi = 0.72$. Because of their finite range we only used correlations up to a distance $r_{\max} = 10$. Fig. 2 illustrates the results of our approach. The saturation to a finite value of the distance between the quantum data and those of the optimized separable state (measured by the norm of the vector $\boldsymbol{g}$, see Fig. 2(a)) and the divergence of the couplings K_r (Fig. 2(b)) clearly indicate the success of entanglement witnessing. The optimal EW operator can be reconstructed in principle from the asymptotic value of the gradient vector $\boldsymbol{g}^{(\infty)}$ as $\hat{\mathcal{W}} = -\sum_{i=1}^{N} \sum_{a\in\{x,y,z\}} \sum_{r=1}^{r_{\max}} w_r\ \hat{\sigma}_a^{(i)} \hat{\sigma}_a^{(i+r)}$ with $w_r = -g_r^{(\infty)}/|\boldsymbol{g}^{(\infty)}|$. In practice, we found a more strongly violated EW inequality using the asymptotic couplings of the effective Hamiltonian, namely $w_r = K_r^{(\infty)}/|\boldsymbol{K}^{(\infty)}|$ – which display a clear spatial structure, shown in Fig. 2(c) (see [18] for the numerical values). The final step of the approach consists in determining the separable bound $B_{\text{sep}} = \min_{\hat{\rho}_p} \text{Tr}(\hat{\rho}_p \hat{\mathcal{W}})$, which can be readily obtained by finding the ground-state energy of the classical Hamiltonian $\mathcal{W}_{\text{cl}} = -\sum_{i=1}^{N} \sum_{r=1}^{r_{\max}} w_r\ \boldsymbol{n}^{(i)} \cdot \boldsymbol{n}^{(i+r)}$. This value can be numerically obtained via temperature annealing [21], as illustrated in Fig. 2(d). There we observe that $B_{\text{sep}}/N = -0.5032$, while the quantum data reach $\langle \hat{\mathcal{W}} \rangle_{\hat{\rho}}/N = -0.6089$. This represents a 21% violation of the separable bound, to be compared with that of $\langle \hat{\boldsymbol{J}}^2 \rangle_{\hat{\rho}}/N = 0.411 < 1/2$, offering a 18% violation of the bound. This result is *not* incremental, because the EW inequality we find is optimal among all those containing two-body correlators. Interestingly, for temperatures $t \gtrsim 1.4$ (at which the PIEW ceases to work) we found numerically impossible to prove that $\hat{\rho}(T)$ is entangled solely based on two-point correlators: this in turn shows that the maximal set of thermal states whose entanglement can be witnessed using two-point correlators is essentially captured by the PIEW. This will not be the case in our next example, in which our approach significantly extends the range of witnessed entangled thermal states.

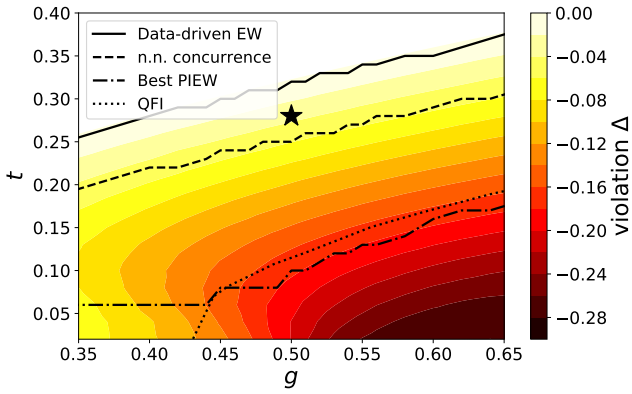

FIG. 3: Data-driven EW for the quantum Ising chain. Phase diagram around the QCP. The star corresponds to $t = 0.28$, $g = 0.5$, at which the quantum data used as input were calculated. The color represents the violation $\Delta = (\langle \hat{\mathcal{W}} \rangle_{\hat{\rho}} - B_{\text{sep}})/N$ of our data-driven EW. The various curves correspond to the temperature below which different entanglement criteria are satisfied (nearest-neighbour concurrence [22]; best PIEW [12]; and quantum Fisher information (QFI) of $\hat{J}_z$ [23].

Quantum Ising chain. Our final example is the quantum Ising model with Hamiltonian $\hat{H} = -J\sum_{i=1}^{N}(\hat{S}_z^{(i)}\hat{S}_z^{(i+1)} + g\hat{S}_x^{(i)})$, where J is the interaction strength and Jg the transverse field. In the ground state, the system displays a quantum critical point (QCP) at $g = g_c = 1/2$ between a ferromagnetic phase ($g < g_c$) and a paramagnetic phase ($g > g_c$) [24]. At finite temperature around the QCP, the system is known to exhibit robust entanglement [23, 25, 26]. Given the symmetries of the correlation functions ($\langle \hat{\sigma}_a^{(i)} \rangle_{\hat{\rho}} = 0$ for $a = y, z$; $\langle \hat{\sigma}_a^{(i)} \hat{\sigma}_b^{(j)} \rangle_{\hat{\rho}} \sim \delta_{ab}$), the classical Hamiltonian tailored to reproduce them is of the form: $H(\{\boldsymbol{n}^{(i)}\}) = -K_x \sum_{i=1}^{N} n_x^{(i)} - \sum_{a=x,y,z} \sum_{i<j} K_a^{|i-j|} n_a^{(i)} n_a^{(j)}$. As input quantum data, we consider the correlation functions of a chain of $N = 64$ spins with PBC at a temperature $t = T/J = 0.28$ for $g = 0.5$ - obtained as well via QMC. Given the finite correlation length, we only used correlators up to a distance $r_{\max} = 20$. Following the same procedure as described for the Heisenberg chain, we find an optimal EW operator which is spatially structured, of the form $\hat{\mathcal{W}} = -w_x \sum_{i=1}^{N} \hat{\sigma}_x^{(i)} - \sum_{a=x,y,z} \sum_{i<j} w_a^{(|i-j|)} \hat{\sigma}_a^{(i)} \hat{\sigma}_a^{(j)}$ (coefficients and separable bound in the Supplemental Material [18]). On Fig. 3, we show that this new EW criterion, optimal for the thermal state at $t = 0.28, g = 0.5$, allows one to prove entanglement for a larger set of ther-

mal states than all the existing criteria in the literature (namely the nearest-neighbour concurrence [22], the PIEW [12], and the quantum Fisher information [23] – see [18] for further details).

Conclusions. We introduced a data-driven method to probe multipartite entanglement in many-body systems. This method relies on mapping separable states onto Boltzmann distributions for a classical field theory on a lattice. The classical degrees of freedom of this field theory are dictated by the considered partitioning of the system. The structure of the corresponding classical Hamiltonian is dictated by the quantum data at hand; and its parameters are optimized in order to fit at best the quantum data. This method allows to exhaustively test the entanglement witnessing capability of a set of quantum data in a time scaling polynomially with the number of parties in the partition (if the size of quantum data is also polynomial); this is guaranteed whenever the classical field theory is not a model of a glass (namely when it does not feature disorder *and* frustration). This opens the way to the systematic certification of entanglement in intermediate-scale quantum devices.

We are very grateful to Antonio Acín for insightful discussions. IF acknowledges support from the Government of Spain (FIS2020-TRANQI and Severo Ochoa CEX2019-000910-S), Fundació Cellex and Fundació Mir-Puig through an ICFO-MPQ Postdoctoral Fellowship, Generalitat de Catalunya (CERCA, AGAUR SGR 1381 and QuantumCAT), and MINECO-EU QuantERA MAQS (funded by The State Research Agency (AEI) PCI2019-111828-2 / 10.13039/501100011033). TR acknowledges support from ANR (EELS project) and QuantERA (MAQS project). Numerical computations have been performed at the Pôle Scientifique de Modélisation Numérique (PSMN).

* Electronic address: irenee.frerot@icfo.eu
† Electronic address: tommaso.roscilde@ens-lyon.fr

[1] Georgescu, I. M., Ashhab, S. & Nori, F. Quantum simulation. *Rev. Mod. Phys.* **86**, 153–185 (2014). URL https://link.aps.org/doi/10.1103/RevModPhys.86.153.

[2] Preskill, J. Quantum computing and the entanglement frontier (2012). 1203.5813.

[3] Fröwis, F., Sekatski, P., Dür, W., Gisin, N. & Sangouard, N. Macroscopic quantum states: Measures, fragility, and implementations. *Rev. Mod. Phys.* **90**, 025004 (2018). URL https://link.aps.org/doi/10.1103/RevModPhys.90.025004.

[4] Paris, M. & Rehacek, J. *Quantum State Estimation.* Lecture Notes in Physics (Springer Berlin Heidelberg, 2004). URL https://www.springer.com/gp/book/9783540223290.

[5] Doherty, A. C., Parrilo, P. A. & Spedalieri, F. M. Detecting multipartite entanglement. *Phys. Rev. A* **71**, 032333 (2005). URL https://link.aps.org/doi/10.1103/PhysRevA.71.032333.

[6] Gühne, O., Reimpell, M. & Werner, R. F. Lower bounds on entanglement measures from incomplete information. *Phys. Rev. A* **77**, 052317 (2008). URL https://link.aps.org/doi/10.1103/PhysRevA.77.052317.

[7] Navascues, M., Baccari, F. & Acín, A. Entanglement marginal problems (2020). 2006.09064.

[8] Werner, R. F. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Phys. Rev. A* **40**, 4277–4281 (1989). URL http://link.aps.org/doi/10.1103/PhysRevA.40.4277.

[9] Gühne, O. & Tóth, G. Entanglement detection. *Physics Reports* **474**, 1–75 (2009). URL http://www.sciencedirect.com/science/article/pii/S0370157309000623.

[10] Frérot, I. & Roscilde, T. Detecting many-body bell nonlocality by solving ising models (2020). URL https://arxiv.org/abs/2004.07796. 2004.07796.

[11] Otherwise one can always reproduce the correlations among these observables using separable states by choosing the local states $|\psi_i\rangle$ to be joint eigenstates of the observables, and reproducing with p the classical correlations among the observables.

[12] Tóth, G., Knapp, C., Gühne, O. & Briegel, H. J. Spin squeezing and entanglement. *Phys. Rev. A* **79**, 042334 (2009). URL https://link.aps.org/doi/10.1103/PhysRevA.79.042334.

[13] Vitagliano, G., Hyllus, P., Egusquiza, I. n. L. & Tóth, G. Spin squeezing inequalities for arbitrary spin. *Phys. Rev. Lett.* **107**, 240502 (2011). URL https://link.aps.org/doi/10.1103/PhysRevLett.107.240502.

[14] d_i complex amplitudes, up to normalization of the state and an arbitrary global phase.

[15] Nguyen, H. C., Zecchina, R. & Berg, J. Inverse statistical problems: from the inverse Ising problem to data science. *Advances in Physics* **66**, 197–261 (2017). URL https://doi.org/10.1080/00018732.2017.1341604.

[16] Jaynes, E. T. Information theory and statistical mechanics. *Phys. Rev.* **106**, 620–630 (1957). URL https://link.aps.org/doi/10.1103/PhysRev.106.620.

[17] Barahona, F. On the computational complexity of ising spin glass models. *Journal of Physics A: Mathematical and General* **15**, 3241–3253 (1982). URL https://doi.org/10.1088%2F0305-4470%2F15%2F10%2F028.

[18] See Supplemental Material (SM) for a discussion of: 1) further details of the optimization algorithm; 2) the quantum data from Quantum Monte Carlo; and 3) the existing entanglement witnesses for the quantum Ising chain.

[19] Tóth, G. Entanglement detection in optical lattices of bosonic atoms with collective measurements. *Phys. Rev. A* **69**, 052327 (2004). URL https://link.aps.org/doi/10.1103/PhysRevA.69.052327.

[20] Wieśniak, M., Vedral, V. & Brukner, Č. Magnetic susceptibility as a macroscopic entanglement witness. *New Journal of Physics* **7**, 258–258 (2005). URL https://doi.org/10.1088%2F1367-2630%2F7%2F1%2F258.

[21] Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. Optimization by simulated annealing. *Science* **220**, 671–680 (1983). URL https://science.sciencemag.org/content/220/4598/671. https://science.sciencemag.org/content/220/4598/671.full.pdf.

[22] Wootters, W. K. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.* **80**,

2245–2248 (1998). URL https://link.aps.org/doi/10.1103/PhysRevLett.80.2245.

[23] Hauke, P., Heyl, M., Tagliacozzo, L. & Zoller, P. Measuring multipartite entanglement through dynamic susceptibilities. *Nat Phys* **12**, 778–782 (2016). URL http://dx.doi.org/10.1038/nphys3700.

[24] Sachdev, S. *Quantum Phase Transitions* (Cambridge University Press, 2011). URL https://books.google.fr/books?id=F3IkpxwpqSgC.

[25] Frérot, I. & Roscilde, T. Quantum critical metrology. *Phys. Rev. Lett.* **121**, 020402 (2018). URL https://link.aps.org/doi/10.1103/PhysRevLett.121.020402.

[26] Frérot, I. & Roscilde, T. Reconstructing the quantum critical fan of strongly correlated systems via quantum correlations. *arXiv:1805.03140 [cond-mat, physics:quant-ph]* (2018). URL http://arxiv.org/abs/1805.03140. ArXiv: 1805.03140.

[27] Creutz, M. Overrelaxation and monte carlo simulation. *Phys. Rev. D* **36**, 515–519 (1987). URL https://link.aps.org/doi/10.1103/PhysRevD.36.515.

[28] Syljuåsen, O. F. & Sandvik, A. W. Quantum monte carlo with directed loops. *Phys. Rev. E* **66**, 046701 (2002). URL https://link.aps.org/doi/10.1103/PhysRevE.66.046701.

[29] Boll, M. *et al.* Spin- and density-resolved microscopy of antiferromagnetic correlations in fermi-hubbard chains. *Science* **353**, 1257–1260 (2016). URL https://science.sciencemag.org/content/353/6305/1257. https://science.sciencemag.org/content/353/6305/1257.full.pdf.

[30] Browaeys, A. & Lahaye, T. Many-body physics with individually controlled rydberg atoms. *Nature Physics* **16**, 132–142 (2020). URL https://doi.org/10.1038/s41567-019-0733-z.

[31] Pezzè, L. & Smerzi, A. Quantum theory of phase estimation. In Tino, G. & Kasevich, M. (eds.) *Proceedings of the International School of Physics "Enrico Fermi"*, 691 (IOS Press, Amsterdam, 2014). URL https://arxiv.org/abs/1411.5164.

Supplemental Material

In this Supplemental Material, we provide: 1) further technical details on the variational algorithm described and implemented for the data presented in the main text; 2) on the generation of quantum data, used as input to our algorithm, by quantum Monte Carlo; 3) on the comparison with existing entanglement criteria. In the attached `.csv` files, the numerical coefficients of the entanglement witnesses discussed in the main text are given.

Details on the algorithm

In the main text, we introduced a variational algorithm to fit a given data set of expectation values by using separable states, represented as Boltzmann distributions over classical Heisenberg spins $\boldsymbol{n}^{(i)}$ on the unit sphere. In the examples discussed in the main text, the data set contains one-qubit expectation values $\langle\hat{\sigma}_a^{(i)}\rangle_{\hat{\rho}}$ and two-qubit correlations $\langle\hat{\sigma}_a^{(i)}\hat{\sigma}_b^{(j)}\rangle_{\hat{\rho}}$. In the examples we considered (namely, the one-dimensional antiferromagnetic Heisenberg model, and the Ising model in a transverse field, both with periodic boundary conditions), correlations $\langle\hat{\sigma}_a^{(i)}\hat{\sigma}_b^{(j)}\rangle_{\hat{\rho}}$ vanish if $a \neq b$. Since we used translationally invariant chains (with periodic boundary conditions), the one-qubit data reduces to the average magnetization $\langle m_a\rangle_{\hat{\rho}} = \sum_{i=1}^N \langle\hat{\sigma}_a^{(i)}\rangle_{\hat{\rho}}/N$, and the the two-qubit correlations depend only on the inter-qubit distance: $\langle C_a^{(r)}\rangle_{\hat{\rho}} = \langle\hat{\sigma}_a^{(i)}\hat{\sigma}_a^{(i+r)}\rangle_{\hat{\rho}}$. In the case of the Heisenberg model, which displays $SU(2)$ invariance, we have $m_a = 0$. In this case, we considered as quantum data $\langle C^{(r)}\rangle_{\hat{\rho}} = \langle C_x^{(r)} + C_y^{(r)} + C_z^{(r)}\rangle_{\hat{\rho}}$.

Correspondingly, the classical Hamiltonian aiming at reproducing the quantum data contains one- and two-body interactions terms (the latter truncated beyond a given distance $r_{\max}$). For the Heisenberg model, we get

$$H = -\sum_i \sum_{r=1}^{r_{\max}} K^{(r)} \boldsymbol{n}^{(i)} \cdot \boldsymbol{n}^{(i+r)} \; ; \tag{5}$$

while for the quantum Ising model, where $m_y = m_z = 0$, we have

$$H = -K_x \sum_{i=1}^N n_x^{(i)} - \sum_{a=x,y,z} \sum_i \sum_{r=1}^{r_{\max}} K_a^{(r)} n_a^{(i)} n_a^{(i+r)} \; . \tag{6}$$

The K's coefficients are the variational parameters of our algorithm, which are optimized in an iterative manner. A simple gradient-descent algorithm consists in iterating the following update rule (for the Ising model):

$$K_x' = K_x - \epsilon[\langle m_x\rangle_p - \langle m_x\rangle_{\hat{\rho}}] \tag{7}$$

$$(K_a^{(r)})' = K_a^{(r)} - \epsilon[\langle C_a^{(r)}\rangle_p - \langle C_a^{(r)}\rangle_{\hat{\rho}}] \tag{8}$$

for $a \in \{x, y, z\}$, and $r \in \{1, 2, \cdots N/2\}$; and (for the Heisenberg model):

$$(K^{(r)})' = K^{(r)} - \epsilon[\langle C^{(r)}\rangle_p - \langle C^{(r)}\rangle_{\hat{\rho}}] \; . \tag{9}$$

In the above equations, $\langle\cdot\rangle_p$ is the expectation value on the Boltzmann distribution for the classical Hamiltonian (whose couplings are the K's coefficients), while $\langle\cdot\rangle_{\hat{\rho}}$ are the target quantum data. As discussed in the main text (see also [10]), ϵ is a small parameter, implementing a numerical gradient descent of the (convex) L function. In practice, we implemented the accelerated gradient-descent algorithm of Nesterov (NAG), with $\epsilon = 0.01$.

Each step of the NAG algorithm requires to compute the Euclidean distance $\boldsymbol{g}$ between the separable data and the quantum data, namely to compute $\langle m_x\rangle_p$ and $\langle C_a^{(r)}\rangle_p$ for the Ising model and $\langle C^{(r)}\rangle_p$ for the Heisenberg model. This was implemented using Markov-chain Monte Carlo. The number of Monte Carlo steps (defined below) implemented at each step of the NAG algorithm was chosen such that the relative error on $\boldsymbol{g}$ be smaller than a given threshold η, which we chose as $\eta = 0.05$ for the Ising

model, and $\eta = 0.1$ for the Heisenberg model. In other words, one step of the NAG algorithm is completed when:

$$\frac{2\sum_\alpha |g_\alpha| \, \mathrm{Err}(g_\alpha)}{|\boldsymbol{g}|^2} < \eta \ , \tag{10}$$

where $\mathrm{Err}(g_\alpha)$ is the error on g_α, as estimated from the Monte Carlo algorithm. Each step of the Monte Carlo algorihm consisted of $2N$ iterations of single-spin Metropolis updates and of single-spin microcanonical overrelaxation updates [27]. The amplitude of the proposed Metropolis updates was adapted along the Monte Carlo simulation so that the move be accepted with frequency 0.5 ± 0.1. Therefore, a single Monte Carlo step consists of $2N$ microcanonical updates, and of N accepted Metropolis updates (on average).

As the variational optimization of the K's parameters progresses along the NAG algorithm, the norm of the gradient $\boldsymbol{g}$ decreases, and therefore an increasing number of Monte Carlo steps is required at each step of the NAG algorithm in order to achieve the required relative precision of η. When the quantum data cannot be fitted by a separable state, $\boldsymbol{g}$ stabilizes to a finite value. The number of steps of the NAG algorithm to achieve this convergence (and therefore the total number of Monte Carlo steps along the whole optimization) depends on the value of $|\boldsymbol{g}|$ as obtained at the end of the optimization. For the examples presented in the main text, about 10^3 steps of the NAG algorithm were necessary, each of them comprising $10^4 \div 10^7$ Monte Carlo steps.

Quantum data from Quantum Monte Carlo

Data-driven entanglement witnessing is fundamentally based on reliable quantum data on quantum many-body systems. Here we chose to use quantum Monte Carlo data for quantum spin chains at finite temperature, obtained using Stochastic Series Expansion [28], which provides numerically exact correlation functions for the model of interest (within the statistical error bar). Finite-temperature equilibrium calculations offer the most reliable source of data for mixed states – which pose the real challenge in terms of entanglement detection, while for pure states any form of connected correlation is an entanglement witness. Beyond their significance in condensed matter physics and quantum statistical physics, the models we chose (quantum Heisenberg and quantum Ising chain) are also of direct relevance to several experiments in quantum simulation, see e.g. [29, 30] for recent examples.

Existing entanglement witnesses

In this section, we provide additional details on the existing entanglement witnesses against which the quantum data of the quantum Ising model were tested (Fig. 3 of the main text).

Concurrence. The concurrence [22] defines a necessary and sufficient condition for the separability of a two-qubits density matrix. We computed the concurrence between nearest-neighbours, after reconstructing the density matrix $\hat{\rho}_{12}$ from the knowledge of one- and two-qubits expectation values $\langle \hat{\sigma}_a^{(1)} \rangle_{\hat{\rho}}$, $\langle \hat{\sigma}_b^{(2)} \rangle_{\hat{\rho}}$ and $\langle \hat{\sigma}_a^{(1)} \hat{\sigma}_b^{(2)} \rangle_{\hat{\rho}}$ (with $a, b \in \{x, y, z\}$) [4]. The dashed line on Fig. 3 defines the temperature below which $\hat{\rho}_{12}$ is entangled. Since the concurrence criterion [22] is based on a subset of the full quantum data we considered (which contains all one- and two-qubits correlations functions, which is equivalent to all two-body reduced density matrices $\hat{\rho}_{ij}$, and not only $\hat{\rho}_{12}$), by construction our data-driven method must detect entanglement in a region of the phase diagram strictly larger than the one detected by the concurrence – a fact clearly visible on Fig. 3.

Permutationally-invariant entanglement witnesses. In Ref. [12], a complete family of 8 entanglement witnesses based on the two-qubits reduced density matrix averaged over all pairs, $\hat{\rho}_{\mathrm{av},2} = 2\sum_{i \neq j} \hat{\rho}_{ij}/[N(N-1)]$, was derived. Equivalently, $\hat{\rho}_{\mathrm{av},2}$ is reconstructed from the knowledge of all one- and two-body correlations averaged over all permutations: $m_a := \sum_{i=1}^N \langle \hat{\sigma}_a^{(i)} \rangle_{\hat{\rho}}$ and $C_{ab} := \sum_{i \neq j} \langle \hat{\sigma}_a^{(i)} \hat{\sigma}_b^{(i)} \rangle_{\hat{\rho}}$. Since m_a and C_{ab} are coarse-grained features of the quantum data we have considered, if an EW inequality is violated by m_a and C_{ab} (namely if one of the 8 EW inequalities of ref. [12] is violated), then our data-driven algorithm must also reconstruct a violated entanglement witnesses – in general, a more strongly violated one. As illustrated on Fig. 3 for the quantum Ising model, for which we tested all 8 criteria for each parameters (t, g) (temperature and transverse field), this is clearly the case.

Quantum Fisher information. The quantum Fisher information (QFI) is another multipartite entanglement witness. Formally, the QFI quantifies the sensitivity of the state ρ to unitary transformations $\hat{\rho}(\phi) = e^{-i\phi\hat{O}} \hat{\rho} e^{i\phi\hat{O}}$ with $\hat{O}$ a quantum observable [31]. The QFI can be expressed as $\mathrm{QFI}(\hat{O}, \hat{\rho}) = 2\sum_{n \neq m} (p_n - p_m)^2 |\langle n|\hat{O}|m\rangle|^2/(p_n + p_m)$, where $\hat{\rho}$ is diagonalized as $\hat{\rho} = \sum_n p_n |n\rangle\langle n|$. Here, we chose for $\hat{O}$ the collective spin $J_z = \sum_{i=1}^N \hat{\sigma}_z^{(i)}/2$, which is optimal to witness entanglement around the quantum critical point of the quantum Ising model [23, 25]. The inequality $\mathrm{QFI}(\hat{J}_z, \hat{\rho}) \leq N$ is satisfied by all separable states, so that a QFI exceeding the system size is an entanglement witness [31]. In general, computing the QFI involves the knowledge of the full density matrix ρ, but the mapping of the quantum Ising chain onto a free-fermion model [24] makes this computation tractable [23]. Notice that computing the

QFI requires knowledge beyond one- and two-body correlators, and therefore it goes beyond the data set we have considered. Hence there is no guarantee a priori that our method exceeds the EW capability of the QFI. Nevertheless, as illustrated on Fig. 3, the parameter region where entanglement is detected by the QFI is broadly included in the region where entanglement is detected via our data-driven algorithm.

Detailed numerical values of the entanglement witnesses

The numerical coefficients of the entanglement witnesses reconstructed by our algorithm are given in this Section. For the Heisenberg model at temperature $T/J = 1$ (Fig. 2 of the main text), we discarded the correlations at distances beyond $r = 10$, and we found:

r	$K^{(r)}$
1	-0.84192229
2	-0.50632705
3	-0.16643027
4	0.03072345
5	0.06820392
6	0.02322838
7	-0.01850183
8	-0.02146924
9	0.00126596
10	0.01267421

These values are given also in the attached file `ddEW_Heisenberg_t1.00.csv`. The separable bound is $E/N = -0.503248446$. For the quantum Ising model (Fig. 3 of the main text), the corresponding coefficients K_x, $K_x^{(r)}$, $K_y^{(r)}$ and $K_z^{(r)}$, together with the separable bound, are given in the file `ddEW_Ising_t0.28_g0.5.csv`.