Sampling scheme for neuromorphic simulation of entangled quantum systems

Stefanie Czischek,[1] Jan M. Pawlowski,[2] Thomas Gasenzer,[1] and Martin Gärttner[1]

[1]*Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany*
[2]*Institut für Theoretische Physik, Ruprecht-Karls-Universität Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany*
(Dated: July 31, 2019)

Due to the complexity of the space of quantum many-body states the computation of expectation values by statistical sampling is, in general, a hard task. Neural network representations of such quantum states which can be physically implemented by neuromorphic hardware could enable efficient sampling. A scheme is proposed which leverages this capability to speed up sampling from so-called neural quantum states encoded by a restricted Boltzmann machine. Due to the complex network parameters a direct hardware implementation is not feasible. We overcome this problem by considering a phase reweighting scheme for sampling expectation values of observables. Applying our method to a set of paradigmatic entangled quantum states we find that, in general, the phase-reweighted sampling is subject to a form of sign problem, which renders the sampling computationally costly. The use of neuromorphic chips could allow reducing computation times and thereby extend the range of tractable system sizes.

I. INTRODUCTION

Simulating quantum many-body systems is considered a hard task for classical computers due to the exponentially growing Hilbert space dimension with system size. While some models, like the transverse-field Ising model, can be solved analytically [1–5], various approximative simulation methods exist. Among the most successful are the density matrix renormalization group, which is based on tensor network states [6–11], quantum Monte Carlo methods [12, 13], or semi-classical phase-space methods [14–17]. However, all these methods exploit specific properties of the quantum states. This limits their range of applicability and requires some a priori knowledge about the system. Quantum simulation of systems by other quantum systems [18–23] as an alternative is plagued by the fragility of quantum states due to decoherence.

Here we explore further a new route leading beyond standard von-Neumann architectures which makes use of brain-inspired approaches [24–26]. Among these, neuromorphic chips emulate neural networks by means of spiking neurons implemented on classical analog hardware. These chips have been shown to enable an efficient implementation of sampling from Boltzmann distributions [25–27]. The sampling on neuromorphic chips can yield a speed-up of at least one order of magnitude compared to classical computers while consuming about three orders of magnitude less energy [28]. The dynamical process performed by the neuromorphic network can be related to Langevin sampling of spin systems [29] which opens a path to the representation of quantum many-body states on the classical hardware.

Recently, parametrizations of quantum many-body states in terms of a restricted Boltzmann machine (RBM) network topology have been proposed [30, 31] and further studied [32–45]. These RBM representations of quantum states, in general, involve complex network parameters (weights and biases) due to the necessity to account for quantum superposition and interference [30, 46, 47]. As a result, the RBM does not resemble a positive definite Boltzmann probability distribution but rather the measure of a Feynman path integral. Therefore, standard RBM sampling and training methods, implementable on neuromorphic chips, cannot be applied.

Here we propose a phase reweighting scheme which absorbs the complex phases into the sampled observables [48–53], while the remaining amplitudes give a Boltzmann distribution of the spin configurations. This enables an implementation of a sampling process from this distribution on neuromorphic hardware to approximately evaluate expectation values according to a quantum Monte Carlo method.

Moreover, an extension of the parametrization to deep Boltzmann machines with multiple hidden layers becomes possible. In contrast to two-layer RBMs with complex weights which require the analytic summation of the hidden spins [30], these networks can be Gibbs sampled. This is crucial since an RBM with a single hidden layer allows to extract directly only diagonal spin observables from the sampled visible spin configurations. Here we devise an extension to deep networks which makes the calculation of arbitrary expectation values of spin operators possible from the visible layer of the network.

In general, the advantages of the reweighting scheme come at the expense of a form of sign problem which renders the sampling costly. However, we argue that neuromorphic hardware implementations could still provide a proportional speedup compared to standard architectures [28]. To benchmark our method, we consider the transverse-field Ising model (TFIM), which is a one-dimensional spin-1/2 chain with nearest-neighbor Ising interactions in a transverse field. The model describes a quantum phase transition between a para- and a ferromagnetic phase, controlled by the relative strength of the interactions and the transverse-field strength. We apply the deep Boltzmann machine representation to perform measurements of the order parameter and correlations in the ground state of the TFIM at the quantum phase transition, which is known to exhibit non-trivial quantum correlations.

To further analyze the ability of the RBM representation to capture quantum mechanical effects, we consider spin systems in strongly entangled states. As paradigmatic examples we choose the Bell state of two spins [54–56], as well as its generalization to larger spin systems, the Greenberger-Horne-Zeilinger (GHZ) states [57].

If the network parameters are purely real, which is the case for measurements in the computational basis (z-basis) for the ground state of the TFIM, we find sampling to be efficient, in the sense that the number of random samples needed in order for the expectation values to be converged is almost independent of the system size. Measurements in other bases require the extension to a deep network. For Bell and GHZ states, where complex network parameters are involved, the number of samples required to reach converged results scales exponentially with the size of the network. This is a manifestation of the sign problem [48–51, 53, 58, 59], which is due to the fact that the phases that need to be averaged in the phase reweighting scheme fluctuate heavily. This leads to strongly growing variances in the statistical sampling process (see Sect. V B). As a result, our method can be used to encode quantum states of limited size, also maximally entangled ones, in classical networks and sample from them, which can be realized efficiently by neuromorphic hardware.

Applications of RBM representation of quantum states include the simulation of ground states and dynamics in closed and open quantum many-body systems [30–38, 43, 45, 60–64] and efficient state tomography by learning the parameters of the state from experimental data [46, 59, 65]. All these applications could benefit from the use of neuromorphic hardware to more efficiently perform learning and sampling tasks.

After having introduced the RBM ansatz for parametrizing quantum states of many-spin systems (Sec. II), we present the phase reweighting scheme (Sec. III) and extensions to deep networks for measuring off-diagonal observables (Sec. IV). In Secs. V and VI, we apply the method to ground states of the TFIM and Bell and GHZ states, respectively, analyzing convergence and sampling efficiency. We draw conclusions in Sec. VII.

II. REPRESENTING QUANTUM STATES WITH RESTRICTED BOLTZMANN MACHINES

The state vector of an array of N spin-1/2 objects can be expressed in terms of product basis states $|\mathbf{v}^z\rangle = |v_1^z\rangle \otimes \cdots \otimes |v_N^z\rangle$, where $v_i^z = \pm 1$ and $|v_i^z\rangle$ are the eigenstates of the Pauli operator σ^z. The many-body wave-function is thus represented by complex coefficients $c_{\mathbf{v}^z} \in \mathbb{C}$ as

$$|\Psi\rangle = \sum_{\{\mathbf{v}^z\}} c_{\mathbf{v}^z} |\mathbf{v}^z\rangle, \qquad (1)$$
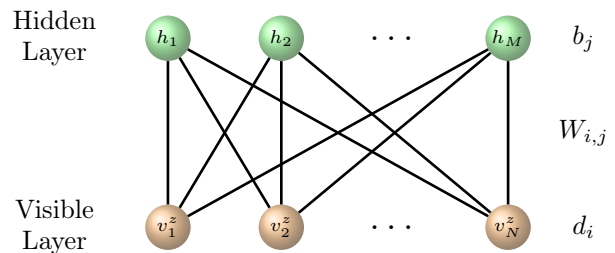

FIG. 1. Restricted Boltzmann machine (RBM) network with N visible neurons $\mathbf{v}^z = (v_1^z, \ldots, v_N^z)$ and M hidden neurons $\mathbf{h}$, cf. Eqs. (2), (3). Each visible neuron v_i^z is connected to each hidden neuron h_j via the weight $W_{i,j}$ and has a bias d_i, while the hidden neurons h_j have biases b_j.

where the sum runs over all basis states. The number of coefficients needed is 2^N, thus scaling exponentially in system size. In order to reduce the number of parameters needed to represent the quantum state to a tractable size, the coefficients $c_{\mathbf{v}^z}$ can be parametrized by means of an artificial neural network, specifically of a restricted Boltzmann machine (RBM) [30, 31, 66].

An RBM consists of N visible and M hidden binary neurons, $v_i^z, h_j \in \{\pm 1\}$, where the visible neurons correspond to the physical spins in the quantum state representation. Each visible neuron v_i^z is connected to each hidden neuron h_j via a weight $W_{i,j}$ and each neuron has an additional bias, d_i for the visible and b_j for the hidden neurons, as illustrated in Fig. 1. This yields the network energy function

$$E_{\mathrm{RBM}}(\mathbf{v}^z, \mathbf{h}; \mathcal{W}) = -\sum_{i=1}^{N}\sum_{j=1}^{M} v_i^z W_{i,j} h_j - \sum_{i=1}^{N} v_i^z d_i \\ -\sum_{j=1}^{M} h_j b_j, \qquad (2)$$

with the set of all weights and biases $\mathcal{W} = (\mathbf{d}, \mathbf{b}, W)$ [67]. An RBM-based parametrization of the unnormalized coefficients $c_{\mathbf{v}^z}$ can be defined as [30, 66]

$$c_{\mathbf{v}^z}(\mathcal{W}) = \sum_{\{\mathbf{h}\}} \exp\left[-E_{\mathrm{RBM}}(\mathbf{v}^z, \mathbf{h}; \mathcal{W})\right]. \qquad (3)$$

While for real-valued network parameters this corresponds to the marginal of a Boltzmann distribution over all neurons of the network, the weights and biases in general need to be complex in order to account for the complex basis projections $c_{\mathbf{v}^z}$ of a quantum state [30, 66].

An unnormalized probability distribution over the visible neurons is instead given by $|c_{\mathbf{v}^z}(\mathcal{W})|^2$, which can be evaluated after the summation over the binary hidden neurons is carried out, giving [30]

$$c_{\mathbf{v}^z}(\mathcal{W}) = \exp\left[\sum_{i=1}^{N} d_i v_i^z\right] \prod_{j=1}^{M} 2\cosh\left[\sum_{i=1}^{N} v_i^z W_{i,j} + b_j\right]. \qquad (4)$$

With this, expectation values of operators that are diagonal in the chosen basis [z-basis, cf. Eq. (1)], $\mathcal{O}^{\mathrm{diag}}$, can be written as

$$
\begin{aligned}
\left\langle \mathcal{O}^{\mathrm{diag}} \right\rangle &= \frac{1}{Z\left(\mathcal{W}\right)} \sum_{\{\mathbf{v}^z\}} \mathcal{O}^{\mathrm{diag}}\left(\mathbf{v}^z\right) \left|c_{\mathbf{v}^z}\left(\mathcal{W}\right)\right|^2 \\
&\approx \frac{1}{Q} \sum_{q=1}^{Q} \mathcal{O}^{\mathrm{diag}}\left(\mathbf{v}_q^z\right),
\end{aligned}
\tag{5}
$$

normalized by

$$
Z\left(\mathcal{W}\right) = \sum_{\{\mathbf{v}^z\}} \left|c_{\mathbf{v}^z}\left(\mathcal{W}\right)\right|^2.
\tag{6}
$$

In the second line of Eq. (5) we approximate the mean value in terms of a sum over Q samples of visible spin configurations $\mathbf{v}^z$ drawn according to the probability distribution $|c_{\mathbf{v}^z}(\mathcal{W})|^2$ via a Metropolis-Hastings algorithm and evaluate $\mathcal{O}^{\mathrm{diag}}(\mathbf{v}_q^z)$ on these samples [30]. The sampling error can be estimated via the variance, as further discussed in Sect. VI B.

Non-diagonal operators, such as the total magnetization $\sum_i \sigma_i^x$, can be evaluated similarly, exploiting their sparsity [30]:

$$
\begin{aligned}
\left\langle \mathcal{O} \right\rangle &= \frac{1}{Z\left(\mathcal{W}\right)} \sum_{\{\mathbf{v}^z\}} \sum_{\{\tilde{\mathbf{v}}^z\}} \left\langle \tilde{\mathbf{v}}^z \left| \mathcal{O} \right| \mathbf{v}^z \right\rangle c_{\mathbf{v}^z}\left(\mathcal{W}\right) c_{\tilde{\mathbf{v}}^z}^*\left(\mathcal{W}\right) \\
&= \frac{1}{Z\left(\mathcal{W}\right)} \sum_{\{\mathbf{v}^z\}} \mathcal{O}^{\mathrm{loc}}\left(\mathbf{v}^z\right) \left|c_{\mathbf{v}^z}\left(\mathcal{W}\right)\right|^2 \\
&\approx \frac{1}{Q} \sum_{q=1}^{Q} \mathcal{O}^{\mathrm{loc}}\left(\mathbf{v}_q^z\right),
\end{aligned}
\tag{7}
$$

where we introduce the local operator

$$
\mathcal{O}^{\mathrm{loc}}\left(\mathbf{v}^z\right) = \sum_{\{\tilde{\mathbf{v}}^z\}} \left\langle \tilde{\mathbf{v}}^z \left| \mathcal{O} \right| \mathbf{v}^z \right\rangle \frac{c_{\tilde{\mathbf{v}}^z}^*\left(\mathcal{W}\right)}{c_{\mathbf{v}^z}^*\left(\mathcal{W}\right)},
\tag{8}
$$

with the star denoting complex conjugation. This expression can be evaluated efficiently if $\mathcal{O}$ is sparse, i.e. $\langle \mathbf{v}^z|\mathcal{O}|\tilde{\mathbf{v}}^z\rangle$ is only non-vanishing for a number of matrix elements that scales polynomially in the number of spins. Physically relevant observables have this property.

The weights and biases in the RBM parametrization are variational parameters which can be adapted to represent a desired wave function. These weights can be found via a variational ansatz, where commonly a stochastic reconfiguration method is used to find ground state representations via energy minimization [30]. In some cases the network parameters can be found analytically as is the case for Bell and GHZ states which we will use below.

In the case of complex network parameters, Eq. (3) can no longer be interpreted as the marginal of a Boltzmann distribution. The exponential factors that are summed are complex and thus do not represent probabilities. In particular, this means that the intuitive procedure for evaluating observables – sampling configurations of all neurons in the network, including hidden ones, and then averaging the values of the observables obtained from the states of the visible neurons – is not applicable.

This motivates us to reformulate the procedure of evaluating observables in a way that allows sampling from a Boltzmann distribution while complex phases are absorbed into the diagonal elements of the considered observable. On the one hand, this enables an extension to deep neural networks with multiple hidden layers. On the other hand, an implementation on neuromorphic hardware setups becomes possible, which is known to efficiently sample from Boltzmann distributions [25–27].

III. PHASE REWEIGHTING SCHEME

To enable a sampling of visible and hidden neurons in the complex RBM from Boltzmann distributions, we consider the exponential of the network energy stated in Eq. (2). We split the weights and biases into real and imaginary parts,

$$
\begin{aligned}
&\exp\left[-E_{\mathrm{RBM}}\left(\mathbf{v}^z, \mathbf{h}; \mathcal{W}\right)\right] \\
&= \exp\left[\sum_{i=1}^{N}\sum_{j=1}^{M} v_i^z W_{i,j}^{\mathrm{R}} h_j + \sum_{i=1}^{N} d_i^{\mathrm{R}} v_i^z + \sum_{j=1}^{M} b_j^{\mathrm{R}} h_j\right] \\
&\quad \times \exp\left[i\left(\sum_{i=1}^{N}\sum_{j=1}^{M} v_i^z W_{i,j}^{\mathrm{I}} h_j + \sum_{i=1}^{N} d_i^{\mathrm{I}} v_i^z + \sum_{j=1}^{M} b_j^{\mathrm{I}} h_j\right)\right] \\
&=: \tilde{P}\left(\mathbf{v}^z, \mathbf{h}; \mathcal{W}^{\mathrm{R}}\right) e^{i\tilde{\varphi}\left(\mathbf{v}^z, \mathbf{h}; \mathcal{W}^{\mathrm{I}}\right)},
\end{aligned}
\tag{9}
$$

with $\mathcal{W} = \mathcal{W}^{\mathrm{R}} + i\mathcal{W}^{\mathrm{I}}$. Here we introduce the probability distribution $\tilde{P}(\mathbf{v}^z, \mathbf{h}; \mathcal{W}^{\mathrm{R}})$ yielding a Boltzmann distribution over visible and hidden neurons, and the phase $\tilde{\varphi}(\mathbf{v}^z, \mathbf{h}; \mathcal{W}^{\mathrm{I}})$.

The basis expansion coefficients can then be expressed as

$$
c_{\mathbf{v}^z}\left(\mathcal{W}\right) = \sum_{\{\mathbf{h}\}} \tilde{P}\left(\mathbf{v}^z, \mathbf{h}; \mathcal{W}^{\mathrm{R}}\right) e^{i\tilde{\varphi}\left(\mathbf{v}^z, \mathbf{h}; \mathcal{W}^{\mathrm{I}}\right)},
\tag{10}
$$

according to Eq. (3). Substituting Eq. (10) into Eq. (5), expectation values of diagonal operators in the z-basis, $\mathcal{O}^{\mathrm{diag}}$, can be expressed as

$$
\begin{aligned}
\left\langle \mathcal{O}^{\mathrm{diag}} \right\rangle = \frac{1}{Z\left(\mathcal{W}\right)} \sum_{\{\mathbf{v}^z\}} \sum_{\{\mathbf{h},\tilde{\mathbf{h}}\}} &\left[\mathcal{O}^{\mathrm{diag}}\left(\mathbf{v}^z\right) e^{i\varphi\left(\mathbf{v}^z, \mathbf{h}, \tilde{\mathbf{h}}; \mathcal{W}^{\mathrm{I}}\right)}\right] \\
&\times P\left(\mathbf{v}^z, \mathbf{h}, \tilde{\mathbf{h}}; \mathcal{W}^{\mathrm{R}}\right),
\end{aligned}
\tag{11}
$$

$$
Z\left(\mathcal{W}\right) = \sum_{\{\mathbf{v}^z\}} \sum_{\{\mathbf{h},\tilde{\mathbf{h}}\}} e^{i\varphi\left(\mathbf{v}^z, \mathbf{h}, \tilde{\mathbf{h}}; \mathcal{W}^{\mathrm{I}}\right)} P\left(\mathbf{v}^z, \mathbf{h}, \tilde{\mathbf{h}}; \mathcal{W}^{\mathrm{R}}\right),
\tag{12}
$$

where we introduce

$$\varphi\left(\mathbf{v}^z, \mathbf{h}, \tilde{\mathbf{h}}; \mathcal{W}^{\mathrm{I}}\right) =: \tilde{\varphi}\left(\mathbf{v}^z, \mathbf{h}; \mathcal{W}^{\mathrm{I}}\right) - \tilde{\varphi}\left(\mathbf{v}^z, \tilde{\mathbf{h}}; \mathcal{W}^{\mathrm{I}}\right),$$
(13)

$$P\left(\mathbf{v}^z, \mathbf{h}, \tilde{\mathbf{h}}; \mathcal{W}^{\mathrm{R}}\right) =: \tilde{P}\left(\mathbf{v}^z, \mathbf{h}; \mathcal{W}^{\mathrm{R}}\right) \tilde{P}\left(\mathbf{v}^z, \tilde{\mathbf{h}}; \mathcal{W}^{\mathrm{R}}\right).$$
(14)

It turns out that the hidden neuron configurations are summed over twice (sums over $\mathbf{h}$ and $\tilde{\mathbf{h}}$) when calculating expectation values. These sums originate from the coefficients of bra- and ket-states, respectively [60].

As $P(\mathbf{v}^z, \mathbf{h}, \tilde{\mathbf{h}}; \mathcal{W}^{\mathrm{R}})$ yields a Boltzmann distribution over the visible and hidden neurons, the sums in Eqs. (11)–(12) can be approximated by summing over samples drawn from these probability distributions. We use standard block Gibbs sampling [67]. In contrast to Eq. (5), not only the operator itself needs to be evaluated for each sample, but also the phase $\varphi(\mathbf{v}^z, \mathbf{h}, \tilde{\mathbf{h}}; \mathcal{W}^{\mathrm{I}})$. Furthermore, this phase appears in the normalization factor $Z(\mathcal{W})$ [cf. Eq. (12)].

The observable is thus reweighted with a complex phase in the evaluation of expectation values, which is why one refers to this ansatz as a phase reweighting scheme. This is a commonly used method in quantum Monte Carlo approaches [48–51, 53, 59]. Note that this method often suffers from a sign problem. If the phases fluctuate heavily they can cancel each other, resulting in an uncontrolled growth of the variance of the sampled quantity, which in turn requires an exponentially growing number of samples [48–51]. A more quantitative account of this will be given in Sect. VI B.

IV.   MEASURING IN DIFFERENT BASES

Having introduced the phase reweighting scheme to evaluate expectation values of diagonal operators by sampling from a Boltzmann distribution, we now derive a scheme for measuring non-diagonal operators that eliminates the use of local operators. This overcomes the problem that efficient evaluation of local operators requires them to be sparse and that their matrix elements need to be evaluated explicitly. More importantly, only the states of the visible spins and the phases associated with a given sample state need to be evaluated for the samples drawn from a Boltzmann distribution. This enables the use of neuromorphic architectures for performing the sampling.

Any hermitian operator can be decomposed into a Pauli string as

$$\mathcal{O} = \sum_{\{\boldsymbol{\alpha}\}} D_{\boldsymbol{\alpha}} \sigma_1^{\alpha_1} \otimes \ldots \otimes \sigma_N^{\alpha_N}$$
(15)

where $\alpha_i \in \{0, 1, 2, 3\}$ and $\sigma_i^{\alpha_i}$ are Pauli operators acting on spin i with $\sigma_i^0 = \mathbb{1}$ and $D_{\boldsymbol{\alpha}}$ are expansion coefficients. Here, we will only consider product operators or Pauli
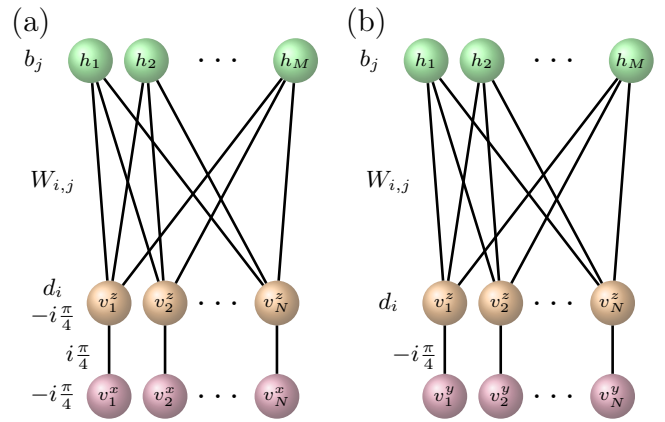

FIG. 2. Setup of the deep neural network parametrization to perform measurements in different spin bases. Panel (a) shows the extension of the RBM with visible neurons v_i^z, hidden neurons h_j, biases d_i, b_j and connecting weights $W_{i,j}$, cf. Fig. 1, to a deep neural network (dNN) to represent the state in the x-basis with visible neurons v_i^x, where the neurons v_i^z turn into hidden neurons which are summed over. v_i^x and v_i^z are connected by weights $i\pi/4$, and biases for v_i^x as well as imaginary parts of the biases for v_i^z appear. Panel (b) shows the corresponding dNN to perform measurements in the y-basis, where the visible neurons v_i^y are added instead of v_i^x in panel (a). Here only connecting weights between v_i^y and v_i^z are required, but no further biases appear.

strings, such as the magnetization of a single spin $\sigma_i^{\alpha_i}$ (not writing out identities) or correlations between two spins $\sigma_i^{\alpha_i} \sigma_j^{\alpha_j}$, but generalization to arbitrary operators is straight forward.

An operator is diagonal in the z-basis if it only involves $\sigma^3 = \sigma^z$-operators (and identities). A non-diagonal spin operator, which involves σ^x and/or σ^y, can be evaluated by rotating the spins locally into the basis in which the operators acting on the corresponding spins are diagonal. In the following we will show that this procedure can be encoded as an additional layer in the neural network representation of the state.

We first consider the Pauli-operator σ_i^x acting on a single spin i. The eigenstates of this operator can be obtained from the z-basis states by applying a rotation,

$$|v_i^x\rangle = \sum_{\{v_i^z\}} u_{z \to x}\left(v_i^x, v_i^z\right) |v_i^z\rangle,$$
(16)

with the entries $u_{z \to x}(v_i^x, v_i^z) = \langle v_i^z | v_i^x \rangle$ of the unnormalized rotation matrix

$$U_{z \to x} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$
(17)

These entries can be written as an exponential function,

$$u_{z \to x}\left(v_i^x, v_i^z\right) = \exp\left[i\frac{\pi}{4}\left(v_i^x v_i^z - v_i^x - v_i^z + 1\right)\right],$$
(18)

taking the form of the exponential of an RBM network energy, similar to Eq. (2). In order to evaluate the ex-

pectation value of σ_i^x we can equivalently apply a rotation to spin i and then evaluate the diagonal observable σ_i^z. This means that we replace the single-particle basis states $|v_i^z\rangle$ appearing in Eq. (1) by Eq. (16), resulting in an additional summation over v_i^z. This additional summation means in our network representation that v_i^z has become a hidden neuron. Since the involved matrix elements can be written as exponential factors of the required form Eq. (18), they just contribute to the overall network energy as any other weights in the network. We hence end up with a deep neural network (dNN) with two hidden layers. If we are interested in measuring the x-magnetization of all spins or correlations $\sigma_i^x \sigma_j^x$, we would add an additional connection to every visible spin as illustrated in Fig. 2(a).

To summarize, we can parametrize the state of a spin in the x-basis by introducing a neuron v_i^x in the RBM parametrization, which is connected to v_i^z via a weight $i\pi/4$ and has a bias of $-i\pi/4$, while v_i^z gets an additional bias of $-i\pi/4$. The overall bias of $i\pi/4$ appearing in Eq. (18) can be neglected as it is an irrelevant global phase factor.

An analogous expression can be derived for a local rotation of the spin state into the y-basis, so that also a neuron v_i^y can be connected to v_i^z, enabling a measurement of the Pauli-operator σ_i^y, as illustrated in Fig. 2(b). The v_i^z neuron again turns into a hidden neuron, yielding the transformation

$$|v_i^y\rangle = \sum_{\{v_i^z\}} u_{z\to y}\left(v_i^y, v_i^z\right)|v_i^z\rangle, \qquad (19)$$

with elements

$$u_{z\to y}\left(v_i^y, v_i^z\right) = \exp\left[i\frac{\pi}{4}\left(1 - v_i^y v_i^z\right)\right], \qquad (20)$$

of the rotation matrix

$$U_{z\to y} = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}. \qquad (21)$$

With this ansatz, any desired Pauli string operator can be measured using the phase reweighting scheme on the corresponding dNN. The network representation can always be set up to represent the local spins in the basis where the applied operator becomes diagonal. Notice that the added network parameters are purely imaginary such that the resulting dNN will always contain complex parameters. We saw in Eqs. (13)–(14) that the evaluation of observables requires to sum twice over all hidden neurons. This will now be the case for all spins v_i^z that get connected to an additional spin v_i^x or v_i^y.

V. GROUND STATES OF THE TRANSVERSE-FIELD ISING MODEL

Having introduced the dNN setup with the phase reweighting scheme, we benchmark the approach on the ground state of the transverse-field Ising model (TFIM) at the quantum critical point.

A. Model and dNN Representation

The one-dimensional TFIM is an integrable model defined on a spin-1/2 chain with N sites via the Hamiltonian

$$H_{\mathrm{TFIM}} = -J\sum_{i=1}^{N} \sigma_i^z \sigma_{(i+1)\%N}^z - h\sum_{i=1}^{N} \sigma_i^x, \qquad (22)$$

with $(i+1)\%N = (i+1)\mathrm{mod}N$ denoting a modulo-N calculation, i.e., we choose periodic boundary conditions. The system undergoes a quantum phase transition at the quantum critical point reached for the transverse magnetic field strength $h = h_{\mathrm{c}} = J$. In the following, we fix the energy scale by setting $J = 1$. The model is integrable and can be solved in terms of a Jordan-Wigner-fermionization [1–5, 68]. It has been studied in great detail and is a common choice to benchmark approximative analytical and numerical methods.

To represent the ground state in the RBM parametrization, the corresponding weights can be found variationally by using stochastic reconfiguration minimizing the system energy [30]. In the following we apply this variational ansatz to find weights and biases representing the ground state at the quantum critical point. Subsequently, given the state representation, we fix the weights and biases and apply the phase reweighting sampling scheme to benchmark its performance when applied in the dNN approach.

We only consider moderate system sizes, $N \leq 10$, so that we can apply exact diagonalization, enabling an exact evaluation of any desired operator, and thus a benchmark of the dNN ansatz in arbitrary bases. We remark that, for these system sizes, sampling of configurations, usually used for evaluating gradients during training, can be omitted in favor of calculating the gradients exactly by summing over all states. This makes the obtained representation more accurate.

B. Results

When training the RBM to represent a ground state of the TFIM, it turns out that it is sufficient to choose the weights purely real since the Hamiltonian is stoquastic, i.e., all its off-diagonal elements in the z-basis are real and non-positive [69, 70]. We train a real RBM with as many hidden as visible neurons, $M = N$, to represent the ground state of the TFIM at the critical point, $h = 1$, and fix the weights. We use the dNN setup according to Fig. 2, together with the phase reweighting scheme with block Gibbs sampling to perform measurements of operators in the z- and x-directions. Here we choose the ground state at the quantum critical point where the entanglement entropy is maximal and grows logarithmically with system size [71]. This demonstrates that quantum effects can be represented in the classical network ansatz. It should be noted that two kinds of
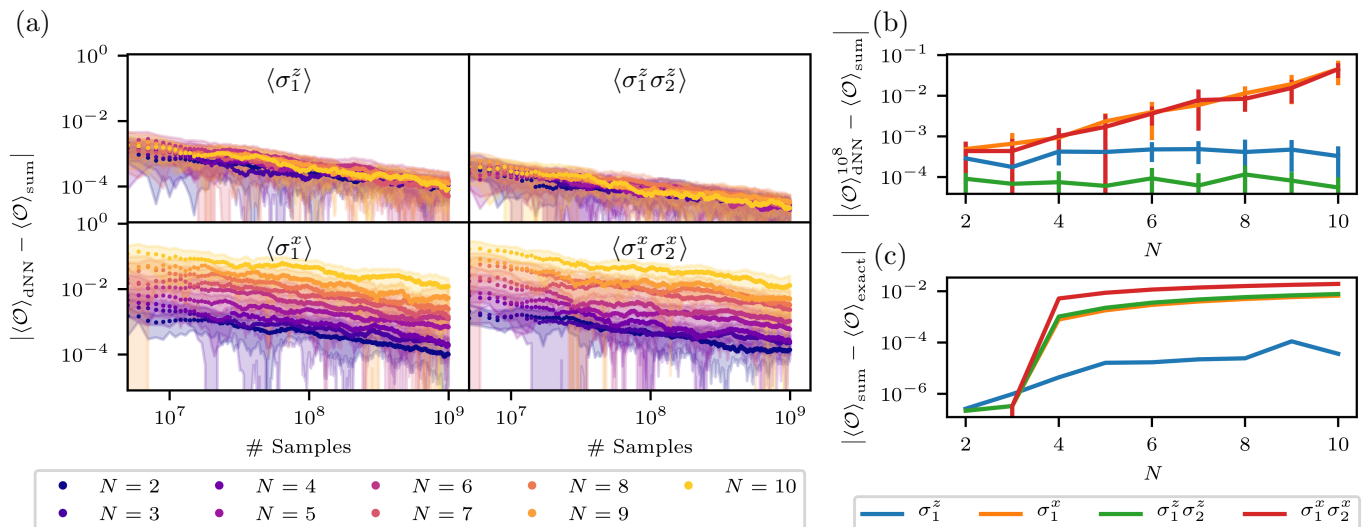
FIG. 3. RBM representation of the TFIM ground state at the quantum critical point: Sampling and representation errors for the phase reweighting scheme. Magnetizations $\langle\sigma_1^k\rangle$ of the first spin and correlations $\langle\sigma_1^k\sigma_2^k\rangle$ between the first two spins in a TFIM chain of N spins, in k-direction, $k = z, x$, have been evaluated. Panel (a) shows the absolute deviation of the observables evaluated with the phase reweighting method, $\langle\mathcal{O}\rangle_{\mathrm{dNN}}$, from the respective value obtained by explicitly summing over all states, $\langle\mathcal{O}\rangle_{\mathrm{sum}}$, as a function of the sample size. The number of spins in the chain is varied from $N = 2$ to $N = 10$, indicated by the colors. The sampling is run ten times for each system size and the results are averaged, where shaded regions denote statistical fluctuations. While we do not find a dependence on the system size for measurements in the z-basis (upper row), we find increasing deviations with larger N for measurements in the x-basis (lower row). In both cases the error decays only slowly and a huge amount of samples is considered. Panel (b) shows the same data evaluated at a fixed sample size of 10^8, as a function of system size N, as indicated in the legend at the bottom. Error bars denote statistical fluctuations within averaging over ten runs. We find the accuracy of the measurements in the z-basis to be approximately constant, while the deviations grow exponentially for measurements in the x-basis. The latter indicate the existence of a sign problem which limits the method to small system sizes. Panel (c) shows the absolute deviation between the observables as calculated in the RBM parametrization but by summing over all configurations, $\langle\mathcal{O}\rangle_{\mathrm{sum}}$, from their values obtained by exact diagonalization, $\langle\mathcal{O}\rangle_{\mathrm{exact}}$. This comparison shows the intrinsic error of the RBM parametrization and illustrates that it dominates the overall error for $N \geq 4$.

imperfections are involved now. First, the representation of the ground state as an RBM is not exact (representation error) and, second, there will be statistical errors due to finite sample sizes (sampling error). We analyze both aspects with the main focus being on the statistical sampling errors.

Figure 3 shows the results for the sampling error, where we vary the system size from $N = 2$ to $N = 10$. Panel (a) shows the absolute deviations between performing the phase reweighting sampling in the dNN representation and summing over all configurations explicitly, evaluating operators in the x-basis using local operators. We study magnetizations (σ_1^z and σ_1^x) and nearest-neighbor correlations in the z- and x-directions ($\sigma_1^z\sigma_2^z$ and $\sigma_1^x\sigma_2^x$). Restricting to the magnetizations of the first spin and the correlations between the first two spins suffices due to translation invariance, which is explicitly implemented in the structure of the dNN weights [30, 72]. We run the sampling ten times for each system size and average the outcomes, with shaded regions denoting the statistical fluctuations.

We find good convergence, especially in the z-basis, where the weights are purely real and the complex phases

vanish. The absolute deviations go down proportional to $1/\sqrt{\# \text{ Samples}}$, as it is expected due to statistical arguments [73]. The error is approximately independent of system size. Considering measurements in the x-basis, we find larger deviations, which is reasonable as now the weights in the network also take imaginary values and phases need to be considered, which can cancel each other in the normalization factor $Z(\mathcal{W})$. If these phases fluctuate heavily, a sign problem can appear. This leads to divergences for too small sample sizes and requires an exponentially growing amount of samples to find convergence for increasing system sizes. An increase in the absolute deviation with growing system size is indeed observed [see lower panels of Fig. 3(a)]. However, we still find stable convergence to the exact solution as expected from statistical reasons up to $N = 10$.

Figure 3(b) shows the absolute deviations of the expectation values of magnetizations and correlations in the x- and z-directions using 10^8 samples, from results when summing over all states explicitly as a function of system size. The underlying data is the same as in panel (a). Here we see more clearly that the absolute deviations do not depend on the system size for measurements in the

z-basis, but they scale exponentially with the system size for measurements in the x-basis (mind the log-scale). In this case the ansatz performs inefficiently, as exponentially many samples are necessary when going to larger system sizes. However, in the present example the sample size is still much smaller than the number of possible states in the network, which here is 2^{5N} as we choose $M = N$. This enables simulations of slightly larger systems than with exact diagonalization using comparable resources.

Figure 3(c) shows the representation error, i. e. the absolute deviations of the observables obtained using the RBM parametrization and summing over all states explicitly from the expected value calculated via exact diagonalization. We find that the deviations grow abruptly larger with increasing system size for $N \geq 4$, which is probably due to the limited representational power of the network. While the state can be parametrized with good accuracy for $N < 4$, it takes a form for $N \geq 4$ which cannot be represented that accurately with the RBM ansatz. The deviations saturate around 10^{-2} for large system sizes, which is still small. This shows that the weights trained in the RBM parametrization represent the exact ground state with good accuracy for the cases considered here. However, the deviations are mostly larger than the ones in panel (a), showing that the overall error is dominated by the representation error and larger sample sizes cannot improve the accuracy any further.

In summary, we find that the ground state of the TFIM can be represented well with the RBM parametrization. It can be sampled using the phase reweighting scheme in the dNN ansatz to perform measurements in the x-basis, but due to exponentially scaling sample sizes it is limited to small system sizes.

VI. BELL AND GHZ STATES

We now apply our dNN approach with phase reweighting to a paradigmatic example of an entangled state, the Bell state [54–56], and its generalization to larger spin systems, the Greenberger-Horne-Zeilinger (GHZ) states [57]. With these examples we intend to assess whether the network can capture genuine quantum features such as entanglement and non-locality manifest in the violation of Bell's inequalities.

A. Model and dNN Representation

We consider a Bell state of two spin-1/2 particles, also called a Bell pair (BP),

$$|\Psi_{\text{BP}}\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle \right). \tag{23}$$

This state has non-classical correlations in the sense that it violates Bell's inequality. For classical systems Bell's inequality cannot be violated under the assumptions of

local realism [54–56]. Specifically, we consider the well-known CHSH-inequality (named after Clauser, Horne, Shimony and Hold) [74, 75]. A CHSH-inequality which is maximally violated by the considered state is given by

$$|\mathcal{B}| = \sqrt{2} \left| \langle \sigma_1^x \otimes \sigma_2^x \rangle - \langle \sigma_1^z \otimes \sigma_2^z \rangle \right| \\ \leq 2. \tag{24}$$

If the measured correlations exceed this bound, they cannot arise in a classical way. However, with the wave function of the Bell state we find

$$\mathcal{B} = 2\sqrt{2} > 2, \tag{25}$$

showing that the CHSH-inequality is violated. It has been shown that this is the maximum reachable value for a quantum state, so that the inequality is maximally violated [76].

To represent the Bell state in the dNN, we need $N = 2$ visible neurons and we show in Appendix A that it is sufficient to choose $M = 1$ hidden neuron. Expressions for the weights in the dNN can be derived analytically by demanding

$$c_{\mathbf{v}^z=[\pm 1, \pm 1]} (\mathcal{W}) \overset{!}{=} 0, \\ c_{\mathbf{v}^z=[\pm 1, \mp 1]} (\mathcal{W}) \overset{!}{=} \frac{1}{\sqrt{2}}, \tag{26}$$

see Appendix A for a full derivation and Table I for possible analytical expressions of the weights. Solving Eq. (26) yields infinitely many possible choices to represent a Bell state. Thus we can directly determine the weights in the dNN and do not need to train them. We can rather perform sampling via the phase reweighting scheme to measure the correlations in the x- and z-directions and see if Bell correlations can be captured with this ansatz.

The generalization of the Bell state to larger system sizes yields the GHZ state with state vector [57]

$$|\Psi_{\text{GHZ}}\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\uparrow \dots \uparrow\rangle + |\downarrow\downarrow \dots \downarrow\rangle \right). \tag{27}$$

This is a genuinely N-partite entangled state which we can represent in the dNN approach. As the GHZ state is the generalization of the Bell state to larger system sizes in the sense of being a superposition between two macroscopically different states, we use it to check the scalability of the dNN ansatz [57, 77, 78]. The weights to represent this GHZ state with a dNN can be calculated analytically by solving

$$c_{\mathbf{v}^z=[\pm 1, \dots, \pm 1]} (\mathcal{W}) \overset{!}{=} \frac{1}{\sqrt{2}}, \\ c_{\mathbf{v}^z \neq [\pm 1, \dots, \pm 1]} (\mathcal{W}) \overset{!}{=} 0, \tag{28}$$

which is solvable for $M = N - 1$ hidden neurons, see Appendix B for a full derivation. We derive a formula for the weights and biases in a general way as a function of the system size N. The final expressions for general system sizes are stated in Table I.

	Bell State (complex weights)	Bell State (imaginary weights)	GHZ State
d_j	$i\frac{\pi}{2}$ for $j = 1$, 0 otherwise	0	0
b_k	$i\frac{\pi}{2}$	0	$i\frac{\pi}{2}$
$W_{j,k}$	$\frac{(-1)^j}{2}\mathrm{arsinh}\left(\frac{1}{\sqrt{8}}\right) + i\frac{\pi}{2}$	$i\left[\frac{(-1)^j}{2}\arccos\left(\frac{1}{\sqrt{8}}\right) - \frac{\pi}{4}\right]$	$\frac{i}{2(N-1)}\arcsin\left[\frac{1}{2^{N-1/2}}\right]$ if $j \neq N$, $k = 1$; $\frac{i}{2}\arcsin\left[\frac{1}{2^{N-1/2}}\right]$ if $j = N$, $k = 1$; $i\frac{\pi}{4}\left(\delta_{j,k-1} + \delta_{j,k}\right)$ if $k \neq 1$.

TABLE I. Analytically calculated weights and biases entering the RBM representation of the Bell and GHZ states, where two possible solutions are given for the Bell state, one with complex values and one with purely imaginary entries. To avoid confusion with the imaginary unit we have used the indices j, k instead of the usual convention i, j to label the visible and hidden units, respectively.

B. Results

We implement a dNN representing a Bell state, consisting of one hidden and two visible neurons and perform the phase reweighting scheme to sample from the underlying Boltzmann distribution using block Gibbs sampling [67]. We consider two possible choices for the network parameters, one with purely imaginary and one with complex weights. The expressions for the weights are given explicitly in the left two columns of Table I.

As the system consists only of two sites, we can solve it exactly and compare the simulation outcome with the exact solution. The expected magnetizations and correlations in the x- and z-directions are given by

$$\langle\sigma_i^x\rangle = \langle\sigma_i^z\rangle = 0, \ \forall \ i \in \{1, 2\},$$
$$\langle\sigma_1^x\sigma_2^x\rangle = -\langle\sigma_1^z\sigma_2^z\rangle = 1. \tag{29}$$

The simulation results of the magnetization in the x- and z-basis are shown in Fig. 4, where for symmetry reasons we only plot the result for the first spin. Shown are results for both choices of the network parameters. The insets depict the corresponding observables together with the exact solutions as functions of sample size, and the main plots show the absolute deviations from the exact result. To compare the convergence to the expectation on grounds of statistical arguments [73] we also display the expected sampling error $\sigma[\langle\mathcal{O}\rangle]/\sqrt{\# \text{ Samples}}$ where $\sigma^2[\langle\mathcal{O}\rangle]$ is the variance. The sampling error can be obtained via error propagation [73],

$$|\langle\mathcal{O}\rangle_{\text{dNN}} - \langle\mathcal{O}\rangle_{\text{exact}}| = \frac{\sigma[\langle\mathcal{O}\rangle_{\text{dNN}}]}{\sqrt{\# \text{ Samples}}}, \tag{30}$$

$$\sigma[\langle\mathcal{O}\rangle_{\text{dNN}}] = \left|\frac{\sigma\left[\text{Re}\langle\Psi|\mathcal{O}|\Psi\rangle_{\text{dNN}}\right]}{\text{Re}\langle\Psi|\Psi\rangle_{\text{dNN}}}\right| + \left|\frac{\text{Re}\langle\Psi|\mathcal{O}|\Psi\rangle_{\text{dNN}}}{\text{Re}\langle\Psi|\Psi\rangle_{\text{dNN}}^2}\sigma\left[\text{Re}\langle\Psi|\Psi\rangle_{\text{dNN}}\right]\right|, \tag{31}$$

where the index "exact" refers to the exact quantum mechanical expectation value. The index "dNN" denotes the average over samples drawn from the dNN.

For such small system sizes as we consider them here, these expressions can be evaluated explicitly by summing over all possible network states. Due to the exact sum over all states the imaginary parts of the expectation values vanish with good accuracy and we hence neglect them. We expect the simulations to follow this decay for a sufficiently large number of samples, when the effects due to fluctuating phases are suppressed. In Fig. 4, the blue line in the main plot denotes the expected convergence according to the explicitly evaluated variance for the case of purely imaginary weights. The expected convergence behavior for the case of complex weights shows a similar decay. We find that the absolute deviation follows the expected sampling error accurately and hence converges to the exact solution, where no clear difference can be observed between the two cases of complex and purely imaginary weights.

To benchmark the phase reweighting scheme on the evaluation of correlations, we directly consider the CHSH-observable, Eq. (24). The result is shown in Fig. 5, which has the same structure as the plots in Fig. 4.

In Fig. 5 we add a dashed line at the classical limit of $\mathcal{B} = 2$ in the inset and at the deviation of the limit from the exact solution, $2\sqrt{2}-2$, in the main plot. The CHSH-inequality is hence violated if the curve in the inset is above and the curve in the main plot is below the dashed line. We find clear convergence in agreement with the explicitly calculated behavior expected from statistical arguments. We also find a violation of the CHSH-inequality after rather short sampling times. However, we already consider 10^6 samples here, which is large compared to the number of network configurations of the dNN which is $2^{N+2M} = 2^4$ or $2^{N+2N+2M} = 2^8$ for measurements in the z- or x-basis, respectively.

To analyze how the sample size necessary to find convergence scales with the system size we now consider GHZ states with $N > 2$. We first consider the GHZ state for $N = 3$ sites. The expected magnetizations and
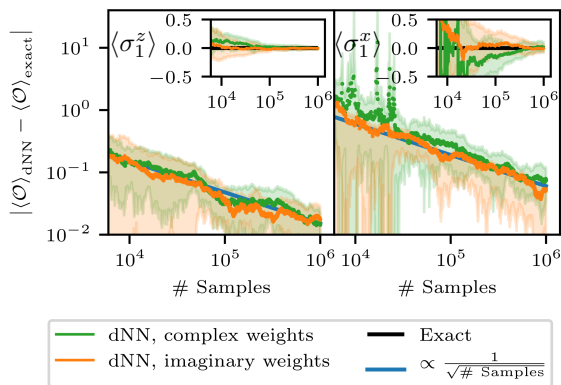
FIG. 4. RBM representation of the Bell state (23): Sampling errors for the phase reweighting scheme. The magnetization of one spin of a Bell pair in the x- and z-directions is evaluated. The graphs show the absolute difference between the exact expectation values and those sampled from the dNN by means of the phase reweighting scheme, as a function of sample size. Two choices of weights are compared, taking either purely imaginary (orange data) or complex values (green data). The blue line indicates the deviation expected on statistical grounds. Insets show the sampled values together with the respective exact solutions. Ten simulation runs are averaged. The shaded regions indicate the statistical fluctuations. For the magnetization in the x-basis, the deviations decay only slowly as function of the sample size while fluctuations resulting from the sum over the complex phases are found for small samples.

correlations are

$$\langle \sigma_i^x \rangle = \langle \sigma_i^z \rangle = 0,$$
$$\langle \sigma_1^x \sigma_2^x \sigma_3^x \rangle = \langle \sigma_i^z \sigma_j^z \rangle = 1, \qquad (32)$$
$$\forall\, i,j \in \{1,2,3\}.$$

We set up a dNN with $M = 2$ hidden neurons and the analytically derived, purely imaginary weights, see the right column of Table I for explicit expressions, and perform the phase reweighting scheme in combination with block Gibbs sampling to measure magnetizations and correlations in the x- and z-directions. The results are shown in Fig. 6, where each panel has the same structure as the plots for the Bell state. We find convergence to the exact solution with the expected dependence on the sample size. Observables in the z-basis show faster convergence due to the smaller network size compared to off-diagonal observables which require an additional hidden layer. The expected convergence behavior can again be evaluated explicitly, as for the Bell state, since the network size is still small. For measurements in the x-basis we find huge fluctuations for small sample sizes, which basically result from the division by the sum over the phases, see Eqs. (11)–(12). This can lead to divergences due to cancellations of the phases. We would expect these fluctuations to vanish for larger sample sizes, however we already consider 10^7 samples here, exceeding the number of network configurations of 2^7 and 2^{13} for representing
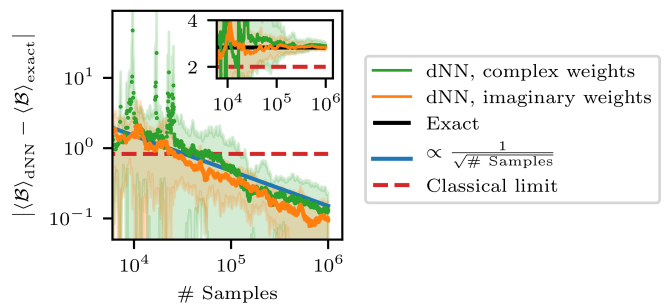

FIG. 5. RBM representation of the Bell state: observable $\mathcal{B}$ entering the CHSH Bell inequality (24). Shown is the same deviation as in Fig. 4, as a function of samples drawn from the dNN representation of a Bell state with phase reweighting ansatz. The weights can be chosen either complex or purely imaginary, indicated as green and orange data, respectively. The main plot shows the absolute deviation of the simulation from the exact solution together with the expected convergence behavior, while the inset shows the direct evaluation of the observables in comparison to the exact solution. Shaded regions indicate statistical fluctuations resulting from averaging over ten runs.

the state in the z- and x-basis, respectively. Besides the fluctuations, the functions decay as expected from statistical arguments. We even find the decay in the simulations slightly below the expected behavior, but a convergence to the blue curve is visible for large sample sizes.

We now increase the system size further and consider a GHZ state with $N = 5$ spins. We can represent it using a dNN with $M = 4$ hidden neurons, see Appendix B. The exact magnetizations and correlations are

$$\langle \sigma_i^x \rangle = \langle \sigma_i^z \rangle = 0,$$
$$\langle \sigma_1^x \sigma_2^x \sigma_3^x \sigma_4^x \sigma_5^x \rangle = \langle \sigma_i^z \sigma_j^z \rangle = 1, \qquad (33)$$
$$\forall\, i,j \in \{1,2,3,4,5\}.$$

The simulation results for these observables are shown in Fig. 7, where we find convergence for measurements in the z-basis according to the expected statistical behavior. This can still be evaluated explicitly for the small network size. However, considering the outcome of measurements in the x-basis we find large fluctuations which do not decrease for the sample sizes considered. This indicates that we are still undersampling and convergence to the exact solution is not yet visible. According to the explicitly calculated error decay, we would expect convergence appearing for larger sample sizes. As we already consider 10^{10} samples, we could not increase the sample size further with the given computational setup. The networks here have 2^{13} and 2^{23} possible configurations to represent the system in the z- and x-basis, respectively.

In summary we find that the necessary sample size required to find convergence scales exponentially with the system or network size. This is due to the appearance of a sign problem, as it is known to be present for phase reweighting sampling schemes in quantum Monte Carlo
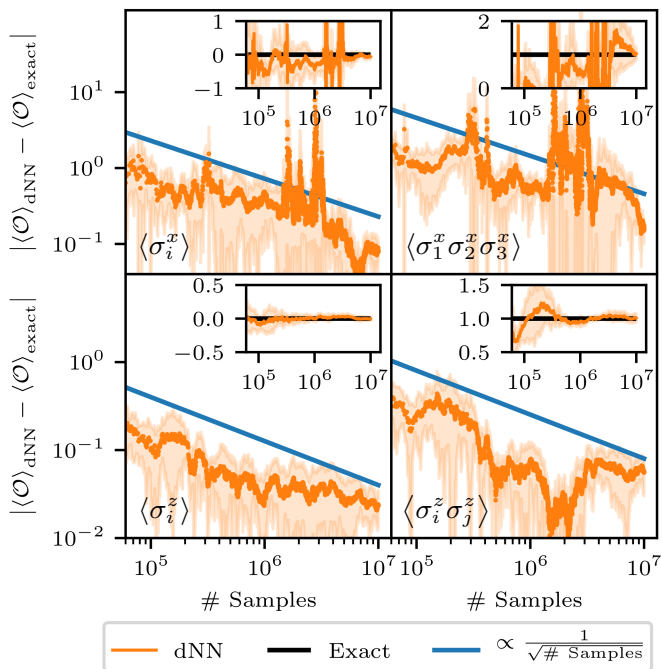
FIG. 6. RBM representation of the GHZ state with $N = 3$ spins: Magnetizations and correlations in the x- and z-bases resulting from sampling by means of the phase reweighting scheme as functions of sample size. Insets show direct evaluations of the operators together with the exact outcome, while main plots show the absolute deviations from the exact solution compared to the explicitly evaluated expected convergence behavior. Five simulation runs are averaged over, with statistical fluctuations indicated by the shaded regions. Additionally, we average over the spin sites i and j due to translation invariance.
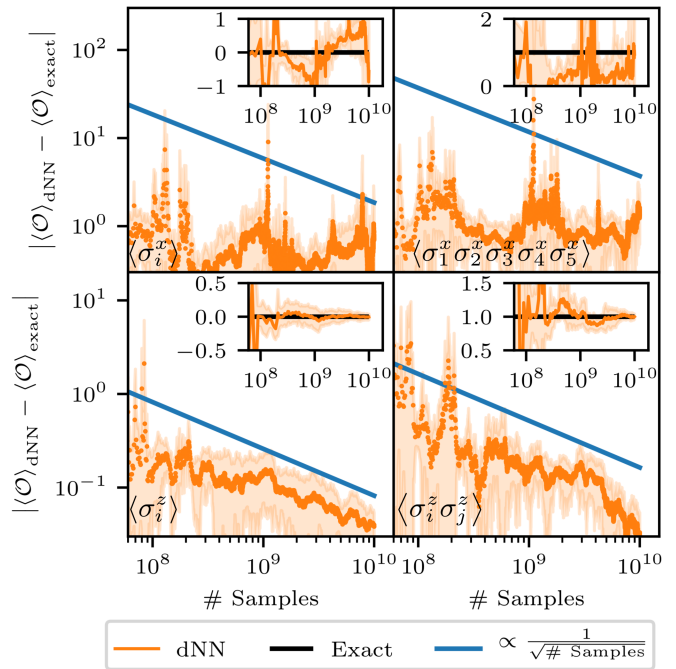
FIG. 7. RBM representation of the GHZ state with $N = 5$ spins: Magnetizations and correlations resulting in the phase reweighting scheme as functions of sample size. Main plots show absolute deviations of the simulations from the exact solution together with the explicitly evaluated expected convergence behavior, while insets show the direct evaluations of the operators together with the exact solution. Shaded regions denote statistical fluctuations resulting from averaging over five simulation runs as well as over the spin sites i and j due to translation invariance.

approaches [48–51]. Intuitively the sign problem results from the sum over the phase factors, which can be distributed broadly on the unit circle in the complex plane. Thus, exponentially many samples are necessary to sum those phases up in the right way and get a stable result.

VII. CONCLUSION

When parametrizing wave functions of quantum spin systems with restricted Boltzmann machines, the weights and biases need to be chosen complex. Thus, the wave function can no longer be viewed as the marginal of a Boltzmann distribution but is rather a sum over complex terms [30, 31, 66, 79]. This prohibits the straight forward implementation of the sampling from quantum states using classical neuromorphic structures. We recover a way to sample the visible and hidden neurons from Boltzmann distributions defined by the real parts of the weights and biases via a phase reweighting scheme, where standard Gibbs sampling can be applied [67]. This enables an extension to deep neural networks with multiple layers, which we use to derive a representation of

spin states in arbitrary bases.

When benchmarking this ansatz on highly entangled spin systems, we find an exponential scaling of the sample size necessary for convergence to the exact solution with increasing system or network size. This can be understood as a sign problem, meaning that the variances of the sampled quantities increase exponentially with system size. The method is hence rendered inefficient for cases where the wave-function coefficients cannot be chosen real and positive [48–51]. However, when representing the ground state of the TFIM at the quantum critical point, we find that the phase reweighting scheme in the dNN ansatz yields accurate results for operators in the z-basis without dependence on the system size. For measurements in the x-basis we find an inefficient exponential scaling of the sample size with the system size. In summary, our ansatz yields a generalization of the RBM parametrization of wave functions to deep networks and enables an implementation of the sampling on neuromorphic hardware, which can efficiently sample from Boltzmann distributions and could provide a speedup and thus shift the limitations to larger system sizes, while it is not expected to in general overcome the curse of dimensionality of the quantum many-body problem [25–27].

Appendix A: Representation of the Bell State

To parametrize the Bell state with an RBM, we need $N = 2$ visible neurons for the two spins and it is sufficient to choose $M = 1$ hidden neuron. We can analytically derive expressions for the weights to represent the corresponding basis expansion coefficients. The state vector reads [54–56]

$$
\begin{aligned}
|\Psi\rangle &= \frac{1}{\sqrt{2}} \left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle \right) \\
&= \sum_{\{\mathbf{v}^z\}} c_{\mathbf{v}^z}(\mathcal{W}) |\mathbf{v}^z\rangle.
\end{aligned} \tag{A1}
$$

Enumerating the two states as $|\downarrow\rangle = |-1\rangle$ and $|\uparrow\rangle = |1\rangle$, we get

$$
\begin{aligned}
c_{\mathbf{v}^z=(\pm1,\pm1)}(\mathcal{W}) &= \exp\left[\pm(d_1+d_2)\right] \\
&\quad \times 2\cosh\left[\pm(W_{1,1}+W_{2,1})+b_1\right] \\
&\overset{!}{=} 0,
\end{aligned} \tag{A2}
$$

$$
\begin{aligned}
c_{\mathbf{v}^z=(\pm1,\mp1)}(\mathcal{W}) &= \exp\left[\pm(d_1-d_2)\right] \\
&\quad \times 2\cosh\left[\pm(W_{1,1}-W_{2,1})+b_1\right] \\
&\overset{!}{=} \frac{1}{\sqrt{2}}.
\end{aligned} \tag{A3}
$$

Here we use the RBM parametrization of the coefficients, see Eqs. (2)–(3) in the main text,

$$
\begin{aligned}
c_{\mathbf{v}^z}(\mathcal{W}) &= \exp\left[\sum_{i=1}^{N} d_i v_i^z\right] \\
&\quad \times \prod_{j=1}^{M} 2\cosh\left[\sum_{i=1}^{N} v_i^z W_{i,j} + b_j\right].
\end{aligned} \tag{A4}
$$

From Eq. (A2) we choose the ansatz

$$
\cosh\left[\pm(W_{1,1}+W_{2,1})+b_1\right] \overset{!}{=} 0 \tag{A5}
$$

$$
\Rightarrow b_1 \pm (W_{1,1}+W_{2,1}) \overset{!}{=} i\pi\left(n_\pm + \frac{1}{2}\right), \quad n_\pm \in \mathbb{Z}, \tag{A6}
$$

$$
\begin{aligned}
\Rightarrow b_1 &= i\pi\left(\frac{n_+ + n_-}{2} + \frac{1}{2}\right), \\
W_{1,1} + W_{2,1} &= i\pi\frac{n_+ - n_-}{2}.
\end{aligned} \tag{A7}
$$

Inserting this into Eq. (A3), we get two expressions depending on whether $n^{(\pm)} = n_+ \pm n_-$ is even or odd,

$$
\begin{aligned}
&\cosh\left[b_1 \pm (W_{1,1}-W_{2,1})\right] \\
&= \begin{cases} \pm i(-1)^{n^{(+)}/2}\sinh\left[W_{1,1}-W_{2,1}\right], & \text{if } n^{(\pm)} \text{ even}, \\ (-1)^{(n^{(+)}+1)/2}\cosh\left[W_{1,1}-W_{2,1}\right], & \text{if } n^{(\pm)} \text{ odd}. \end{cases}
\end{aligned} \tag{A8}
$$

Considering the even case, Eqs. (A2)–(A3) turn into

$$
\begin{aligned}
i(-1)^{n^{(+)}/2}\sinh\left[d_1-d_2\right]\sinh\left[W_{1,1}-W_{2,1}\right] &\overset{!}{=} 0, \\
i(-1)^{n^{(+)}/2}\cosh\left[d_1-d_2\right]\sinh\left[W_{1,1}-W_{2,1}\right] &\overset{!}{=} \frac{1}{\sqrt{8}}.
\end{aligned} \tag{A9}
$$

From this, it follows that

$$
\begin{aligned}
d_1 - d_2 &= i\pi\left(m_\mathrm{e} + \frac{1}{2}\right), \quad m_\mathrm{e} \in \mathbb{Z}, \\
W_{1,1} - W_{2,1} &= (-1)^{n^{(+)}/2+m_\mathrm{e}+1}\operatorname{arsinh}\left[\frac{1}{\sqrt{8}}\right].
\end{aligned} \tag{A10}
$$

Together with Eq. (A7) we get a solutions for even $n^{(\pm)}$,

$$
\begin{aligned}
W_{1,1} &= (-1)^{n^{(+)}/2+m_\mathrm{e}+1}\frac{1}{2}\operatorname{arsinh}\left[\frac{1}{\sqrt{8}}\right] + i\frac{\pi}{4}n^{(-)}, \\
W_{2,1} &= (-1)^{n^{(+)}/2+m_\mathrm{e}}\frac{1}{2}\operatorname{arsinh}\left[\frac{1}{\sqrt{8}}\right] + i\frac{\pi}{4}n^{(-)}, \\
d_1 - d_2 &= i\pi\left(m_\mathrm{e} + \frac{1}{2}\right), \\
b_1 &= i\pi\left(\frac{n^{(+)}}{2} + \frac{1}{2}\right).
\end{aligned} \tag{A11}
$$

For the case of odd $n^{(\pm)}$, Eqs. (A2)–(A3) yield

$$
\begin{aligned}
(-1)^{(n^{(+)}+1)/2}\cosh\left[d_1-d_2\right]\cosh\left[W_{1,1}-W_{2,1}\right] &\overset{!}{=} \frac{1}{\sqrt{8}}, \\
(-1)^{(n^{(+)}+1)/2}\sinh\left[d_1-d_2\right]\cosh\left[W_{1,1}-W_{2,1}\right] &\overset{!}{=} 0.
\end{aligned} \tag{A12}
$$

This leads to

$$
\begin{aligned}
d_1 - d_2 &= i\pi m_\mathrm{o}, \quad m_\mathrm{o} \in \mathbb{Z}, \\
W_{1,1} - W_{2,1} &= i\left[\pi\left(\frac{n^{(+)}+1}{2} + m_\mathrm{o}\right) \right. \\
&\quad \left. + (-1)^{(n^{(+)}+1)/2+m_\mathrm{o}}\arccos\left(\frac{1}{\sqrt{8}}\right)\right].
\end{aligned} \tag{A13}
$$

So we get for odd $n^{(\pm)}$ the solutions

$$W_{1,1} = i \left[(-1)^{(n^{(+)}+1)/2+m_o} \frac{1}{2}\arccos\left(\frac{1}{\sqrt{8}}\right) \right.$$
$$\left. + \frac{\pi}{4}\left(n^{(+)} + n^{(-)} + 1 + 2m_o\right) \right],$$
$$W_{2,1} = i \left[(-1)^{(n^{(+)}+1)/2+m_o+1} \frac{1}{2}\arccos\left(\frac{1}{\sqrt{8}}\right) \right.$$
$$\left. + \frac{\pi}{4}\left(n^{(-)} - n^{(+)} - 1 - 2m_o\right) \right],$$
$$d_1 - d_2 = i\pi m_o,$$
$$b_1 = i\frac{\pi}{2}\left(n^{(+)} + 1\right). \tag{A14}$$

The solutions are highly degenerate, as three integers can be chosen arbitrarily, so we find infinitely many possibilities to choose the weights. In the main text we focus on two specific choices, one with complex and one with purely imaginary weights. The expressions for complex weights are given by choosing $n_+ = 1$, $n_- = -1$, so that we consider the case of even $n^{(\pm)}$. Additionally we choose $m_e = 0$, yielding

$$d_1 - d_2 = i\frac{\pi}{2} \Rightarrow d_1 = i\frac{\pi}{2}, \ d_2 = 0,$$
$$b_1 = i\frac{\pi}{2},$$
$$W_{1,1} = -\frac{1}{2}\text{arsinh}\left(\frac{1}{\sqrt{8}}\right) + i\frac{\pi}{2}, \tag{A15}$$
$$W_{2,1} = \frac{1}{2}\text{arsinh}\left(\frac{1}{\sqrt{8}}\right) + i\frac{\pi}{2}.$$

The second case we consider in the main text is the choice of purely imaginary weights, which we reach by setting

$n_+ = -1$, $n_- = 0$, so that $n^{(\pm)}$ is odd. By also setting $m_o = 0$, we get

$$d_1 - d_2 = 0 \Rightarrow d_1 = d_2 = 0,$$
$$b_1 = 0,$$
$$W_{1,1} = i\left(-\frac{1}{2}\arccos\left[\frac{1}{\sqrt{8}}\right] - \frac{\pi}{4}\right), \tag{A16}$$
$$W_{2,1} = i\left(\frac{1}{2}\arccos\left[\frac{1}{\sqrt{8}}\right] - \frac{\pi}{4}\right).$$

These are the explicit values for the weights and biases as stated in Table I in the main text.

Appendix B: Representing the GHZ-State

For a general spin-1/2 system with N sites, the Greenberger-Horne-Zeilinger (GHZ) state is described by [57]

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\uparrow\ldots\uparrow\rangle + |\downarrow\downarrow\ldots\downarrow\rangle\right). \tag{B1}$$

The GHZ state is a strongly entangled quantum state consisting of at least three spin-1/2 particles [77, 78]. To derive the weights representing such a GHZ state in the RBM parametrization, we first consider the case of $N = 3$ sites. We add $M = 2$ hidden neurons to the neural network and consider the parametrization of the basis state expansion coefficients as stated in Eq. (3) in the main text. This provides a set of four equations,

$$c\left(\mathbf{v}^z = [\pm1, \pm1, \pm1]; \mathcal{W}\right) = \exp\left[\pm(d_1 + d_2 + d_3)\right] 4\cosh\left[b_1 \pm (W_{1,1} + W_{2,1} + W_{3,1})\right]\cosh\left[b_2 \pm (W_{1,2} + W_{2,2} + W_{3,2})\right]$$
$$\overset{!}{=} \frac{1}{\sqrt{2}}, \tag{B2}$$

$$c\left(\mathbf{v}^z = [\pm1, \pm1, \mp1]; \mathcal{W}\right) = \exp\left[\pm(d_1 + d_2 - d_3)\right] 4\cosh\left[b_1 \pm (W_{1,1} + W_{2,1} - W_{3,1})\right]\cosh\left[b_2 \pm (W_{1,2} + W_{2,2} - W_{3,2})\right]$$
$$\overset{!}{=} 0, \tag{B3}$$

$$c\left(\mathbf{v}^z = [\pm1, \mp1, \pm1]; \mathcal{W}\right) = \exp\left[\pm(d_1 - d_2 + d_3)\right] 4\cosh\left[b_1 \pm (W_{1,1} - W_{2,1} + W_{3,1})\right]\cosh\left[b_2 \pm (W_{1,2} - W_{2,2} + W_{3,2})\right]$$
$$\overset{!}{=} 0, \tag{B4}$$

$$c\left(\mathbf{v}^z = [\mp1, \pm1, \pm1]; \mathcal{W}\right) = \exp\left[\pm(d_2 + d_3 - d_1)\right] 4\cosh\left[b_1 \pm (W_{2,1} + W_{3,1} - W_{1,1})\right]\cosh\left[b_2 \pm (W_{2,2} + W_{3,2} - W_{1,2})\right]$$
$$\overset{!}{=} 0. \tag{B5}$$

There are many ways to solve this set of equations, but in the following we only look for one possible solution.

Thus, from Eq. (B3) we choose

$$\cosh\left[b_1 \pm (W_{1,1} + W_{2,1} - W_{2,2})\right] \overset{!}{=} 0. \tag{B6}$$

From this it follows,

$$b_1 = i\pi \left(\frac{n_1^+ + n_1^-}{2} + \frac{1}{2} \right),$$

$$W_{1,1} + W_{2,1} - W_{3,1} = i\pi \frac{n_1^+ - n_1^-}{2}. \tag{B7}$$

Analogously, we can demand from Eq. (B4)

$$\cosh\left[b_2 \pm (W_{1,2} - W_{2,2} + W_{3,2}) \right] \overset{!}{=} 0$$

$$\Rightarrow b_2 = i\pi \left(\frac{n_2^+ + n_2^-}{2} + \frac{1}{2} \right), \quad n_2^\pm \in \mathbb{Z} \tag{B8}$$

$$W_{1,2} - W_{2,2} + W_{3,2} = i\pi \frac{n_2^+ - n_2^-}{2},$$

and from Eq. (B5)

$$\cosh\left[b_2 \pm (-W_{1,2} + W_{2,2} + W_{3,2}) \right] \overset{!}{=} 0$$

$$\Rightarrow -W_{1,2} + W_{2,2} + W_{3,2} = i\pi \frac{n_2^+ - n_2^-}{2}. \tag{B9}$$

As we only look for a single possible solution, we consider the simplest case with $n_i^\pm = 0$ for $i = 1, 2$. This yields

$$b_1 = b_2 = i\frac{\pi}{2},$$

$$W_{1,1} + W_{2,1} - W_{3,1} = 0, \tag{B10}$$

$$W_{3,2} = 0, \ W_{1,2} = W_{2,2}.$$

Plugging these results into Eq. (B2) and considering only the cosh-terms gives

$$\cosh\left[i\frac{\pi}{2} \pm (W_{1,1} + W_{2,1} + W_{3,1}) \right] \cosh\left[i\frac{\pi}{2} \pm 2W_{1,2} \right]$$

$$= \{\pm i\sinh\left[W_{1,1} + W_{2,1} + W_{3,1} \right]\} \{\pm i\sinh\left[2W_{1,2} \right]\}$$

$$= -\sinh\left[W_{1,1} + W_{2,1} + W_{3,1} \right] \sinh\left[2W_{1,2} \right]$$

$$= -i\sinh\left[W_{1,1} + W_{2,1} + W_{3,1} \right], \tag{B11}$$

where we again simplify the expression by choosing $W_{1,2} = i\pi/4$ in the last line. With this choice, Eq. (B2) becomes

$$-i\exp\left[\pm (d_1 + d_2 + d_3) \right] \sinh\left[W_{1,1} + W_{2,1} + W_{3,1} \right]$$

$$= -i\left(\cosh\left[d_1 + d_2 + d_3 \right] \pm i\sinh\left[a_1 + a_2 + a_3 \right] \right)$$

$$\times \sinh\left[W_{1,1} + W_{2,1} + W_{3,1} \right]$$

$$\overset{!}{=} \frac{1}{4\sqrt{2}} \tag{B12}$$

$$\Rightarrow \sinh\left[d_1 + d_2 + d_3 \right] \overset{!}{=} 0 \quad \Rightarrow d_1 + d_2 + d_3 = 0$$

$$\Rightarrow -i\sinh\left[W_{1,1} + W_{2,1} + W_{3,1} \right] \overset{!}{=} \frac{1}{4\sqrt{2}}$$

$$\Rightarrow W_{1,1} + W_{2,1} + W_{3,1} = i\arcsin\left[\frac{1}{4\sqrt{2}} \right]. \tag{B13}$$

Given these conditions, we can choose one possible solution with the weights

$$d_1 = d_2 = d_3 = 0,$$

$$b_1 = b_2 = i\frac{\pi}{2},$$

$$W_{1,1} = W_{2,1} = \frac{i}{4}\arcsin\left[\frac{1}{4\sqrt{2}} \right],$$

$$W_{3,1} = \frac{i}{2}\arcsin\left[\frac{1}{4\sqrt{2}} \right],$$

$$W_{1,2} = W_{2,2} = i\frac{\pi}{4}, \ W_{3,2} = 0. \tag{B14}$$

These are the weights we choose in the main text for $N = 3$ sites.

From these results we can see that the choice of the weights to the first hidden neuron guarantees the normalized coefficients if all spins are equal and the zero coefficients if the third spin is flipped compared to the other two. The weights to the second hidden neuron guarantee the zero coefficients for the case that spins one and two have opposite sign, so that all cases are covered. This behavior can be generalized to an arbitrary number N of spins and yields the weights to represent a GHZ state in a neural network with $M = N - 1$ hidden neurons,

$$d_j = 0, \qquad b_k = i\frac{\pi}{2},$$

$$W_{j \neq N, 1} = \frac{i}{2(N-1)}\arcsin\left[\frac{1}{2^{N-1/2}} \right],$$

$$W_{N,1} = \frac{i}{2}\arcsin\left[\frac{1}{2^{N-1/2}} \right], \tag{B15}$$

$$W_{j,k \neq 1} = i\frac{\pi}{4}\left(\delta_{j,k-1} + \delta_{j,k} \right),$$

$$\forall j \in \{1, \ldots, N\}, \ k \in \{1, \ldots, M\}.$$

To avoid confusion with the imaginary unit, we have replaced the indices i, j of the visible and hidden neurons by the indices j, k, respectively. These are the expressions quoted in Table I in the main text and it has been checked that this choice of weights represents the GHZ state for small systems, $N \leq 6$. The weights for the calculations with $N = 5$ sites in the main text are chosen according to these equations.

[1] P. Pfeuty, Ann. Phys. (NY) **57**, 79 (1970).

[2] P. Calabrese, F. H. L. Essler, and M. Fagotti, J. Stat. Mech: Theory Exp. **2012**, P07016 (2012).

[3] P. Calabrese, F. H. L. Essler, and M. Fagotti, J. Stat. Mech: Theory Exp. **2012**, P07022 (2012).

[4] E. Lieb, T. Schultz, and D. Mattis, Ann. Phys. **16**, 407 (1961).

[5] S. Sachdev, *Quantum Phase Transitions*, 2nd ed. (Cambridge University Press, 2011).

[6] U. Schollwöck, Ann. Phys. **326**, 96 (2011).

[7] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[8] G. Vidal, Phys. Rev. Lett. **93**, 040502 (2004).

[9] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, J. Stat. Mech.: Theory Exp. **2004**, P04005 (2004).

[10] S. Sharma, S. Suzuki, and A. Dutta, Phys. Rev. B **92**, 104306 (2015).

[11] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, Phys. Rev. B **94**, 165116 (2016).

[12] S. Sorella, M. Casula, and D. Rocca, J. Chem. Phys. **127**, 014105 (2007).

[13] S. Sorella, Phys. Rev. B **64**, 024512 (2001).

[14] A. Polkovnikov, Ann. Phys. **325**, 1790 (2010).

[15] J. Schachenmayer, A. Pikovski, and A. M. Rey, New J. Phys. **17**, 065009 (2015).

[16] J. Schachenmayer, A. Pikovski, and A. M. Rey, Phys. Rev. X **5**, 011022 (2015).

[17] L. Pucci, A. Roy, and M. Kastner, Phys. Rev. B **93**, 174302 (2016).

[18] R. P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).

[19] I. Bloch, J. Dalibard, and N. Sylvain, Nat. Phys. **8**, 267 (2012).

[20] R. Blatt and C. F. Roos, Nat. Phys. **8**, 277 (2012).

[21] S. Braun, M. Friesdorf, S. S. Hodgman, M. Schreiber, J. P. Ronzheimer, A. Riera, M. del Rey, I. Bloch, J. Eisert, and U. Schneider, PNAS **112**, 3641 (2015).

[22] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Nature **551**, 579 (2017).

[23] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe, Nature **551**, 601 (2017).

[24] M. Di Ventra and F. L. Traversa, J. Appl. Phys. **123**, 180901 (2018).

[25] M. A. Petrovici, *Form Versus Function: Theory and Models for Neuronal Substrates*, edited by S. I. Publishing (Springer International Publishing, 2016).

[26] M. A. Petrovici, J. Bill, I. Bytschok, J. Schemmel, and K. Meier, Phys. Rev. E **94**, 042312 (2016).

[27] A. F. Kungl, S. Schmitt, J. Klähn, P. Müller, A. Baumbach, D. Dold, A. Kugele, E. Müller, C. Koke, M. Kleider, C. Mauch, O. Breitwieser, L. Leng, N. Gürtler, M. Güttler, D. Husmann, K. Husmann, A. Hartel, V. Karasenko, A. Grübl, J. Schemmel, K. Meier, and M. A. Petrovici, arXiv:1807.02389v3 [cs.NE] (2019).

[28] T. Wunderlich, A. F. Kungl, E. Müller, A. Hartel, Y. Stradmann, S. A. Aamir, A. Grübl, A. Heimbrecht, K. Schreiber, D. Stöckel, C. Pehle, S. Billaudelle, G. Kiene, C. Mauch, J. Schemmel, K. Meier, and M. A. Petrovici, Front. Neurosci. **13**, 260 (2019).

[29] L. Kades and J. Pawlowski, arxiv:1901.05214 [cs.NE] (2019).

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

[31] H. Saito, J. Phys. Soc. Jpn. **86**, 093001 (2017).

[32] X. Gao and L.-M. Duan, Nat. Commun. **8**, (2017).

[33] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. X **7**, 021021 (2017).

[34] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Phys. Rev. B **96**, 205152 (2017).

[35] R. Kaubruegger, L. Pastori, and J. C. Budich, arXiv:1710.04713 [cond-mat.str-el] (2017).

[36] N. Freitas, G. Morigi, and V. Dunjko, Int. J. Quantum Inf. **16**, 1840008 (2018).

[37] G. Carleo, Y. Nomura, and M. Imada, Nat. Commun. **9** (2018).

[38] P. Teng, Phys. Rev. E **98**, 033305 (2018).

[39] H. Huang, Phys. Rev. E **98**, 062313 (2018).

[40] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, Phys. Rev. X **8**, 011006 (2018).

[41] S. R. Clark, J. Phys. A: Math. Theor. **51**, 135301 (2018).

[42] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, Phys. Rev. B **97**, 085104 (2018).

[43] S. Czischek, M. Gärttner, and T. Gasenzer, Phys. Rev. B **98**, 024311 (2018).

[44] C. Pehle, K. Meier, M. Oberthaler, and C. Wetterich, arXiv:1810.10335 [quant-ph] (2018).

[45] T. Westerhout, N. Astrakhantsev, K. S. Tikhonov, M. Katsnelson, and A. A. Bagrov, arXiv:1907.08186 [cond-mat.dis-nn] (2019).

[46] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. **14**, 447 (2018).

[47] G. Torlai and R. G. Melko, Phys. Rev. Lett. **120**, 240503 (2018).

[48] M. Troyer and U.-J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005).

[49] K. N. Anagnostopoulos and J. Nishimura, Phys. Rev. D **66**, 106008 (2002).

[50] T. Nakamura, N. Hatano, and H. Nishimori, J. Phys. Soc. Jpn. **61**, 3494 (1992).

[51] E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, Phys. Rev. B **41**, 9301 (1990).

[52] G. Torlai, J. Carrasquilla, M. T. Fishman, R. G. Melko, and M. P. A. Fisher, arXiv:1906.04654 [quant-ph] (2019).

[53] D. Hangleiter, I. Roth, D. Nagaj, and J. Eisert, arXiv:1906.02309 [quant-ph] (2019).

[54] J. S. Bell, Physics **1**, 195 (1964).

[55] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).

[56] J. S. Bell and A. Aspect, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*, 2nd ed. (Cambridge University Press, 2004).

[57] D. M. Greenberger, M. A. Horne, and A. Zeilinger, "Going beyond Bell's theorem," in *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, edited by M. Kafatos (Springer Netherlands, Dordrecht, 1989) pp. 69–72.

[58] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Sci. Rep. **7**, 8823 (2017).

[59] G. Torlai, B. Timar, E. P. L. van Nieuwenburg, H. Levine, A. Omran, A. Keesling, H. Bernien, M. Greiner, V. Vuleti, M. D. Lukin, R. G. Melko, and M. Endres, arXiv:1904.08441 [quant-ph] (2019).

[60] M. J. Hartmann and G. Carleo, Phys. Rev. Lett. **122**, 250502 (2019).

[61] A. Nagy and V. Savona, Phys. Rev. Lett. **122**, 250501 (2019).

[62] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, Phys. Rev. Lett. **122**, 250503 (2019).

[63] N. Yoshioka and R. Hamazaki, Phys. Rev. B **99**, 214306 (2019).

[64] S. Lu, X. Gao, and L.-M. Duan, Phys. Rev. B **99**, 155136 (2019).

[65] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, Nat. Mach. Intell. **1**, 155 (2019).

[66] G. Torlai and R. G. Melko, Phys. Rev. B **94**, 165134 (2016).

[67] G. E. Hinton, "A practical guide to training restricted Boltzmann machines," in *Neural Networks: Tricks of the Trade: Second Edition*, edited by G. Montavon, G. B. Orr, and K.-R. Müller (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012) pp. 599–619.

[68] M. Karl, H. Cakir, J. C. Halimeh, M. K. Oberthaler, M. Kastner, and T. Gasenzer, Phys. Rev. E **96**, 022110 (2017).

[69] I. D. Kivlichan, "On the complexity of stoquastic hamiltonians," (2015).

[70] S. Bravyi, D. P. DiVencenzo, R. Oliveira, and B. M. Terhal, Quant. Inf. Comp. **8**, 0361 (2008).

[71] F. Iglói and Y.-C. Lin, J. Stat. Mech.: Theory Exp. **2008**, P06004 (2008).

[72] K. Sohn and H. Lee, in *Proceedings of the 29th International Conference on Machine Learning*, ICML'12 (Omnipress, USA, 2012) pp. 1339–1346.

[73] R. E. Caflisch, Acta Numer. **7**, 149 (1998).

[74] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[75] J. F. Clauser and M. A. Horne, Phys. Rev. D **10**, 526 (1974).

[76] B. S. Cirel'son, Lett. Math. Phys. **4**, 93 (1980).

[77] W. Dür, G. Vidal, and J. I. Cirac, Phys. Rev. A **62**, 062314 (2000).

[78] N. Gisin and H. Bechmann-Pasquinucci, Phys. Lett. A **246**, 1 (1998).

[79] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, arXiv:1903.10563 [physics.comp-ph] (2019).