Variational classical networks for dynamics in interacting quantum matter

Roberto Verdel, Markus Schmitt, Yi-Ping Huang, Petr Karpov, Markus Heyl

Variational classical networks for dynamics in interacting quantum matter

Roberto Verdel [●],[1] Markus Schmitt [●],[2] Yi-Ping Huang [●],[3,4] Petr Karpov [●],[1,5,*] and Markus Heyl [●][1]

[1]*Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Straße 38, Dresden 01187, Germany*
[2]*University of California, Berkeley, California 94720, USA*
[3]*Paul Scherrer Institute, Forschungsstrasse 111, 5232 Villigen, Switzerland*
[4]*Department of Physics, National Tsing Hua University, Hsinchu 30013, Taiwan*
[5]*National University of Science and Technology "MISiS", Moscow 119991, Russia*

Dynamics in correlated quantum matter is a hard problem, as its exact solution generally involves a computational effort that grows exponentially with the number of constituents. While remarkable progress has been witnessed in recent years for one-dimensional systems, much less has been achieved for interacting quantum models in higher dimensions, since they incorporate an additional layer of complexity. In this work, we employ a variational method that allows for an efficient and controlled computation of the dynamics of quantum many-body systems in one and higher dimensions. The approach presented here introduces a variational class of wave functions based on complex networks of classical spins akin to artificial neural networks, which can be constructed in a controlled fashion. We provide a detailed prescription for such constructions and illustrate their performance by studying quantum quenches in one- and two-dimensional models. In particular, we investigate the nonequilibrium dynamics of a genuinely interacting two-dimensional lattice gauge theory, the quantum link model, for which we have recently shown—employing the technique discussed thoroughly in this paper—that it features disorder-free localization dynamics [P. Karpov *et al.*, Phys. Rev. Lett. **126**, 130401 (2021)]. The present work not only supplies a framework to address purely theoretical questions but also could be used to provide a theoretical description of experiments in quantum simulators, which have recently seen an increased effort targeting two-dimensional geometries. Importantly, our method can be applied to any quantum many-body system with a well-defined classical limit.

I. INTRODUCTION

One of the main challenges in quantum many-body dynamics is that unless the model under study is exactly solvable, the numerical overhead required to find an exact solution grows, in general, exponentially with the number of degrees of freedom. In the last decades the development of powerful computational techniques has nevertheless seen impressive progress, largely motivated by the experimental advances in realizing and controlling isolated quantum systems far away from equilibrium [1–16]. The majority of the advances have been achieved for one-dimensional (1D) systems, for which there exist now a set of reliable methods that can simulate efficiently the dynamics of lattice models. The primary example of such set of techniques is tensor network algorithms such as the time-dependent density matrix renormalization group (or its variants TEBD and tMPS) [17–24], which uses a matrix product state [25–27] representation of the wave function and solves the dynamics, for instance, via a Trotter decomposition of the evolution operator. Yet, this approach is generally restricted due to a rapid growth of entanglement, and a substantial increment of its computational complexity in higher dimensions. A recent alternative consists of encoding quantum states in various types of networks of classical degrees of freedom, such as artificial neural networks (ANNs) [28–30] and perturbative classical networks (pCNs) [31]. On the other hand, the description of quantum dynamics in higher dimensions is facing severe limitations. In spite of a few very recent efforts in two dimensions (2D) using tensor networks [32–41], artificial neural networks [30,42,43], or numerical linked cluster expansion [44], solving the quantum dynamics of 2D (and higher-dimensional) interacting systems remains one of the central challenges in computational quantum physics.

In this work, we introduce a numerical framework that allows for an efficient solution of the dynamics of quantum lattice models in one and higher dimensions. In our method, the many-body wave function is represented as a complex network of classical spins akin to ANNs, with couplings among the spins that are taken as variational parameters, and which are then optimized via a time-dependent variational principle (TDVP) [28,45], with the key advantage that the classical networks do not face numerical instabilities as they have been observed for ANNs [30]. The architecture of these *variational classical networks* (VCNs) can be derived systematically, in a fashion similar to that in the case of their relative pCNs [31].

*Corresponding author: karpov.petr@gmail.com

A crucial property of VCNs is that they inherit the *controlled* character that arises from the perturbative nature of pCNs. In addition, the optimization step introduced with the TDVP allows us to mitigate several inherent drawbacks of pCNs, such as being forcibly limited to weak quantum fluctuations and short timescales. Furthermore, there exist situations, as discussed subsequently, where VCNs may bear a reduced computational complexity compared to similar state-of-the-art techniques, while still yielding sufficiently accurate results.

We employ our technique to study the dynamics of two quantum spin systems: First, we consider various quenches in the paradigmatic 1D transverse field Ising model (TFIM), which is exactly solvable, thereby serving as testing ground for our method. Next, we tackle the more challenging case of a genuinely interacting 2D lattice gauge theory, namely, the quantum link model (QLM). As shown in our recent paper [46], this system features disorder-free localization [47–51], a mechanism for ergodicity breaking in homogeneous systems due to local constraints imposed by gauge invariance. We find that the methodology presented here is particularly well suited for describing the dynamics of the 2D QLM in the nonergodic regime, allowing us to probe timescales and system sizes that could hardly be accessed with any other state-of-the-art computational scheme.

The present work serves as a companion paper to Ref. [46], where we briefly introduced the VCN method but only for the specific case of the 2D QLM and omitted some technical details. Here we present the method in its full generality and show how it can be applied to any interacting quantum system in arbitrary dimensions, as long as a classical limit can be defined for the model under consideration.

The outline of the remainder of this paper is as follows: In Sec. II we state the general problem and settings for the construction and subsequent application of our method. Further, we show how to derive, in general, the structure of VCNs (Sec. II B), and recall the way in which the TDVP operates (Sec. II C). Next, using our methodology we simulate quantum quenches in the paradigmatic 1D TFIM in Sec. III, and in the genuinely interacting 2D QLM in Sec. IV. Some concluding remarks, including possible applications to recent experiments in quantum simulators, are discussed in the last section.

II. THE METHOD

In this section, we explain the basic idea of our method. It contains two crucial ingredients: (i) an efficient compression of the wave function in terms of networks of classical degrees of freedom, and (ii) a suitable procedure to optimize such networks so as to get an accurate description of the encoded time-dependent quantum state.

A. General problem and generative machines

Throughout this work, we shall consider systems of N spin-1/2 degrees of freedom. We work in the computational basis of spin configurations $s = (s_1, s_2, \ldots, s_N)$, with $s_i = \uparrow, \downarrow$. In this representation, the state vector can be expanded as

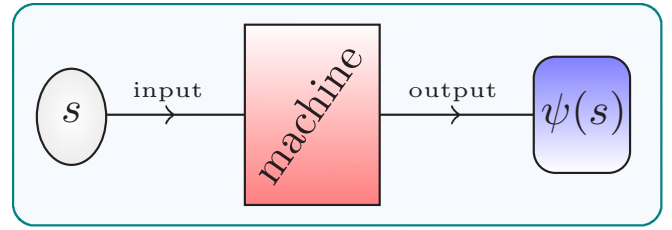$$|\psi(t)\rangle = \sum_{\{s\}} \psi(s, t)|s\rangle, \qquad (1)$$

FIG. 1. Schematic representation of a generative machine that takes as input a spin configuration s, and computes on-the-fly an approximation of the corresponding wave function amplitude $\psi(s)$. Such machine can be, for example, an ANN or, as employed in this work, a variational classical network (VCN).

where the amplitudes $\psi(s, t) \equiv \langle s|\psi(t)\rangle$ contain the full information about the system, from which, in principle, all physical quantities can be computed. However, there is a fundamental computational limitation associated with the expansion (1), as it involves exponentially many terms.

A possible way to overcome this difficulty consists of using a generative machine that approximates the wave function *on the fly*, rather than storing all the individual coefficients; see Fig. 1. This approach received increased attention recently, when general-purpose ANNs were proposed as generative models [28]. Here, we provide an alternative class of generative machines akin to ANNs, which can be constructed according to a controlled prescription as explained in the following. For the time being, however, let us consider a generic generative machine $\psi_\eta(s)$, such that

$$\psi(s, t) \approx \psi_\eta(s), \qquad (2)$$

where η refers to a set of complex parameters that carry the time dependence, i.e., $\eta \equiv \eta(t) = (\eta_1(t), \eta_2(t), \ldots, \eta_K(t)) \in \mathbb{C}^K$, and which are optimized variationally.

A representation of the wave function in terms of such generative machine renders the problem of computing physical quantities tractable. In effect, the expectation value of an observable O with matrix elements $\langle s|O|s'\rangle = O_{s,s'}$, can be written as

$$\langle \psi_\eta|O|\psi_\eta\rangle = \sum_{\{s\}} |\psi_\eta(s)|^2 O_\eta(s), \qquad (3)$$

where $O_\eta(s) = \sum_{\{s'\}} O_{s,s'}\psi_\eta(s')/\psi_\eta(s)$. Typical local observables are such that $\langle s|O|s'\rangle$ is sparse. Consequently, getting $O_\eta(s)$ requires only a polynomial computational overhead. Hence, assuming a normalized wave function, the expectation value (3) can be calculated efficiently via a Monte Carlo sampling of the distribution $|\psi_\eta(s)|^2$. Note that a compression of the wave function such as Eq. (2) will be efficient as long as the overall number K of parameters is significantly less than the dimension of the Hilbert space.

Let us also note that, in general, a generative machine refers to a model that can generate samples according to some target distribution, which in this case is $|\psi(s, t)|^2$. Remarkably, generative machines such as that in Eq. (2), which we shall regard in the following, not only achieve the task mentioned above but also give direct access to the complex amplitudes $\psi(s, t)$.

B. Variational classical networks

As mentioned before, this work aims to introduce a class of adequate generative machines to represent the many-body wave function. In the following, we give a controlled prescription to construct such generative models.

1. General settings

Let us consider a Hamiltonian of the form

$$H = H_0 + \gamma V, \tag{4}$$

where H_0 represents a classical system in the sense that it is diagonal in the computational basis,

$$H_0|s\rangle = E_s|s\rangle, \tag{5}$$

and the off-diagonal perturbation γV, with γ playing the role of a small parameter, accounts for quantum fluctuations.

We are interested in the nonequilibrium dynamics generated by the Hamiltonian (4). This can be obtained by solving the time-dependent Schrödinger equation, which admits the formal solution (in units such that $\hbar = 1$)

$$|\psi(t)\rangle = e^{-iHt}|\psi_0\rangle, \tag{6}$$

where $|\psi_0\rangle$ denotes the initial state. In general, it is challenging to determine the action of the evolution operator e^{-iHt}, onto the basis vectors. However, whenever the Hamiltonian H can be split as in Eq. (4), it is possible to carry out a perturbative treatment by working in the interaction picture, in which the evolution operator can be written as

$$e^{-iHt} = e^{-iH_0 t} W_\gamma(t), \tag{7}$$

where

$$W_\gamma(t) = \mathcal{T} \exp\left[-i\gamma \int_0^t dt' V(t')\right], \tag{8}$$

where $\mathcal{T}$ is the time-ordering operator, and with $V(t)$ satisfying the equation of motion:

$$-i\frac{d}{dt} V(t) = [H_0, V(t)]. \tag{9}$$

Within these settings, the many-body wave function amplitudes are given by

$$\psi(s, t) = e^{-iE_s t} \langle s|W_\gamma(t)|\psi_0\rangle. \tag{10}$$

The task now is to calculate the right-hand side Eq. (10). Classical networks provide a possible solution, as detailed below.

2. Cumulant expansion and pCNs

The right-hand side of Eq. (10) can be computed in a controlled way by means of a cumulant expansion [52], namely,

$$\langle s|W_\gamma(t)|\psi_0\rangle = \langle s|\psi_0\rangle \exp\left[\sum_{n=1}^{\infty} \frac{(-i\gamma)^n}{n!} \int_0^t dt_1 \int_0^t dt_2\right.$$
$$\left. \times \cdots \int_0^t dt_n \langle \mathcal{T} V(t_1)V(t_2)\cdots V(t_n)\rangle_c\right], \tag{11}$$

where $\langle \cdot \rangle_c$ denotes the cumulant average. For example, for the lowest-order corrections, we have

$$\langle A \rangle_c \equiv \frac{\langle s|A|\psi_0\rangle}{\langle s|\psi_0\rangle}, \tag{12}$$

$$\langle AB \rangle_c \equiv \frac{\langle s|AB|\psi_0\rangle}{\langle s|\psi_0\rangle} - \frac{\langle s|A|\psi_0\rangle \langle s|B|\psi_0\rangle}{\langle s|\psi_0\rangle^2}. \tag{13}$$

This expansion allows us to write down the wave function as

$$\psi(s, t) = e^{\mathcal{H}_{\text{eff}}(s,t)}, \tag{14}$$

with $\mathcal{H}_{\text{eff}}(s, t)$ defined by Eqs. (10) and (11). Representations of the wave function in the form of a Boltzmann-like factor are quite adequate to compute physical observables via a Monte Carlo procedure [31,53–55]. In order to gain some insight about the physical content of the function $\mathcal{H}_{\text{eff}}(s, t)$, let us restrict ourselves for a moment to a simple initial product state, namely, an equally weighted superposition of the spin configurations:

$$|\psi_0\rangle = |\rightarrow\rangle \equiv \bigotimes_{i=1}^{N} \frac{1}{\sqrt{2}}[|\uparrow\rangle_i + |\downarrow\rangle_i]. \tag{15}$$

This initial state is particularly convenient as $\psi_0(s) = 2^{-N/2}$ for all s, and hence $\psi_0(s)$ drops out in all the cumulant averages. In this scenario, and upon performing the integrals in the cumulant expansion (11), the function $\mathcal{H}_{\text{eff}}$ adopts, in general, the following form:

$$\mathcal{H}_{\text{eff}}(s, t) = \sum_l C_l(t)\Phi_l(s). \tag{16}$$

That is, $\mathcal{H}_{\text{eff}}$ can be regarded as the effective Hamiltonian of a classical spin system with complex couplings $C_l(t)$, and with spin interactions given by the functions $\Phi_l(s)$, which are local provided that the quantum Hamiltonian is local, too. Situations where quenches from the initial state (15) are of physical interest are discussed in posterior sections.

In some cases, it is possible to recast the systems defined by Eq. (16), as conventional classical statistical mechanical models. For example, the effective model corresponding to the 1D TFIM, discussed later in detail (see Sec. III), contains, up to first order in the cumulant expansion, the following terms [31]:

$$\Phi_1(s) = \sum_i s_i s_{i+1}, \quad \Phi_2(s) = \sum_i s_i s_{i+2}, \tag{17}$$

which define a 1D classical Ising model with nearest and next-to-nearest neighbor interactions.

On the other hand, the system (16) can also be visualized as a *network* of classical spins with connectivity specified by the functions $\Phi_l(s)$. Hence, the models defined by an effective Hamiltonian as in Eq. (16) are called perturbative classical networks (pCNs) [31]. In Fig. 2, we display the network representation, up to second order, of the classical spin model that emerges when considering a translationally invariant 1D TFIM.

Let us emphasize that our approach works as well for initial states other than the equally weighted superposition (15). Importantly, this is true not only for translationally invariant initial states but also for nonuniform ones, as will be shown
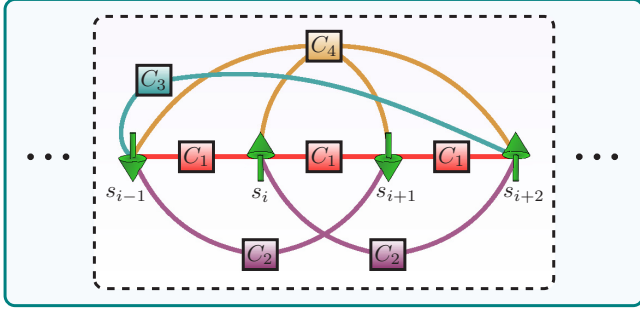
FIG. 2. Local structure of a classical network for the 1D TFIM in a translationally invariant lattice, containing up to second-order terms. The network's connectivity is defined by functions such as those in Eq. (17) [see also Eqs. (34) and (35), for this particular example]. The nodes of the network correspond to classical spins s_i, whereas the links of the network are specified by the couplings C_l.

later. To consider a more general case, let us assume that the perturbation consists of a sum of local terms, namely, $V = \sum_\alpha v_\alpha$, where the overall number of terms is polynomial in system size. Thus, we can write an equation of motion for each of the individual terms:

$$-i\frac{d}{dt}v_\alpha(t) = [H_0, v_\alpha(t)] = e^{iH_0 t}[H_0, v_\alpha]e^{-iH_0 t}. \quad (18)$$

The commutator $[H_0, v_\alpha]$ measures, essentially, the energy difference between two eigenstates of H_0 in a transition induced by v_α, i.e., $v_\alpha|s_1\rangle = |s_2\rangle$. Indeed, one can readily prove that

$$[H_0, v_\alpha]|s_1\rangle = (E_{s_2} - E_{s_1})v_\alpha|s_1\rangle. \quad (19)$$

Therefore, Eq. (18) can be rewritten as

$$\frac{d}{dt}v_\alpha(t) = i\Omega_\alpha v_\alpha(t), \quad (20)$$

with Ω_α being a diagonal operator in the computational basis, which measures the energy difference in a transition induced by v_α. This equation admits the formal solution $v_\alpha(t) = e^{i\Omega_\alpha t}v_\alpha$. Thus, if the effective Hamiltonian is written as $\mathcal{H}_{\text{eff}} = \sum_{n=0}^\infty \mathcal{H}^{(n)}$, where the zeroth-order term is $\mathcal{H}^{(0)}(s, t) = -iE_s t + \ln[\psi_0(s)]$, and the subsequent orders are defined by Eq. (11), one can write, for example, the first-order correction as

$$\mathcal{H}^{(1)}(s, t) = -i\gamma \sum_{\{s'\}} \frac{\psi_0(s')}{\psi_0(s)} \sum_\alpha \langle s|v_\alpha|s'\rangle \int_0^t dt' e^{i\Omega_\alpha(s)t'}, \quad (21)$$

and likewise for higher-order terms. Note that the matrix $\langle s|v_\alpha|s'\rangle$ is typically sparse for physical systems with few-body couplings.

Let us remark that the cumulant expansion (11) goes beyond conventional time-dependent perturbation theory, since the corrections considered here effectively account for a resummation of several terms that appear in a standard perturbative expansion [31]. However, pCNs face their own limitations, too. In particular, they are inherently restricted to weak quantum fluctuations (small γ) to ensure that we can safely truncate the expansion (11). Besides, the description of the evolution of observables will eventually break down, since

resonant processes may be present, giving rise to secular terms that limit a correct description to timescales of order $\mathcal{O}(1/\gamma)$ [31]. Nonetheless, one can still benefit from the framework introduced here, while mitigating the drawbacks mentioned before. This is achieved by constructing adequate *variational* wave functions with a network architecture that is inherited from a corresponding pCN, as argued in the following.

3. Variational ansatz

The main idea of this work was already outlined in the previous paragraph: building upon the structure of an underlying pCN, one can construct variational classical networks (VCNs) that are then used as generative machines to compute the dynamics. Importantly, the resulting VCNs will inherit the *controlled* character of the cumulant expansion, in that the accuracy of the approximation can be improved systematically by reducing the value of γ, or by taking into account higher-order cumulants.

First, let us consider classical networks of the form given in Eq. (16). The corresponding VCN can be obtained simply by letting the couplings C_l be variational parameters, that is,

$$\mathcal{H}_{\text{VCN}}(s; \eta(t)) = \sum_l \eta_l(t)\Phi_l(s), \quad (22)$$

where $\eta(t)$ denotes a set of complex variational parameters.

In the more general case such as Eq. (21), one can build the corresponding VCN by noting that $\Omega_\alpha(s)$ take a finite number of discrete values for any s. Thus one can simply introduce a variational parameter for each value of $\Omega_\alpha(s)$. To fix ideas let us consider the first-order correction given in Eq. (21), and let us denote as Λ_{Ω_α} the set of all possible values of $\Omega_\alpha(s)$. After rewriting the integral

$$\int_0^t dt' e^{i\Omega_\alpha(s)t'} = \sum_{\Omega \in \Lambda_{\Omega_\alpha}} \delta_{\Omega_\alpha(s),\Omega} \int_0^t dt' e^{i\Omega t'}, \quad (23)$$

we can introduce a set of variational parameters so that the corresponding first-order variational effective Hamiltonian reads

$$\mathcal{H}_{\text{VCN}}^{(1)}(s; \eta(t)) = -i\gamma \sum_{\{s'\}} \frac{\psi_0(s')}{\psi_0(s)} \sum_\alpha \langle s|v_\alpha|s'\rangle$$
$$\times \sum_{\Omega \in \Lambda_{\Omega_\alpha}} \delta_{\Omega_\alpha(s),\Omega} \eta_\Omega^{(1)}(t), \quad (24)$$

and likewise for higher-order terms. In either case, the concomitant wave function amplitudes take the form

$$\psi_{\text{VCN}}(s; \eta(t)) = e^{\mathcal{H}_{\text{VCN}}(s;\eta(t))}. \quad (25)$$

Let us point out that the more higher-order corrections are included in the architecture of a classical network, the more quantum correlations can, in principle, be taken into account. Thus, the cumulant expansion (11) provides us with a controlled procedure to generate generative machines of the form given in Eq. (25), which allows for a systematic addition of terms that can potentially encode more and more nonlocal quantum correlations. Before turning to the applications, let us discuss the variational procedure that is used to optimize the resulting VCNs.

C. Time-dependent variational principle

The TDVP [28,45,53,54] is a way of optimizing a time-dependent variational ansatz $\psi_\eta(s)$, where η denotes a set of complex time-dependent variational parameters, i.e., $\eta(t) = (\eta_1(t), \eta_2(t), \ldots, \eta_K(t))$. Such trial wave function could be, e.g., a Jastrow ansatz [56,57], an ANN [28], or, as presented in this work, a VCN. In essence, the TDVP is a procedure that establishes an equivalence between the time-dependent Schrödinger equation and a system of first-order differential equations that govern the dynamics of the variational parameters, namely,

$$\sum_{k'} \mathcal{S}_{k,k'} \dot{\eta}_{k'} = -iF_k, \qquad (26)$$

where the overdot denotes differentiation with respect to time and with the following definitions:

$$\mathcal{S}_{k,k'} := \langle O_k^* O_{k'} \rangle - \langle O_k^* \rangle \langle O_{k'} \rangle, \qquad (27)$$

which is the so-called covariance matrix, and

$$F_k := \langle E_{\mathrm{loc}} O_k^* \rangle - \langle E_{\mathrm{loc}} \rangle \langle O_k^* \rangle. \qquad (28)$$

These quantities are expressed in terms of the local energy $E_{\mathrm{loc}}(s) := \frac{\langle s|H|\psi_\eta \rangle}{\langle s|\psi_\eta \rangle}$, and the variational derivatives, $O_k(s) := \frac{\partial \ln \psi_\eta(s)}{\partial \eta_k}$.

In order to quantify the accuracy of this TDVP, let us introduce the Fubini-Study metric $\mathcal{D}_{\mathrm{FS}}$, which measures the distance between the exact evolution during a small time interval δt: $e^{-i\delta t H}|\psi_\eta\rangle$, and the variational evolution $|\psi_{\eta+\delta\dot{\eta}}\rangle$. Its definition is the following:

$$\mathcal{D}_{\mathrm{FS}}(\varphi, \phi)^2 := \arccos\left(\sqrt{\frac{\langle\varphi|\phi\rangle\langle\phi|\varphi\rangle}{\langle\varphi|\varphi\rangle\langle\phi|\phi\rangle}}\right)^2. \qquad (29)$$

Thus, one can define a relative residual error as [28,30]

$$r^2(t) := \frac{\mathcal{D}_{\mathrm{FS}}(|\psi_{\eta+\delta\dot{\eta}}\rangle, e^{-i\delta t H}|\psi_\eta\rangle)^2}{\mathcal{D}_{\mathrm{FS}}(|\psi_\eta\rangle, e^{-i\delta t H}|\psi_\eta\rangle)^2}, \qquad (30)$$

which can be measured, too, by performing a Monte Carlo sampling of $|\psi_\eta(s)|^2$. In practice, we use a second-order expansion of (29) to compute Eq. (30). Moreover, in the following, we shall consider the integrated residual error $R^2(t) := \int_0^t dt' r^2(t')$. Note that Eq. (26) can be derived by minimizing the numerator in Eq. (30) with respect to $\dot{\eta}^*$ [28,30].

III. QUENCHES IN THE 1D TFIM

A. Model

As a first illustration of our method, we study several quantum quenches in the archetypal 1D TFIM, whose Hamiltonian for N spins on a ring reads

$$H_{\mathrm{TFIM}} = -J \sum_{i=1}^{N} \sigma_i^z \sigma_{i+1}^z - h \sum_{i=1}^{N} \sigma_i^x, \qquad (31)$$

where σ_i^μ ($\mu = x, y, z$) are the Pauli matrices at site i, $J > 0$ is the exchange constant that sets the overall energy scale, and h is a transverse magnetic field.

Let us recall some features of the model in Eq. (31). First of all, the 1D TFIM is integrable by means of a Jordan-Wigner transformation [58]; hence, comparison with analytical solutions is at our disposal. Moreover, this model features both equilibrium [59] and dynamical [60] quantum phase transitions. Indeed, the Hamiltonian (31) undergoes an equilibrium quantum phase transition at $h_c/J = 1$ [59], where the critical point separates a ferromagnetic phase ($h < h_c$) from a paramagnetic one ($h > h_c$). Its dynamical quantum phase transition (DQPT) is signaled by non-analyticities in the many-body dynamics [60,61], and occurs when quenching across the underlying equilibrium quantum phase transition. In this respect, it is interesting to study quenches that cross the critical point. Below, we shall consider quenches from the paramagnetic point $h_0 = \infty$, which corresponds to the initial state in Eq. (15), to both the ferromagnetic and the paramagnetic phases (see details below). Note that the first type of quench is precisely of interest for the study of DQPTs. From an experimental viewpoint, both probing the dynamics of this model and engineering the relevant initial state in Eq. (15) are now feasible tasks with current technologies in quantum simulators in various settings [4,9,10].

Before discussing the details of the quench dynamics simulations and the corresponding results, let us first construct the VCNs associated to the 1D TFIM.

B. VCNs for the 1D TFIM

The corresponding pCNs for TFIMs have been recently derived elsewhere [31]. Here, we review the main steps of such calculations. First of all, according to the general settings established in Sec. II B, we take the Ising interaction term as reference Hamiltonian, i.e., $H_0^{\mathrm{TFIM}} = -J \sum_{i=1}^{N} \sigma_i^z \sigma_j^z$, and the transverse field as the perturbation, namely, $\gamma V^{\mathrm{TFIM}} = -h \sum_i \sigma_i^x$ (identifying $\gamma = -h$).

Using the basic commutation relations of the Pauli matrices, one can readily show that

$$[H_0^{\mathrm{TFIM}}, \sigma_j^x] = -2J(\sigma_{j-1}^z + \sigma_{j+1}^z)\sigma_j^z\sigma_j^x, \qquad (32)$$

where we emphasize again that this commutator measures the change in energy in a transition induced by σ_j^x, between eigenstates of H_0^{TFIM}. The solution to the equation of motion Eq. (20) for $\sigma_j^x(t)$ therefore reads

$$\sigma_j^x(t) = e^{-2i(\sigma_{j-1}^z + \sigma_{j+1}^z)\sigma_j^z Jt} \sigma_i^x$$

$$= \left[\cos^2(2Jt) - \sin^2(2Jt)\sigma_{j-1}^z\sigma_{j+1}^z \right.$$

$$\left. - \frac{i}{2}\sin(4Jt)(\sigma_{j-1}^z + \sigma_{j+1}^z)\sigma_j^z\right]\sigma_i^x, \qquad (33)$$

where the second step follows from Euler's formula. This solution leads to a first-order pCN of the form anticipated in Eq. (17). Indeed, plugging the solution (33) in Eq. (21) and using the fact that $\psi_0(s) = 2^{-N/2}$ for all s, for the initial state in Eq. (15), one gets

$$\mathcal{H}_{\mathrm{TFIM}}^{(1)} = C_0^{(1)}(t)N + C_1^{(1)}(t)\sum_{i=1}^{N} s_i s_{i+1} + C_2^{(1)}(t)\sum_{i=1}^{N} s_i s_{i+2}, \qquad (34)$$

where the explicit form of the coefficients $C_l^{(1)}$ can be easily deduced from Eqs. (33) and (21). As previously explained, the classical network defined above can be turned into a VCN, simply by regarding the couplings $C_l^{(1)}$ as variational parameters.

The perturbatively motivated structure of the VCN can be systematically expanded by straightforwardly plugging Eq. (33) into higher-order terms in Eq. (11). By potentiating $\sum_j \sigma_j^x(t)$, more and more nonlocal couplings are generated. In fact, at order k couplings up to distance $k+1$ are generated (see Ref. [31] for details). Hence, we can systematically increase the VCN by adding all distinct classical coupling terms up to a given distance d, which are compatible with the system's symmetries. For example, when considering second-order corrections, the terms

$$\mathcal{H}_{\text{TFIM}}^{(2)} = C_1^{(2)}(t) \sum_{i=1}^{N} s_{i-1} s_i s_{i+1} s_{i+2} + C_2^{(2)}(t) \sum_{i=1}^{N} s_i s_{i+3},$$
(35)

that respect $\mathbb{Z}_2$ and lattice symmetries, and which expand up to a distance $d = 3$, would be added to the effective Hamiltonian, $\mathcal{H}_{\text{TFIM}} = \mathcal{H}_{\text{TFIM}}^{(1)} + \mathcal{H}_{\text{TFIM}}^{(2)}$ (see Fig. 2).

Notice that the possible number of coupling terms in the classical network at $d = N$ equals the dimension of the Hilbert space of the quantum system. In the presence of translational invariance, lattice-inversion symmetry, and $\mathbb{Z}_2$ symmetry, the symmetry-allowed couplings can be obtained by generating the corresponding symmetry-reduced computational basis and then identifying the domain wall configuration in each computational basis state with a coupling term in the variational wave function. In the absence of $\mathbb{Z}_2$ symmetry the computational basis configurations themselves correspond to coupling terms.

Finally, let us point out that the classical networks presented here for the 1D TFIM can be mapped onto certain types of ANNs [31].

C. Quench protocol and results

As mentioned before, quenches in the 1D TFIM across the critical point comprise a DQPT. In that respect, an interesting class of quenches consists of going from the paramagnetic to the ferromagnetic phase. Here, we concentrate precisely on the aforementioned situation as well as on quenches within the paramagnetic phase. In particular, we consider the initial state $|\rightarrow\rangle$ given in Eq. (15), which corresponds to the point $h_0 = \infty$. Next, we compute the unitary dynamics generated by the Hamiltonian in Eq. (31) with $h/J < 1$ (ferromagnetic) and $h/J > 1$ (paramagnetic).

We compare results for the dynamics of the TFIM obtained in three different ways: exact, pCN, and VCN. For the exact solution we exploit the integrability of the model. Via a Jordan-Wigner transformation the spin system is mapped to a model of noninteracting fermions [58], for which closed-form expressions can be obtained for all quantities of interest [59]. The results shown are for a translationally invariant chain in the thermodynamic limit. With pCN and VCN, we consider systems with $N = 50$ sites and periodic boundary conditions. On the timescales shown there is no finite-size effect in the
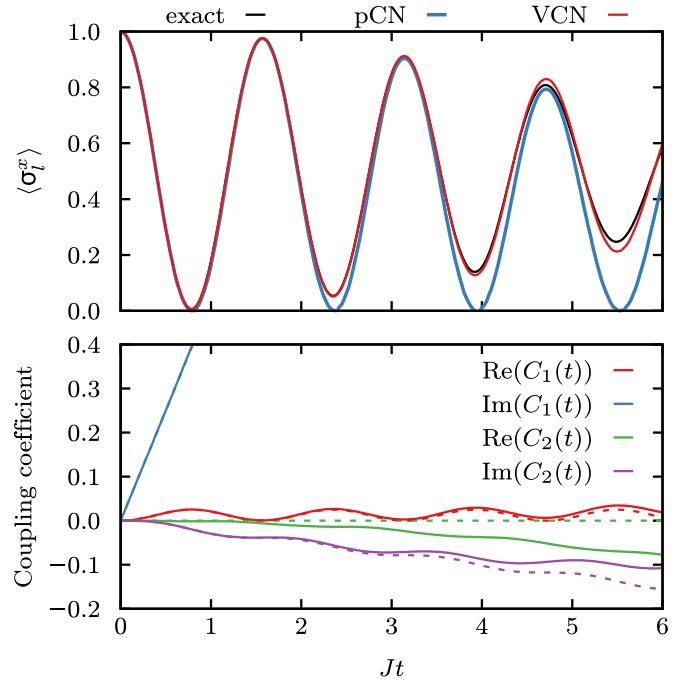

FIG. 3. Comparison of first-order pCN, first-order VCN, and the exact solution of the TFIM with $N = 50$, $h/J = 0.1$. Top: Dynamics of the transverse magnetization $\langle \sigma_l^x \rangle$. The black curve shows the exact solution obtained via fermionization in the thermodynamic limit. Bottom: Evolution of the perturbative (dashed lines) and variational (solid lines) couplings of the classical network defined in Eq. (34).

observables. To obtain the time-evolved VCN we initialize all network couplings with zero and integrate the TDVP equation using a second-order consistent integrator with adaptive time step. Expectation values with respect to $|\psi(s)|^2$ are estimated using 8×10^4 samples generated by a single-spin-flip Markov chain Monte Carlo.

The results of the quench dynamics are shown in Figs. 3–6. First, in Fig. 3, we compare the performance of the first-order pCN given in Eq. (34) and its associated VCN, in a quench to $h/J = 0.1$. As illustrated for the dynamics of the transverse magnetization $\langle \sigma_l^x \rangle$, both approaches capture very accurately the short-time behavior. However, it is the variational ansatz that yields a much more accurate description at longer times. Interestingly, when looking at the evolution of the perturbative and variational couplings, Fig. 3 (bottom), we can see that their dynamics start to differ approximately at the point where discrepancies in the evolution of observables are first noted.

In Fig. 4, we study the overall performance of various VCNs with different coupling distance d, when quenching to the ferromagnetic phase ($h/J = 0.3$) and the paramagnetic one ($h/J = 3$), left and right columns in Fig. 4, respectively. As a principal result, we observe that the accuracy is systematically improved upon increasing the coupling distance of the VCNs. This is not only observed from the real-time evolution of the transverse magnetization $\langle \sigma_l^x \rangle$, and the next-to-nearest neighbor correlation function $\langle \sigma_l^z \sigma_{l+2}^z \rangle$, but also from the integrated residuals $R^2(t)$, which show a systematic error convergence by increasing d. Next, focusing on the transverse magnetization $\langle \sigma_l^x \rangle$, we see that, in both quenches, it quickly relaxes to a steady-state value: While at weak
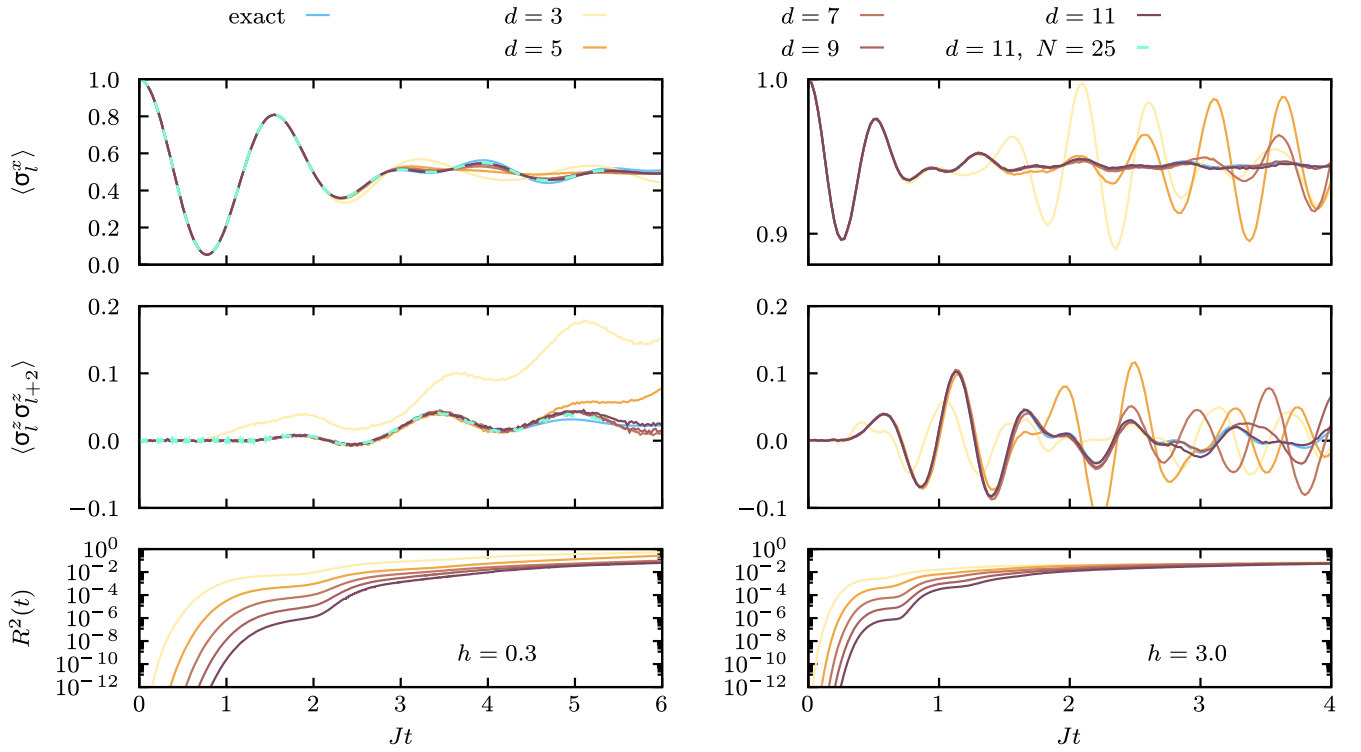
FIG. 4. Dynamics of the TFIM in quenches from $h_0 = \infty$ to $h/J = 0.3$ (left column) and $h/J = 3$ (right column), with $N = 50$. Convergence of VCN solution to the exact dynamics is shown as a function of the coupling distance d (see main text). Upper panels: Transverse magnetization $\langle \sigma_l^x \rangle$. Middle panels: Next-to-nearest neighbor correlation function $\langle \sigma_l^z \sigma_{l+2}^z \rangle$. Lower panels: Integrated relative residual $R^2(t)$. Results for a different system size ($N = 25$) are also shown in the quench to $h/J = 0.3$ (left column) for the VCN with $d = 11$.

transverse field this feature can be well captured by all the considered VCNs, the situation becomes more challenging when the value of h/J is large. Nevertheless, even in the latter case, the dynamics computed with the VCNs with the largest coupling distances regarded here ($d = 9, 11$) follow

very closely the actual relaxation of $\langle \sigma_l^x \rangle$. As for the next-to-nearest neighbor correlation function $\langle \sigma_l^z \sigma_{l+2}^z \rangle$, it is found that correlations at this distance are rather small in the quench to the ferromagnetic phase, whereas they are larger and oscillate between positive and negative values before decaying to zero
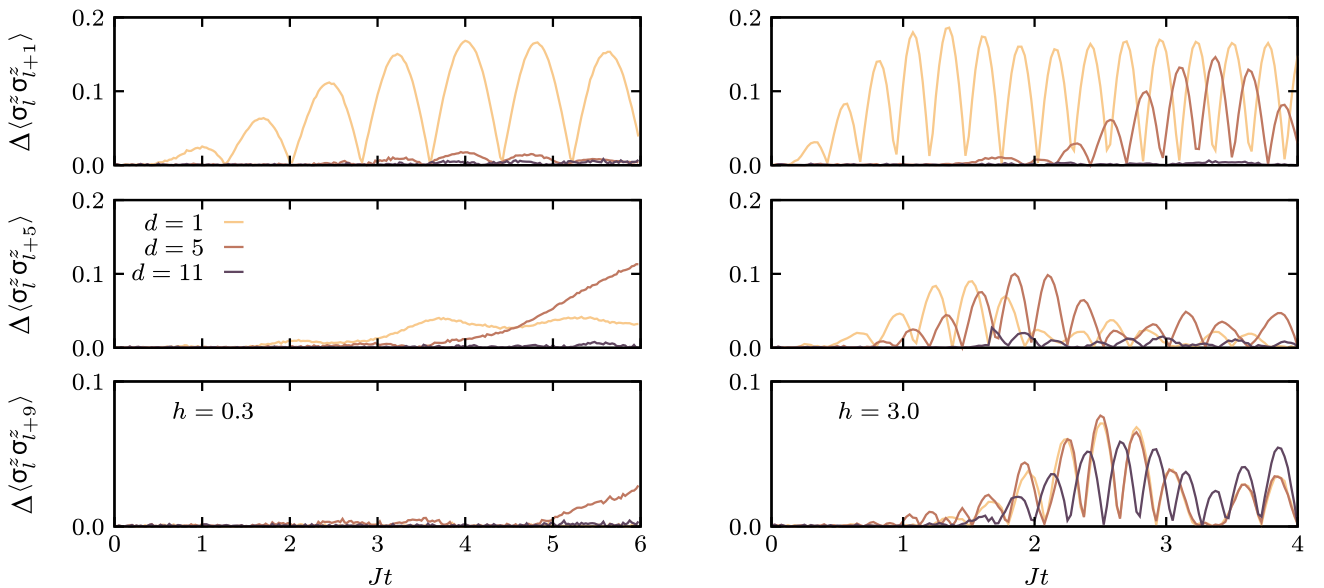

FIG. 5. Accuracy of correlations in the TFIM at different distances in the quench to $h/J = 0.3$ (left column) and $h/J = 3$ (right column): The quantity $\Delta \langle \sigma_l^z \sigma_{l+r}^z \rangle$ denotes the absolute value of the difference between the TDVP result and the exact result from free fermions.
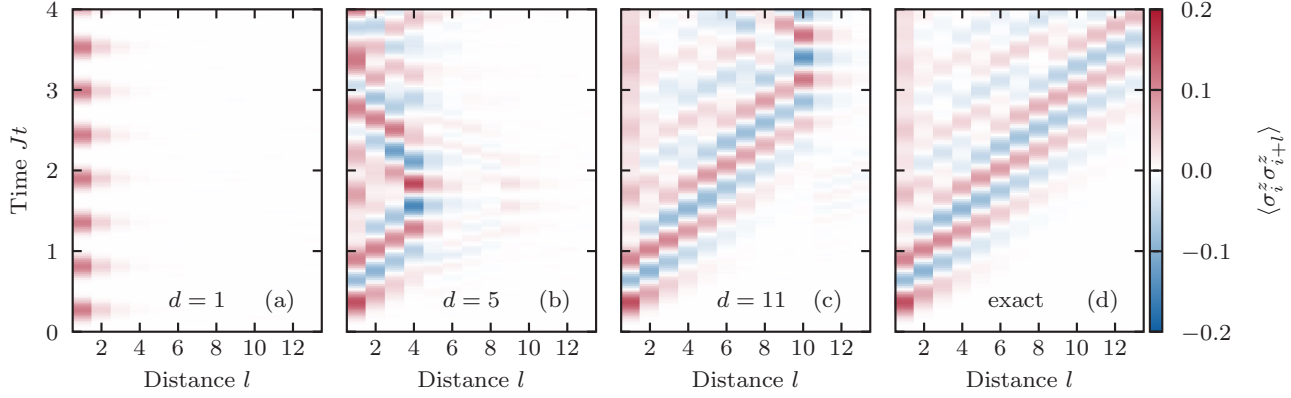
FIG. 6. Correlation spreading in the Ising model for the quench to $h/J = 3$, with VCNs with (a) $d = 1$, (b) $d = 5$, (c) $d = 11$: Spreading is only captured up to the coupling distance. The exact correlation dynamics is also shown in (d).

in the quench within the paramagnetic phase. In both cases, however, it is again the highest-order VCNs that yield a better description of this correlation function, as expected. In Fig. 4 results corresponding to a different system size ($N = 25$) are also shown for comparison in the quench to $h/J = 0.3$, with $d = 11$. There are no appreciable finite-size effects observed up to the accessed timescales.

The accuracy of correlations as a function of the coupling distance d is analyzed in Fig. 5, for the two quenches considered before. In this figure we plot the deviation of the TDVP results from the exact dynamics $\Delta \langle \sigma_l^z \sigma_{l+r}^z \rangle$, in terms of two-point correlation functions $\langle \sigma_l^z \sigma_{l+r}^z \rangle$, at various distances r. As a general remark, we observe that the deviations from the exact result are systematically decreased by increasing the coupling distance d, in agreement with the results in Fig. 4. Also, it should be noted that the smaller the coupling distance d is, the earlier the deviations from the exact dynamics occur, as expected from the underlying perturbatively motivated structure of the VCNs. Moreover, we observe that in the case of large transverse field the deviations, in general, grow more than in the quench to weak fields. This is due to the fact that in the dynamics with a large transverse field, correlations develop significantly at all the considered distances in the relevant timescales, see Fig. 6 below, whereas at weak transverse field the dynamics is more local and hence correlations at large distances are rather small (see, for instance, the correlation function on the left column of Fig. 4). Lastly, note that the oscillations observed for some of the deviations arise from the fact that the variational results oscillate around the exact solution, as can be seen in Fig. 4.

Finally, it is instructive to look at the correlation spreading when using VCNs with different coupling distance d. This is illustrated in Fig. 6 for the quench to $h/J = 3$, with three VCNs with $d = 1, 5, 11$, as well as the exact solution. The results shown on this figure reveal another crucial feature of VCNs: The distance for which a VCN can adequately capture the propagation of correlations is exactly determined by the coupling distance d. Thus, we see that the spreading of correlations can be well captured in a controlled manner by increasing the coupling distance in the structure of the VCN. Although this result is obtained for the Ising model, we expect that it holds in general.

IV. QUENCHES IN THE 2D $U(1)$ QLM

A. Model

We now show that the method presented in this paper can also be used to tackle more challenging problems such as the quantum dynamics of nonintegrable models in higher dimensions. Concretely, we study a 2D $U(1)$ quantum link model (QLM), where gauge fields are represented by spin-$1/2$ operators, defined on the links of a square lattice, and without matter degrees of freedom (see Fig. 7). The Hamiltonian reads

$$H_{\text{QLM}} = -\mathcal{J} \sum_\square (U_\square + U_\square^\dagger) + \lambda \sum_\square (U_\square + U_\square^\dagger)^2, \quad (36)$$

where the sum goes over elementary plaquettes $\square$, and $U_\square = S_{x,\hat{i}}^+ S_{x+\hat{i},\hat{j}}^+ S_{x+\hat{j},\hat{i}}^- S_{x,\hat{j}}^-$, where $S_{x,\hat{\mu}}^{+/-}$ are the standard raising/lowering spin-$1/2$ operators defined on the link that connects sites x and $x + \hat{\mu}$, with the unit vectors on the 2D lattice being denoted $\hat{i}, \hat{j}$. For simplicity, we use spin variables normalized to 1; i.e., the eigenvalues of $S_{x,\hat{\mu}}^z$ are $s_{x,\hat{\mu}} = \pm 1$. This model is relevant within the context of both lattice quantum electrodynamics [62] and quantum spin liquids [63]. Experimentally, there are several proposals for realizing
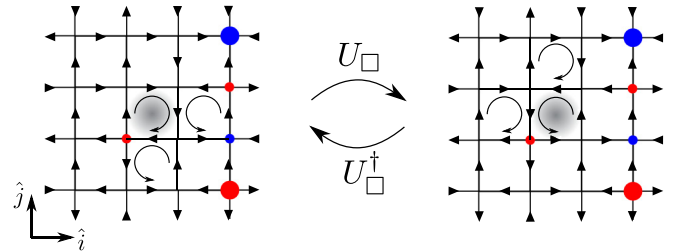

FIG. 7. The 2D $U(1)$ QLM. Quantum links and electric fluxes are represented by spin-$1/2$ operators. In this figure, electric fluxes (spins) are depicted as arrows with the following convention: Right/up (left/down) arrows correspond to $s_{x,\hat{\mu}} = +1$ ($s_{x,\hat{\mu}} = -1$), where $s_{x,\hat{\mu}}$ are the eigenvalues of $S_{x,\hat{\mu}}^z$. Quantum dynamics is induced by plaquette-flip operators $U_\square, U_\square^\dagger$, where their action is schematically illustrated on the shaded plaquettes. A non-flippable plaquette is annihilated by these operators. Red (blue) circles indicate sites with a positive (negative) charge Q_x.

lattice gauge theories in various quantum simulator settings [62,64–67] that could be relevant for the 2D $U(1)$ QLM studied here. In what follows, we also consider periodic boundary conditions.

In this theory, the associated electric flux is given by the z component of the spin-1/2 operator, i.e., $E_{x,\hat{\mu}} \equiv S^z_{x,\hat{\mu}}$. Thus, quantum links and electric field operators satisfy canonical commutation relations, i.e., $[S^z_{x,\hat{\mu}}, S^+_{x',\hat{\mu}'}] = S^+_{x,\hat{\mu}}\delta_{x,x'}\delta_{\mu,\mu'}$ and $[S^z_{x,\hat{\mu}}, S^-_{x',\hat{\mu}'}] = -S^-_{x,\hat{\mu}}\delta_{x,x'}\delta_{\mu,\mu'}$. Moreover, the Hamiltonian (36) is invariant under local $U(1)$ transformations with generator [62] $G(x) = \sum_\mu (E_{x,\hat{\mu}} - E_{x-\hat{\mu},\hat{\mu}})$, so that $[H_{\mathrm{QLM}}, G] = 0$. This allows us to structure the Hilbert space into so-called superselection sectors [49], with an associated background charge distribution $\{Q_\alpha\}$ satisfying Gauss law, i.e., $G(x)|\psi\rangle = Q_x|\psi\rangle$, for all x. Thus, every superselection sector consists of only physical states fulfilling that the incoming and outgoing fluxes equal the charge Q_x at a given site x, specified by the distribution $\{Q_\alpha\}$ defining the corresponding sector.

As illustrated in Fig. 7, the dynamics in the QLM is generated by the first term on the right-hand side of Eq. (36), where the plaquette operators $U_\square$, $U^\dagger_\square$ induce tunneling processes between configurations with *flippable* plaquettes, in which the electric fluxes form a loop with clockwise or anticlockwise orientation (a non-flippable plaquette is annihilated by such operators). The second contribution on the right-hand side of Eq. (36) acts as a potential energy term that favors those configurations with a larger number of flippable plaquettes for $\lambda < 0$. When $\lambda = -\infty$, the Hamiltonian in Eq. (36) has two $\mathbb{Z}_2$-symmetry-broken ground states, with the electric fluxes arranged such that all the plaquettes in the system are flippable; i.e., the four spins in any plaquette form a closed loop with either clockwise or anticlockwise orientation. These states, together with all other states kinetically connected to them, define a sector that we shall refer to as the *fully flippable* (FF) sector. We note that the FF sector is a subset of the superselection sector with zero charge: $G(x)|\psi\rangle = 0 \; \forall \; x$.

The out-of-equilibrium dynamics of the model studied in this section has recently gained increased attention. For instance, it has been found that DQPTs occur during the unitary dynamics that follows a sudden perturbation in the Hamiltonian in Eq. (36), when starting from one of the ground states in the FF sector [68]. Further, in a recent paper [46], we have also shown that the model under consideration may undergo a localization-to-ergodic transition, constituting an example of a genuinely interacting 2D theory featuring disorder-free localization [47–51]. Indeed, it is found that quenches from the initial state $|\rightarrow\rangle$ to a QLM with $\mathcal{J}/\lambda < 0$ give rise to localized behavior, such as a significant suppression of transport and a limited spread of correlations, whose origin is linked to a *fragmentation* [69–72] of the Hilbert space into kinetically disconnected regions due to hard local constraints imposed by gauge invariance. Let us remark that in the present context, the initial state $|\rightarrow\rangle$ in Eq. (15) corresponds to an equally weighted superposition of all electric flux (spin) configurations. On the other hand, the dynamics within the FF sector, for example, displays ergodic behavior, characterized by a propagation of correlations and transport quantities throughout the whole system. Subsequently, we shall consider
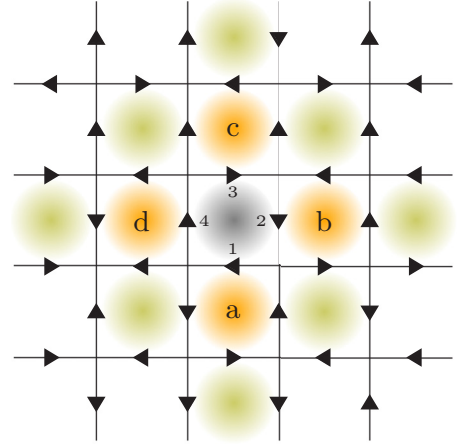

FIG. 8. Convention for the definition of the operators in Eqs. (39)–(42). The projectors in these equations are defined on the bonds of the neighboring plaquettes (orange), except for those that are shared with the reference plaquette (gray). Labeling of plaquettes (a, b, c, d) and of spins within a plaquette (1,2,3,4) is done according to their relative position as shown in this figure. Plaquettes in orange are relevant for first-order cumulants, whereas second-order cumulants involve also the plaquettes in green.

precisely the two scenarios mentioned above. However, let us first introduce the VCNs for the QLM.

B. VCNs for the 2D QLM

Using the general scheme introduced in Sec. II B, we derive pCNs and the corresponding VCNs for the QLM. The classical limit of the theory under consideration is $H^{\mathrm{QLM}}_0 = \lambda \sum_\square (U_\square + U^\dagger_\square)^2$. Quantum fluctuations are then induced by the term $\gamma V^{\mathrm{QLM}} = -\mathcal{J}\sum_\square (U_\square + U^\dagger_\square)$, where we recognize $\gamma = -\mathcal{J}$. Next, we solve the equation of motion for the operator $U_\square$, that is, $-i\frac{d}{dt}U_\square(t) = [H_0, U_\square]$ (and likewise for $U^\dagger_\square$). We get

$$U_\square(t) = e^{i\lambda\Omega_\square t}U_\square, \quad U^\dagger_\square(t) = e^{-i\lambda\Omega_\square t}U^\dagger_\square, \quad (37)$$

where the operator $\Omega_\square$ commutes with both $U_\square$ and $U^\dagger_\square$, and is given by

$$\Omega_\square = \sum_{p\in\mathcal{P}_\square} (-\mathcal{A}_p + \mathcal{B}_p), \quad (38)$$

where $\mathcal{P}_\square = \{a, b, c, d\}$ denotes the set of neighboring plaquettes around a given plaquette (see Fig. 8), and the operators $\mathcal{A}_p$ and $\mathcal{B}_p$ are given by

$$\mathcal{A}_a = P^\uparrow_{a,1}P^\uparrow_{a,2}P^\downarrow_{a,4}, \quad \mathcal{B}_a = P^\downarrow_{a,1}P^\downarrow_{a,2}P^\uparrow_{a,4}, \quad (39)$$

$$\mathcal{A}_b = P^\uparrow_{b,1}P^\uparrow_{b,2}P^\downarrow_{b,3}, \quad \mathcal{B}_b = P^\downarrow_{b,1}P^\downarrow_{b,2}P^\uparrow_{b,3}, \quad (40)$$

$$\mathcal{A}_c = P^\uparrow_{c,2}P^\downarrow_{c,3}P^\downarrow_{c,4}, \quad \mathcal{B}_c = P^\downarrow_{c,2}P^\uparrow_{c,3}P^\uparrow_{c,4}, \quad (41)$$

$$\mathcal{A}_d = P^\uparrow_{d,1}P^\downarrow_{d,3}P^\downarrow_{d,4}, \quad \mathcal{B}_d = P^\downarrow_{d,1}P^\uparrow_{d,3}P^\uparrow_{d,4}. \quad (42)$$

In these definitions, the operators $P^{\uparrow,\downarrow}_{p,i}$ are projectors onto one of the components of the spins defined on the bonds of the four neighboring plaquettes; e.g., $P^\uparrow_{a,2} = S^+_2 S^-_2 = (1 + S^z_2)/2$

projects onto the $s_2 = +1$ component of the spin living on the second link of plaquette a, which lies underneath the reference plaquette as in Fig. 8. Note that the very same spin of this example is also the fourth one in the plaquette on the right side of a. The convention for enumerating the spins in a given plaquette and labeling neighboring plaquettes is shown in Fig. 8.

1. First-order ansatz

Using λ as the unit of energy, the first-order correction of the cumulant expansion, Eq. (21), gives

$$\mathcal{H}^{(1)}_{\text{pCN–QLM}} = i\frac{\mathcal{J}}{\lambda} \sum_{\square} \left(\frac{\psi_0(s_\square)}{\psi_0(s)} \mathcal{F}_\square(s)^2 \int_0^t dt' e^{i\lambda\omega_\square(s)t'} \right),$$

$$(43)$$

where $\omega_\square(s) := \mathcal{F}_\square(s)\Omega_\square(s)$, and $\mathcal{F}_\square := U_\square U_\square^\dagger - U_\square^\dagger U_\square$ is a diagonal operator with possible matrix elements $\mathcal{F}_\square(s) = +1, -1, 0$, when the reference plaquette (indicated by the subscript $\square$) is flippable and has an anticlockwise orientation $(+1)$, a clockwise orientation (-1), or it is not flippable (0). Also, we use the notation $|s_\square\rangle \equiv (U_\square + U_\square^\dagger)|s\rangle$; i.e., $s_\square$ differs from s by the flipping of a single plaquette. $\Omega_\square(s)$ denotes the diagonal entries of the operator introduced in Eq. (38).

Following the prescription given by Eqs. (23) and (24), we can introduce a variational parameter for each one of the nine possible values that the integer variable $\omega_\square(s) \in [-4, 4]$ can take. Thus, using the short-hand notation

$$\eta^{(1)}_{\omega_\square(s)}(t) := \sum_{\omega=-4}^{4} \delta_{\omega_\square(s),\omega} \eta^{(1)}_\omega(t),$$

$$(44)$$

we can write down the corresponding variational effective Hamiltonian as

$$\tilde{\mathcal{H}}^{(1)}_{\text{QLM}} = i\frac{\mathcal{J}}{\lambda} \sum_{\square} \left(\frac{\psi_0(s_\square)}{\psi_0(s)} \times [\mathcal{F}_\square(s)]^2 \eta^{(1)}_{\omega_\square(s)}(t) \right), \quad (45)$$

which contains up to nine independent (assuming translational invariance) "first-order" variational parameters $\eta^{(1)}_\omega(t)$. On the other hand, the "zeroth-order" contribution reads

$$\tilde{\mathcal{H}}^{(0)}_{\text{QLM}} = \ln[\psi_0(s)] - iE_s^* \eta^{(0)}(t), \quad (46)$$

with a single variational parameter $\eta^{(0)}(t)$ and $E_s^* \equiv E_s/\lambda$ being dimensionless. Thus, a VCN built upon the zeroth- and first-order cumulants reads

$$\mathcal{H}^{(1)}_{\text{QLM}} = \tilde{\mathcal{H}}^{(0)}_{\text{QLM}} + \tilde{\mathcal{H}}^{(1)}_{\text{QLM}}. \quad (47)$$

The first-order VCN defined above contains a maximum coupling distance (along either the $\hat{i}$ or $\hat{j}$ direction) between plaquettes equal to 2. Indeed, for a given plaquette $\square$, the function $\omega_\square(s)$ involves only the nearest neighboring plaquettes that are shown in Fig. 8. Thus, when plugged into the expression of the classical network, this gives rise to terms where the plaquettes that are separated the most are, for example, b and d in Fig. 8. In terms of parallel spins [e.g., spins on links $(x, x + \hat{j})$ and $(x + l \cdot \hat{i}, x + \hat{j})$, with l an integer], the maximal coupling distance is equal to 3.

Finally, let us note that the ansatz (47) can be explicitly recast as a classical Ising-like spin model with multiple (up to 16) spin interaction terms. Indeed, this can be achieved by rewriting the constraints in Eq. (44) in terms of the projectors $P^{\uparrow,\downarrow}_{p,i}$. For instance, for $\omega = 4$, we have

$$\delta_{\omega_\square(s),\omega=4} = P^\downarrow_{\square,1} P^\downarrow_{\square,2} P^\uparrow_{\square,3} P^\uparrow_{\square,4} P^\uparrow_{a,1} P^\uparrow_{a,2} P^\downarrow_{a,4} P^\uparrow_{b,1} P^\uparrow_{b,2} P^\downarrow_{b,3} P^\uparrow_{c,2} P^\downarrow_{c,3} P^\downarrow_{c,4} P^\uparrow_{d,1} P^\downarrow_{d,3} P^\downarrow_{d,4}. \quad (48)$$

2. Second-order ansatz

For most of the numerical results we shall present next, it becomes important to make use of a second-order VCN. Thus, let us show explicitly how to construct such ansatz. The second-order cumulant contains two terms as indicated in Eq. (13). Using the solutions in Eq. (37), we have

$$\frac{\langle s|V(t')V(t'')|\psi_0\rangle}{\langle s|\psi_0\rangle} = \sum_{\square',\square''} \frac{\psi_0(s_{\square',\square''})}{\psi_0(s)} [\mathcal{F}_{\square'}(s)]^2 [\mathcal{F}_{\square''}(s_{\square'})]^2 e^{i\lambda\omega_{\square'}(s)t'} e^{i\lambda\omega_{\square''}(s_{\square'})t''}, \quad (49)$$

$$\frac{\langle s|V(t')|\psi_0\rangle\langle s|V(t'')|\psi_0\rangle}{\langle s|\psi_0\rangle^2} = \sum_{\square',\square''} \frac{\psi_0(s_{\square'})\psi_0(s_{\square''})}{\psi_0(s)^2} [\mathcal{F}_{\square'}(s)]^2 [\mathcal{F}_{\square''}(s)]^2 e^{i\lambda\omega_{\square'}(s)t'} e^{i\lambda\omega_{\square''}(s)t''}, \quad (50)$$

where $|s_{\square',\square''}\rangle \equiv (U_{\square'} + U_{\square'}^\dagger)(U_{\square''} + U_{\square''}^\dagger)|s\rangle$. These equations then give the needed corrections to form a second-order pCN, and will be the basis to build the corresponding second-order VCN. Let us note at this point the following aspect concerning the locality of the second-order cumulant. Each of the two contributions, Eqs. (49) and (50), might, in principle, give rise to couplings at all distances. However, when we subtract them to form the overall second-order correction [see Eq. (13)], most of the resulting terms cancel out, leaving only couplings up to some (local) coupling distance d. For instance, if we consider the initial state $|\rightarrow\rangle$

in Eq. (15), the ratios of initial amplitudes in the equations above reduce to 1. Then, one can easily verify that the only nonvanishing contributions arise from overlapping plaquettes, i.e., plaquettes sharing one common link (gray and orange plaquettes in Fig. 8) or from plaquettes connected by a common neighboring plaquette (gray and green plaquettes in Fig. 8). In effect, for plaquettes separated by more than one intermediate plaquette (i.e., plaquettes outside the colored region in Fig. 8), we have that $\mathcal{F}_{\square''}(s_{\square'}) \equiv \mathcal{F}_{\square''}(s)$ and $\omega_{\square''}(s_{\square'}) \equiv \omega_{\square''}(s)$, and hence, Eqs. (49) and (50) become identical. In this case, the coupling distance (along either the $\hat{i}$ or $\hat{j}$ direction) is $d = 4$

for plaquettes, and $d = 5$ for parallel spins. We note that this remark also holds for other initial states that can be written as a product of local terms, as long as supports of sufficiently separated local terms do not overlap with each other, as is the case for all initial states considered in this work.

$$\int_0^t dt' \int_0^{t'} dt'' e^{i\lambda\omega_{\square'}(s)t'} e^{i\lambda\omega_{\square''}(s_{\square'})t''} = \sum_{\omega_1,\omega_2=-4}^{4} \delta_{\omega_{\square'}(s),\omega_1}\delta_{\omega_{\square''}(s_{\square'}),\omega_2} \int_0^t dt' \int_0^{t'} dt'' e^{i\lambda\omega_1 t'} e^{i\lambda\omega_2 t''}, \tag{51}$$

for Eq. (49), and likewise for Eq. (50). In this case, there are 81 possible combinations of the tuple (ω_1, ω_2); hence, we introduce 81 second-order variational parameters $\eta^{(2)}_{\omega_1,\omega_2}(t)$. Thus, using the short-hand notation

$$\eta^{(2)}_{\omega_{\square'}(s),\omega_{\square''}(s_{\square'})}(t) := \sum_{\omega_1,\omega_2=-4}^{4} \delta_{\omega_{\square'}(s),\omega_1}\delta_{\omega_{\square''}(s_{\square'}),\omega_2}\eta^{(2)}_{\omega_1,\omega_2}(t), \tag{52}$$

and upon integrating Eq. (49), one can make the substitution

$$\int_0^t dt' \int_0^{t'} dt'' \frac{\langle s|V(t')V(t'')|\psi_0\rangle}{\langle s|\psi_0\rangle} \longleftarrow \sum_{\square',\square''} \frac{\psi_0(s_{\square',\square''})}{\psi_0(s)} \mathcal{F}_{\square'}(s)^2 \mathcal{F}_{\square''}(s_{\square'})^2 \eta^{(2)}_{\omega_{\square'}(s)\omega_{\square''}(s_{\square'})}(t). \tag{53}$$

Analogously, the corresponding substitution of Eq. (50) is

$$\int_0^t dt' \int_0^{t'} dt'' \frac{\langle s|V(t')|\psi_0\rangle\langle s|V(t'')|\psi_0\rangle}{\langle s|\psi_0\rangle^2} \longleftarrow \sum_{\square',\square''} \frac{\psi_0(s_{\square'})\psi_0(s_{\square''})}{\psi_0(s)^2} \mathcal{F}_{\square'}(s)^2 \mathcal{F}_{\square''}(s)^2 \tilde{\eta}^{(2)}_{\omega_{\square'}(s),\omega_{\square''}(s)}(t), \tag{54}$$

with $\tilde{\eta}^{(2)}_{\omega_{\square'}(s),\omega_{\square''}(s)}(t)$ defining a further 81 variational parameters. Combining the previous two equations, we obtain the full, variational second-order correction:

$$\tilde{\mathcal{H}}^{(2)}_{\text{QLM}} = -\frac{\mathcal{J}^2}{\lambda^2} \tilde{\sum}_{\square',\square''} \left(\frac{\psi_0(s_{\square',\square''})}{\psi_0(s)} \eta^{(2)}_{\omega_{\square'}(s),\omega_{\square''}(s_{\square'})}(t) \right) + \frac{\mathcal{J}^2}{\lambda^2} \tilde{\tilde{\sum}}_{\square',\square''} \left(\frac{\psi_0(s_{\square'})\psi_0(s_{\square''})}{\psi_0(s)^2} \tilde{\eta}^{(2)}_{\omega_{\square'}(s),\omega_{\square''}(s)}(t) \right), \tag{55}$$

where $\tilde{\sum}_{\square',\square''} := \sum_{\square',\square''}[\mathcal{F}_{\square'}(s)]^2[\mathcal{F}_{\square''}(s_{\square'})]^2$ and $\tilde{\tilde{\sum}}_{\square',\square''} := \sum_{\square',\square''}[\mathcal{F}_{\square'}(s)]^2[\mathcal{F}_{\square''}(s)]^2$. Thus, a VCN including up to second-order cumulants reads

$$\mathcal{H}^{(2)}_{\text{QLM}} = \tilde{\mathcal{H}}^{(0)}_{\text{QLM}} + \tilde{\mathcal{H}}^{(1)}_{\text{QLM}} + \tilde{\mathcal{H}}^{(2)}_{\text{QLM}}, \tag{56}$$

with the expressions given in Eqs. (46), (45), and (55).

Finally, let us point out that we restrict the resulting VCNs to have the same locality as the underlying pCNs; i.e., the sums over plaquette $\square''$ in (55) are restricted to the nearest or next-nearest neighbors of plaquette $\square'$ (see Fig. 8).

C. Quench protocol and results

In order to test the performance of the VCNs defined in Eqs. (47) and (56), we investigate several quantum quenches in the 2D QLM. First, we benchmark our method by considering a setting where it is still possible to perform exact diagonalization (ED) for relatively large system sizes. This is achieved by regarding a quasi-1D ladder with dynamics restricted to a given superselection sector. In this type of setting, we study quenches from both uniform and nonuniform initial states. Next, we consider situations well beyond the scope of ED. Namely, we regard a truly two-dimensional configuration, and quenches from the initial state in Eq. (15), which involves all the superselection sectors. As explained below, for the particular physical scenario considered in this case (disorder-free localization), our method is able to yield sufficiently accurate results for long timescales that could hardly be accessed with any other state-of-the-art numerical technique. Unless otherwise stated, in all the examples discussed in this section, the TDVP equations are solved using a 4th-order Runge-Kutta integrator with step size $\Delta t = 0.1\lambda^{-1}$.

Now let us convert the second-order pCN into a second-order VCN. In the same spirit as for the first-order ansatz, the integrals involved in the second-order correction can be written as

1. Benchmark in quasi-1D ladders: Uniform initial states

As a first benchmark, we consider a quasi-1D ladder of 2×10 plaquettes, i.e., 40 spins, with dynamics restricted to a given superselection sector. Concretely, we consider the FF sector, consisting of all the states with zero charge that are kinetically connected to the two maximally flippable configurations, as explained in Sec. IV A. As shown below, with these settings, one can still carry out efficient ED calculations. Indeed, the key idea is that, since the Hamiltonian in Eq. (36) adopts a block-diagonal form, with each block corresponding to a given superselection sector, a state that belongs to one of such sectors will remain in that sector during the course of the dynamics generated by H_{QLM}. Thus, one needs to consider only the portion of the Hilbert space that is relevant for the chosen sector.

The quench protocol considered here is the following: We initialize the system in an equally weighted superposition of all the basis states spanning the FF sector. We denote such a superposition as $|\rightarrow\rangle_{\text{FF}}$. Next, we evolve the system with the
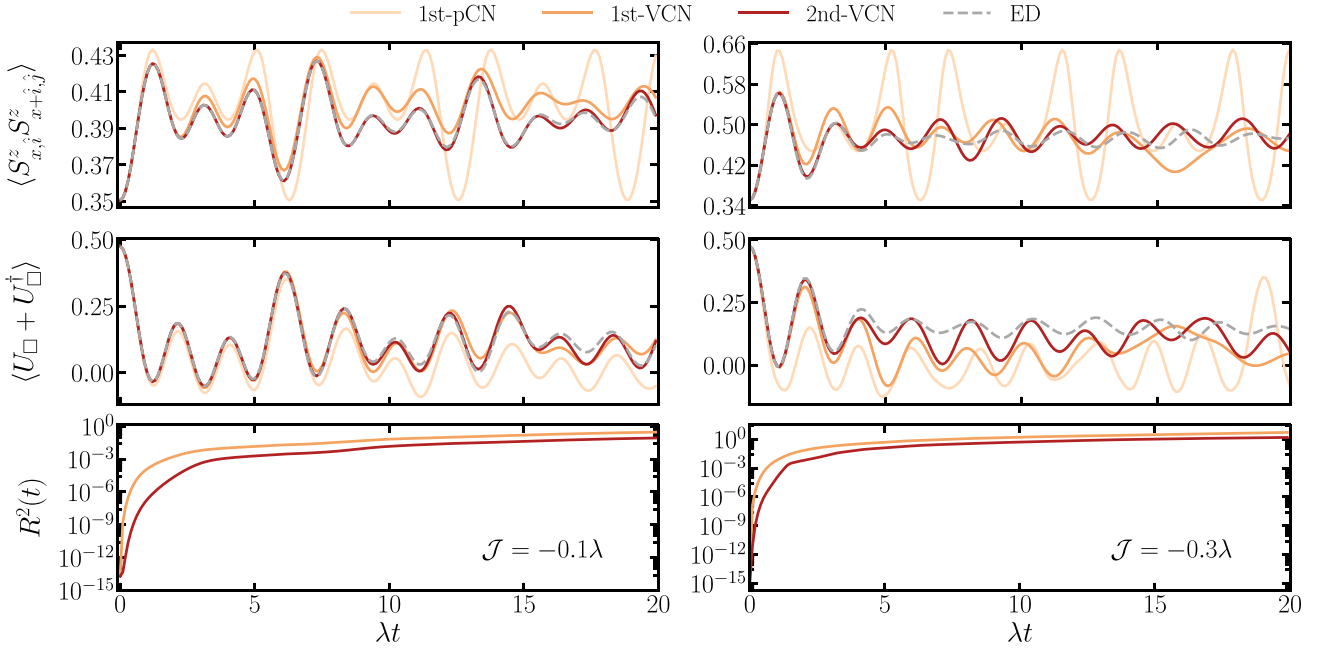
FIG. 9. Comparison of perturbative, variational, and exact dynamics in a quasi-1D ladder of 2×10 plaquettes (40 spins): Left and right columns show the quench dynamics from $| \rightarrow \rangle_{FF}$ to $\mathcal{J}/\lambda = -0.1$ and $\mathcal{J}/\lambda = -0.3$, respectively. Upper row: Dynamics of the nearest-neighbor spin-spin correlator $\langle S^z_{x,i} S^z_{x+\hat{i},\hat{j}} \rangle$. Middle row: Dynamics of the (dimensionless) mean kinetic energy per plaquette $\langle U_\square + U_\square^\dagger \rangle$. Lower row: Integrated residuals $R^2(t)$, of the first- and second-order VCNs.

Hamiltonian in Eq. (36), with a finite value of $\mathcal{J}/\lambda$. Later, we will study similar quenches but starting from nonuniform initial states. Note that for a 2×10 system, the dimension of the FF sector is 17 906. Therefore, in the TDVP calculations, we can also carry out an exact enumeration of states; i.e., we can explicitly sum over all relevant spin configurations in expressions such as the expectation value in Eq. (3), rather than performing a Monte Carlo sampling.

Let us emphasize that the dynamics in the FF sector is interesting because significant correlations may develop in the entire spatial extent of the system, as we have recently shown [46]. Therefore, this constitutes a highly challenging scenario for our method, since a VCN can adequately capture the buildup of correlations up to a distance specified, essentially, by the order of the cumulant expansion, as we have already argued.

The results of this benchmark are shown in Fig. 9, for quenches to $\mathcal{J}/\lambda = -0.1$ and $\mathcal{J}/\lambda = -0.3$, left and right columns of Fig. 9, respectively. We compare the dynamics computed with a first-order pCN, a first-order VCN [Eq. (47)], a second-order VCN [Eq. (56)], and via ED. Particularly, we focus on two observables: the nearest-neighbor spin-spin correlation function $\langle S^z_{x,i} S^z_{x+\hat{i},\hat{j}} \rangle$ (i.e., $\langle S^z_1 S^z_2 \rangle$, with the convention in Fig. 8), and the mean kinetic energy (up to a factor of $-\mathcal{J}$) per plaquette $\langle U_\square + U_\square^\dagger \rangle$. Besides, we also show the integrated relative residuals $R^2(t)$ of the first- and second-order VCNs, as a measure of their accuracy.

As expected, it is the second-order VCN that provides the most accurate results in the studied quenches. This can be seen from both the evolution of the observables and the growth of the residuals. In particular, for the quench to $\mathcal{J}/\lambda = -0.1$, the second-order VCN captures remarkably well correlations at short distances (left column, upper row), as well as the

oscillatory behavior of off-diagonal observables (left column, middle row), in the entire range of accessed timescales $\lambda t = 20$. As anticipated, in the quench with a bigger strength of the perturbation $\mathcal{J}/\lambda = -0.3$, the description in terms of all of the considered wave functions breaks down at earlier times. Yet, the second-order VCN still yields a rather excellent agreement with ED up to significant timescales $\lambda t \sim 7$. In all cases, the VCNs outperform the first-order pCN regarded here.

Next, we point out that, as illustrated better in the second quench (right panel of Fig. 9), the oscillations of both $\langle S^z_{x,i} S^z_{x+\hat{i},\hat{j}} \rangle$ and $\langle U_\square + U_\square^\dagger \rangle$ *decay* to some steady-state value in agreement with the expectation of ergodic behavior in the FF sector [46]. Importantly, such decay of the oscillations is captured by the TDVP solutions. On the contrary, the pCN fails to capture this crucial feature, although it gets right the frequency of the oscillations up to considerable timescales. Finally, let us emphasize that the results displayed in Fig. 9 corroborate once more the *controlled* nature of VCNs, in that the accuracy of a VCN can be improved systematically either by adding higher-order terms of the cumulant expansion or by decreasing the strength of quantum fluctuations.

2. Benchmark in quasi-1D ladders: Nonuniform initial states

Here, we consider the same setting as before, namely, a 2×10 quasi-1D ladder, with dynamics restricted to the FF sector. The chosen initial state, however, is nonuniform. In particular, the initial condition is created by adding a line defect with subextensive energy contribution to the state $| \rightarrow \rangle_{FF}$. This is achieved by applying the operator

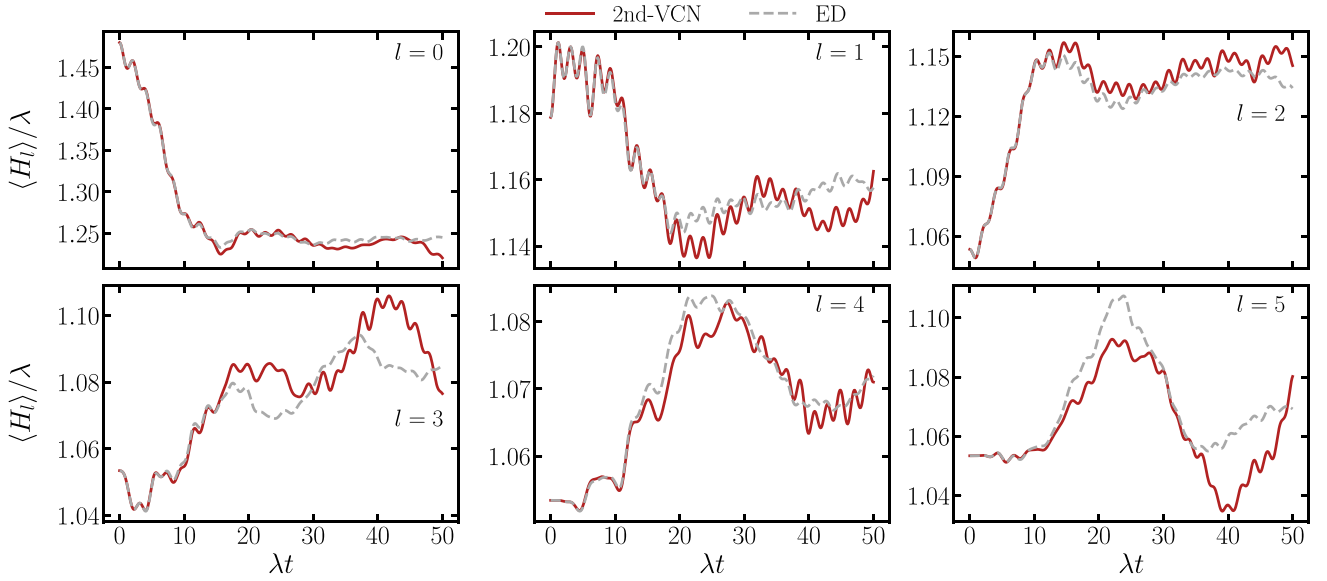$$\mathcal{P} = \prod_{\square \in \mathfrak{c}_0} 1 + (U_\square + U_\square^\dagger)^2 \qquad (57)$$

FIG. 10. Energy of the lth column with $l = 0, \dots, 5$, for the quench from the nonuniform initial state $\mathcal{P}|\rightarrow\rangle_{FF}$, containing an excess of energy around $l = 0$ (see main text) to $\mathcal{J}/\lambda = -0.1$. We compare the TDVP solution using a second-order VCN [Eq. (56)] with ED results.

upon the state $|\rightarrow\rangle_{FF}$, where $\mathcal{C}_d$ denotes the set of plaquettes in the lth column. Thus, $\mathcal{P}|\rightarrow\rangle_{FF}$ represents a state with a line energy defect along column $l = 0$. The resulting quench dynamics obtained using both a second-order VCN and ED are analyzed in terms of the evolution of the total energy of column l, which is given by

$$H_l = \sum_{\square \in \mathcal{C}_l} -\mathcal{J}(U_\square + U_\square^\dagger) + \lambda(U_\square + U_\square^\dagger)^2. \quad (58)$$

The results are shown in Fig. 10. One can observe that the second-order VCN captures quite well, in a quantitative way, the propagation of the line defect at all distances up to a time $\lambda t \approx 20$. After this point, the TDVP solution is not exact anymore, but it still follows qualitatively the exact dynamics to the largest accessed time, $\lambda t = 50$. The quantity in Eq. (58) can serve as a probe to distinguish between ergodic and nonergodic behavior in the QLM as done in Ref. [46].

3. Quenches in a 2D lattice

After having assessed the performance of our method in quasi-1D ladders, we now consider a truly 2D setting. Moreover, we shall regard regimes where access to ED is computationally prohibited. Concretely, the following quench protocol is carried out: The system is initialized in the state $|\rightarrow\rangle$ given in Eq. (15) and then evolved with the Hamiltonian in Eq. (36) at finite $\mathcal{J}/\lambda$. As mentioned before, in the present context the state $|\rightarrow\rangle$ can be thought of as an equally weighted superposition of all the electric flux (spin) configurations, thereby involving all the superselection sectors. Thus, apart from very small system sizes, ED techniques become impractical. In the following, we show results for a system of 100×100 plaquettes (2×10^4 spins). In all the calculations considered here, a Metropolis Monte Carlo sampling was performed with 10^6 sweeps at each time instance, and single-spin-flip updates.

As explained in our recent work [46], it is interesting to investigate quenches from the initial state $|\rightarrow\rangle$ to the QLM in Eq. (36), as this scenario gives rise to (disorder-free) many-body localized dynamics, even though both the Hamiltonian and the initial condition are homogeneous. In effect, when quenching from the initial state $|\rightarrow\rangle$, not only the transport of energy is highly suppressed, but also the spreading of correlations is drastically constrained [46]. This means that relevant correlations can only develop at very short distances. Due to this reason, we expect low-order VCNs to capture very well the main features of the exact quantum dynamics in the present case.

The resulting dynamics is shown in Figs. 11 and 12, where quenches to the points $\mathcal{J}/\lambda = -0.1$ and $\mathcal{J}/\lambda = -0.3$ are considered, respectively. Let us focus first on the quench to $\mathcal{J}/\lambda = -0.1$ (Fig. 11). The left column shows the spatiotemporal buildup of quantum correlations in terms of two correlation functions, namely, the spin-spin correlation function $\langle S_0^z(t) S_l^z(t)\rangle \equiv \langle S_{x,\hat{j}}^z(t) S_{x+l\cdot\hat{i},\hat{j}}^z(t)\rangle$, between parallel spins separated by a distance l, and the connected two-point plaquette-flip correlation function $\langle (U_0 + U_0^\dagger)(U_l + U_l^\dagger)\rangle_C \equiv \langle (U_0 + U_0^\dagger)(U_l + U_l^\dagger)\rangle - \langle (U_0 + U_0^\dagger)\rangle^2$. In both cases, we compare the results obtained using a first-order VCN (left-hand side of the color map), and a second-order VCN (right-hand side of the color map). As a first observation, we note the rather confined spreading of correlations. As already mentioned, this is due to the disorder-free localization mechanism that takes place in the situation considered here. We stress that this is not an artifact of the VCNs. Indeed, a first-order VCN, see Eq. (47), has a coupling distance $d = 2$, for plaquettes, and $d = 3$, for parallel spins. On the other hand, a second-order VCN, see Eq. (56), has coupling distances $d = 4, 5$, for plaquettes and parallel spins, respectively. Therefore, according to our previous analysis (see Fig. 6, in particular), we expect that both of these VCNs would be able to account for the spreading of correlations, at least up to the
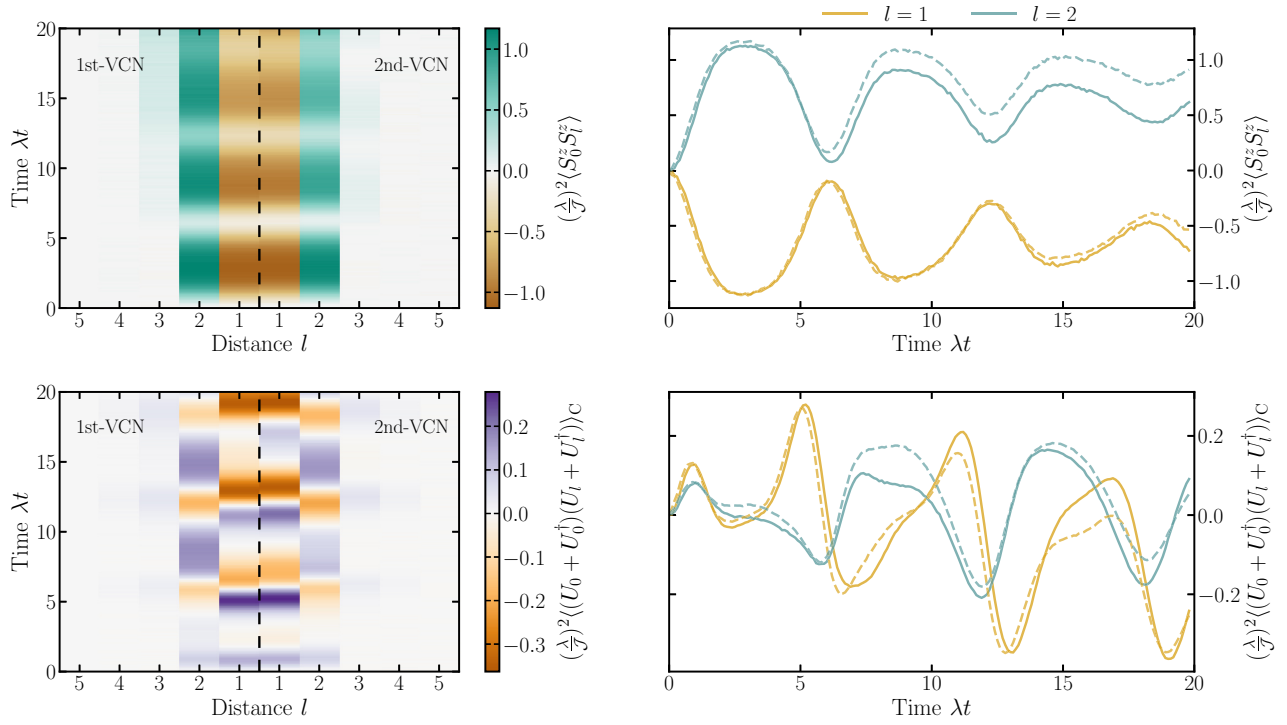
FIG. 11. Correlation dynamics in the 2D QLM for a quench from $|\rightarrow\rangle$ to $\mathcal{J}/\lambda = -0.1$. Left column: Spatiotemporal buildup of quantum correlations computed with a first-order VCN (left side), and a second-order VCN (right side). Upper panel: Spin-spin correlation function $\langle S_0^z S_l^z \rangle \equiv \langle S_{x,\hat{j}}^z S_{x+l\cdot\hat{i},\hat{j}}^z \rangle$, between parallel links separated by a distance l; lower panel: plaquette-plaquette connected correlation function $\langle (U_0 + U_0^\dagger)(U_l + U_l^\dagger) \rangle_C$ (see main text). Right column: Individual cuts along the distances with the dominant signal, $l = 1, 2$, of the corresponding quantities on the left column; second-order VCN (solid), first-order VCN (dashed). Results for a 100×100 system (2×10^4 spins).

aforementioned distances. However, we find that significant correlations are developed predominantly no further than a distance $l = 3$, at least up to the accessed timescales.

A second observation is that the results obtained with both VCNs exhibit a very similar qualitative behavior. On the right column of Fig. 11, we make a more quantitative comparison by plotting individual cuts along the distances with the largest signal, i.e., $l = 1, 2$. We notice that the oscillations in both cases have pretty much the same frequency, whereas the

observed discrepancies are potentially due to the fact that the first-order ansatz cannot quite capture the decay of oscillations, as also seen in Fig. 9 for the quasi-1D ladder.

Finally, in Fig. 12, we show the dynamics of correlations for the quench to $\mathcal{J}/\lambda = -0.3$. We note that, apart from the fact that the dynamics gets accelerated due to the bigger strength of the off-diagonal perturbation, remarks similar to the ones for the previous quench also hold in this case. In particular, we see that the correlation spreading is constrained,
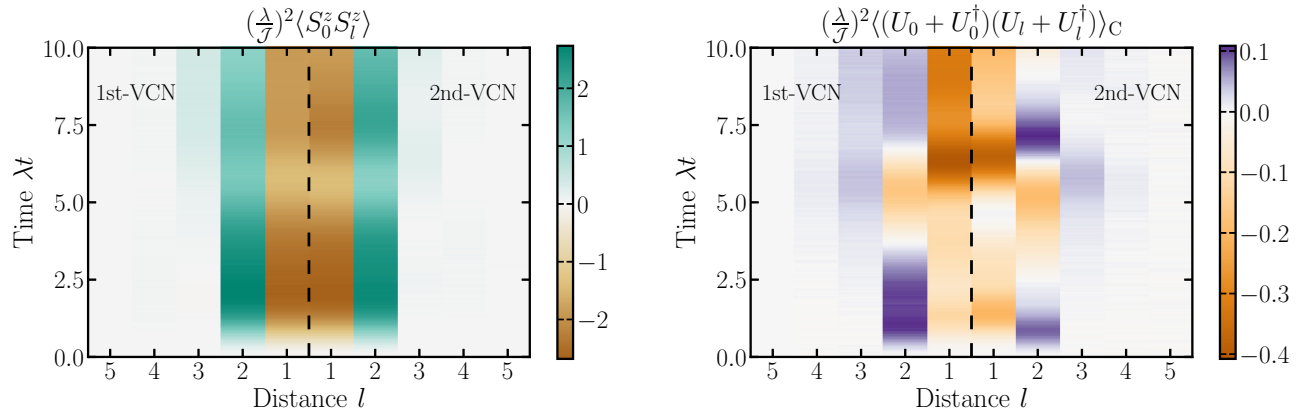

FIG. 12. Correlation dynamics in the 2D QLM for a quench from $|\rightarrow\rangle$ to $\mathcal{J}/\lambda = -0.3$. Left panel: Spin-spin correlation function $\langle S_0^z(t) S_l^z(t) \rangle \equiv \langle S_{x,\hat{j}}^z(t) S_{x+l\cdot\hat{i},\hat{j}}^z(t) \rangle$, between parallel links separated by a distance l. Right panel: Plaquette-plaquette connected correlation function $\langle (U_0 + U_0^\dagger)(U_l + U_l^\dagger) \rangle_C$ (see main text). Results obtained with a first-order VCN (left side), and a second-order VCN (right side), for a 100×100 system (2×10^4 spins). Here, a time step $\Delta t = 0.05\lambda^{-1}$ was used to integrate the TDVP equations.

too, with significant correlations growing only in the spatial region $l \leqslant 4$. Once more, this is a manifestation of the disorder-free localization phenomenon that occurs in the dynamics of the 2D QLM for the type of quenches studied in this section.

V. SUMMARY AND OUTLOOK

We have introduced a numerical variational scheme for the study of dynamics in correlated quantum systems in one and higher dimensions. Our method relies on an efficient representation of the many-body wave function, in terms of complex networks of classical spin variables. This class of variational wave functions, termed VCNs, is similar to ANNs. Crucially, VCNs can be constructed according to a controlled prescription, as it has been underlined and explicitly shown in this paper. Going beyond Ref. [46] where VCNs were introduced only for the particular case of the 2D QLM, here we presented VCNs as a general numerical framework that can be applied to any interacting quantum lattice model, regardless of spatial dimensionality, provided that an expansion around a well-defined classical limit is possible.

We have illustrated the way in which our method works and its range of applicability by studying quantum quenches in two very distinct models. First, we considered the paradigmatic 1D TFIM, which serves as an ideal testing ground for our method as it can be solved exactly [58]. Moreover, the 1D TFIM presents the advantage that the resulting VCNs have a relatively simple and intuitive form, which is in fact that of classical Ising models. Further, for this example we have shown, in a rather intuitive manner, how the perturbatively motivated architecture of VCNs can be systematically built upon by incorporating classical couplings that expand over a certain coupling distance, which is closely related to the order of the underlying cumulant expansion. Then, we characterized the performance of various VCNs in terms of the real-time evolution of two-point correlation functions at varying distance, local off-diagonal observables, and the integrated residuals. In general, we found that the basic architecture of VCNs allows for a systematic improvement in the accuracy of such results by adding higher-order terms according to the controlled procedure referred to above. We have verified that this statement remains true when varying system size and quenching to different points in the parameter space of the associated Hamiltonian.

Next, we studied the out-of-equilibrium dynamics in a genuinely interacting two-dimensional lattice gauge theory, a $U(1)$ QLM. As opposed to the Ising model, in this case the resulting VCNs are highly nonconventional spin models; see, for example, Eq. (48). Yet, the same general ideas apply in their construction. We have shown this by explicitly calculating first- and second-order VCNs, which were then used for computing the dynamics of the QLM in different scenarios. On the one hand, we considered quasi-1D ladders of 2×10 (40 spins), with dynamics restricted to the FF sector. This allowed us to use ED for the purpose of benchmarking. We emphasize, however, that this is a rather challenging scenario for our method, as it has been found [46] that the dynamics in the FF sector embodies considerable quantum correlations that propagate throughout the entire

system. Yet, a remarkable quantitative agreement with ED was obtained. In particular, we found that second-order VCNs yield very accurate results for short-range correlations and local off-diagonal observables, up to long timescales. Importantly, we tested our method using not only homogeneous initial conditions but also nonuniform ones with a spatial energy inhomogeneity along a given column. Once again, we found that second-order VCNs are able to accurately describe the propagation of such a line defect at all distances up to long timescales. Next, we considered situations where ED and other state-of-the-art techniques become computationally inaccessible. Namely, we studied quenches starting from an equally weighted superposition of all spin configurations, thereby involving all superselection sectors. This part of our study was done in systems with 100×100 plaquettes (2×10^4 spins). Here, however, the crucial observation that the QLM exhibits disorder-free localization dynamics [46] enabled us to employ reliably our VCNs, since such localization mechanism yields a strong suppression of correlation spreading.

Overall, the quantum quenches discussed above have allowed us to characterize the accuracy and versatility of our methodology. As a principal observation, we note that upon adding higher-order couplings in the structure of a VCN, its accuracy can be improved in a controlled way. In addition, higher-order couplings also account for another crucial feature: correlation spreading is properly captured up to a spatial scale determined by the maximum coupling distance included in the architecture of the VCN. On the other hand, at a fixed order of the cumulant expansion, the accuracy can also be systematically increased by reducing the strength of the off-diagonal perturbation. As a rule of thumb, we expect a low-order VCN to give sufficiently accurate results up to a timescale set by the inverse of the perturbation strength [31]. However, such timescales can be further prolonged when considering higher-order VCNs. Besides, for a given order of the cumulant expansion, in general, a VCN outperforms its corresponding pCN at all relevant timescales, and in some cases, the former can account for important features of the quantum dynamics, which are just beyond the scope of the latter, such as the relaxation of observables toward a steady-state value. We have checked that all the remarks made above also hold when varying system size, and quenching to different points in the parameter space of the studied Hamiltonians.

Naturally, our method is not exempt from limitations and drawbacks, which can be understood from the considerations made in the previous paragraph. First and foremost, the addition of higher-order terms is accompanied by an exponential growth in complexity, so in practice, we are limited to a given order of the cumulant expansion. Moreover, we also know that a VCN will fail at capturing correlations beyond their maximum coupling distance, as illustrated very clearly in Fig. 6. Such coupling distance is also determined by the order of the cumulant expansion. Also, in general, we expect a breakdown in the description of the evolution of observables after a timescale set by the inverse of the perturbation strength, at least, as far as low-order VCNs are concerned.

Regarding possible further applications, there are several interesting routes that one could explore employing VCNs. A particularly promising one consists of formulating hybrid

approaches that combine VCNs with ANNs, so as to mitigate some of the drawbacks sustained by both kinds of generative machines. Also, as we have already mentioned, the applicability of the approach presented in this paper is not fundamentally restricted by spatial dimensionality. Thus, in this respect, the following models arise as natural extensions of the systems studied here: 2D TFIMs, 2D bosonic and fermionic Hubbard models, 2D QLMs in the presence of a dynamical matter field, 3D quantum spin ice models, as well as other kinetically constrained models. Lastly, regarding VCNs as generative machines also raises the possibility of employing them for tasks other than solving the real-time dynamics of quantum many-body systems. In particular, VCNs could be used for addressing ground-state search problems, with an optimization procedure guided by a conventional variational principle that minimizes the energy functional, rather than using a TDVP.

Finally, let us emphasize that the approach presented in this paper not only allows us to address theoretical questions regarding higher-dimensional systems but also could provide a theoretical description of recent experiments in two dimensions using quantum simulators. Among such experiments are studies of two-dimensional many-body localization dynamics employing bosons [8] and fermions [6,11], experiments probing transport properties of various 2D lattice systems [5,16], as well as investigations on other aspects of the quench dynamics of several 2D quantum spin systems [7,12,15]. Most of the aforementioned experiments explore regimes that are still inaccessible to simulation methods relying on classical resources, thereby calling for a larger input from theoretical and numerical sides. We believe that the method introduced in this work does constitute an important step into this direction.

[1] M. Greiner, O. Mandel, T. W. Hänsch, and I. Bloch, Nature (London) **419**, 51 (2002).

[2] T. Kinoshita, T. Wenger, and D. S. Weiss, Nature (London) **440**, 900 (2006).

[3] W. S. Bakr, A. Peng, M. E. Tai, R. Ma, J. Simon, J. I. Gillen, S. Fölling, L. Pollet, and M. Greiner, Science **329**, 547 (2010).

[4] J. Simon, W. S. Bakr, R. Ma, M. E. Tai, P. M. Preiss, and M. Greiner, Nature (London) **472**, 307 (2011).

[5] S. Hild, T. Fukuhara, P. Schauß, J. Zeiher, M. Knap, E. Demler, I. Bloch, and C. Gross, Phys. Rev. Lett. **113**, 147205 (2014).

[6] S. S. Kondov, W. R. McGehee, W. Xu, and B. DeMarco, Phys. Rev. Lett. **114**, 083002 (2015).

[7] H. Labuhn, D. Barredo, S. Ravets, S. de Léséleuc, T. Macrì, T. Lahaye, and A. Browaeys, Nature (London) **534**, 667 (2016).

[8] J.-y. Choi, S. Hild, J. Zeiher, P. Schauß, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross, Science **352**, 1547 (2016).

[9] 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 (London) **551**, 579 (2017).

[10] P. Jurcevic, H. Shen, P. Hauke, C. Maier, T. Brydges, C. Hempel, B. P. Lanyon, M. Heyl, R. Blatt, and C. F. Roos, Phys. Rev. Lett. **119**, 080501 (2017).

[11] P. Bordia, H. Lüschen, S. Scherg, S. Gopalakrishnan, M. Knap, U. Schneider, and I. Bloch, Phys. Rev. X **7**, 041047 (2017).

[12] V. Lienhard, S. de Léséleuc, D. Barredo, T. Lahaye, A. Browaeys, M. Schuler, L.-P. Henry, and A. M. Läuchli, Phys. Rev. X **8**, 021070 (2018).

[13] S. Erne, R. Bücker, T. Gasenzer, J. Berges, and J. Schmiedmayer, Nature (London) **563**, 225 (2018).

[14] D. Barredo, V. Lienhard, S. de Léséleuc, T. Lahaye, and A. Browaeys, Nature (London) **561**, 79 (2018).

[15] E. Guardado-Sanchez, P. T. Brown, D. Mitra, T. Devakul, D. A. Huse, P. Schauß, and W. S. Bakr, Phys. Rev. X **8**, 021069 (2018).

[16] E. Guardado-Sanchez, A. Morningstar, B. M. Spar, P. T. Brown, D. A. Huse, and W. S. Bakr, Phys. Rev. X **10**, 011042 (2020).

[17] M. A. Cazalilla and J. B. Marston, Phys. Rev. Lett. **88**, 256403 (2002).

[18] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).

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

[20] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, J. Stat. Mech. (2004) P04005.

[21] S. R. White and A. E. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004).

[22] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).

[23] M. Zwolak and G. Vidal, Phys. Rev. Lett. **93**, 207205 (2004).

[24] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, Ann. Phys. (NY) **411**, 167998 (2019).

[25] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quantum Inf. Comput. **7**, 401 (2007).

[26] F. Verstraete, V. Murg, and J. Cirac, Adv. Phys. **57**, 143 (2008).

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

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

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

[30] M. Schmitt and M. Heyl, Phys. Rev. Lett. **125**, 100503 (2020).

[31] M. Schmitt and M. Heyl, SciPost Phys. **4**, 013 (2018).

[32] V. Murg, F. Verstraete, and J. I. Cirac, Phys. Rev. A **75**, 033605 (2007).

[33] J. Hauschild, F. Pollmann, and F. Heidrich-Meisner, Phys. Rev. A **92**, 053629 (2015).

[34] M. P. Zaletel, R. S. K. Mong, C. Karrasch, J. E. Moore, and F. Pollmann, Phys. Rev. B **91**, 165112 (2015).

[35] T. Hashizume, I. P. McCulloch, and J. C. Halimeh, arXiv:1811.09275.

[36] P. Czarnik, J. Dziarmaga, and P. Corboz, Phys. Rev. B **99**, 035115 (2019).

[37] C. Hubig and J. I. Cirac, SciPost Phys. **6**, 31 (2019).

[38] A. Kshetrimayum, M. Goihl, and J. Eisert, Phys. Rev. B **102**, 235132 (2020).

[39] C. Hubig, A. Bohrdt, M. Knap, F. Grusdt, and J. I. Cirac, SciPost Phys. **8**, 21 (2020).

[40] T. Hashizume, J. C. Halimeh, and I. P. McCulloch, Phys. Rev. B **102**, 035115 (2020).

[41] B. Kloss, Y. B. Lev, and D. R. Reichman, SciPost Phys. **9**, 70 (2020).

[42] G. Fabiani and J. H. Mentink, SciPost Phys. **7**, 4 (2019).

[43] I. López-Gutiérrez and C. B. Mendl, arXiv:1912.08831.

[44] J. Richter, T. Heitmann, and R. Steinigeweg, SciPost Phys. **9**, 31 (2020).

[45] P. Kramer and M. Saraceno, *Geometry of the Time-Dependent Variational Principle in Quantum Mechanics* (Springer, Berlin, 1981).

[46] P. Karpov, R. Verdel, Y.-P. Huang, M. Schmitt, and M. Heyl, Phys. Rev. Lett. **126**, 130401 (2021).

[47] A. Smith, J. Knolle, D. L. Kovrizhin, and R. Moessner, Phys. Rev. Lett. **118**, 266601 (2017).

[48] A. Smith, J. Knolle, R. Moessner, and D. L. Kovrizhin, Phys. Rev. Lett. **119**, 176601 (2017).

[49] M. Brenes, M. Dalmonte, M. Heyl, and A. Scardicchio, Phys. Rev. Lett. **120**, 030601 (2018).

[50] I. Papaefstathiou, A. Smith, and J. Knolle, Phys. Rev. B **102**, 165132 (2020).

[51] P. A. McClarty, M. Haque, A. Sen, and J. Richter, Phys. Rev. B **102**, 224303 (2020).

[52] R. Kubo, J. Phys. Soc. Jpn. **17**, 1100 (1962).

[53] G. Carleo, F. Becca, M. Schiró, and M. Fabrizio, Sci. Rep. **2**, 243 (2012).

[54] G. Carleo, F. Becca, L. Sanchez-Palencia, S. Sorella, and M. Fabrizio, Phys. Rev. A **89**, 031602(R) (2014).

[55] G. Carleo, L. Cevolani, L. Sanchez-Palencia, and M. Holzmann, Phys. Rev. X **7**, 031026 (2017).

[56] R. Jastrow, Phys. Rev. **98**, 1479 (1955).

[57] B. Blaß and H. Rieger, Sci. Rep. **6**, 38185 (2016).

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

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

[60] M. Heyl, A. Polkovnikov, and S. Kehrein, Phys. Rev. Lett. **110**, 135704 (2013).

[61] M. Heyl, Rep. Prog. Phys. **81**, 054001 (2018).

[62] U.-J. Wiese, Ann. Phys. (Berlin) **525**, 777 (2013).

[63] M. Hermele, M. P. A. Fisher, and L. Balents, Phys. Rev. B **69**, 064404 (2004).

[64] L. Tagliacozzo, A. Celi, A. Zamora, and M. Lewenstein, Ann. Phys. (NY) **330**, 160 (2013).

[65] A. W. Glaetzle, M. Dalmonte, R. Nath, I. Rousochatzakis, R. Moessner, and P. Zoller, Phys. Rev. X **4**, 041037 (2014).

[66] D. Marcos, P. Widmer, E. Rico, M. Hafezi, P. Rabl, U.-J. Wiese, and P. Zoller, Ann. Phys. (NY) **351**, 634 (2014).

[67] A. Celi, B. Vermersch, O. Viyuela, H. Pichler, M. D. Lukin, and P. Zoller, Phys. Rev. X **10**, 021057 (2020).

[68] Y.-P. Huang, D. Banerjee, and M. Heyl, Phys. Rev. Lett. **122**, 250401 (2019).

[69] P. Sala, T. Rakovszky, R. Verresen, M. Knap, and F. Pollmann, Phys. Rev. X **10**, 011047 (2020).

[70] V. Khemani, M. Hermele, and R. Nandkishore, Phys. Rev. B **101**, 174204 (2020).

[71] T. Rakovszky, P. Sala, R. Verresen, M. Knap, and F. Pollmann, Phys. Rev. B **101**, 125126 (2020).

[72] Z.-C. Yang, F. Liu, A. V. Gorshkov, and T. Iadecola, Phys. Rev. Lett. **124**, 207602 (2020).