Disorder-free localization in an interacting two-dimensional lattice gauge theory

P. Karpov,1, 2,∗ R. Verdel,1 Y.-P. Huang,3 M. Schmitt,4 and M. Heyl1

1Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Straße 38, Dresden 01187, Germany
2National University of Science and Technology “MISIS”, Moscow, Russia
3The Paul Scherrer Institute, Forschungsstrasse 111, 5232 Villigen, Switzerland
4Department of Physics, University of California at Berkeley, Berkeley, CA 94720, USA

(Dated: March 10, 2020)

Disorder-free localization has been recently introduced as a mechanism for ergodicity breaking in low-dimensional homogeneous lattice gauge theories caused by local constraints imposed by gauge invariance. We show that also genuinely interacting systems in two spatial dimensions can become nonergodic as a consequence of this mechanism. Specifically, we prove nonergodic behavior in the quantum link model by obtaining a rigorous bound on the localization-delocalization transition through a classical correlated percolation problem implying a fragmentation of Hilbert space on the nonergodic side of the transition. We study the quantum dynamics in this system by means of an efficient and perturbatively controlled representation of the wavefunction in terms of a variational network of classical spins akin to artificial neural networks. We identify a distinguishing dynamical signature by studying the propagation of line defects, yielding different light cone structures in the localized and ergodic phases, respectively. The methods we introduce in this work can be applied to any lattice gauge theory with finite-dimensional local Hilbert spaces irrespective of spatial dimensionality.

Introduction. Systems with local constraints play an important role in various physical contexts ranging from strongly correlated electrons [1] and frustrated magnets [2, 3] to fundamental theories of matter such as quantum electro- and chromodynamics [4], where constraints take the form of local gauge symmetries. The equilibrium properties of such systems have been extensively studied over the last decades, but only recently their nonequilibrium dynamics has moved into focus. In particular, local constraints have emerged as a new paradigm for ergodicity breaking, besides the two known archetypical scenarios caused by localization due to strong disorder or integrability. Systems with local constraints can exhibit rare nonergodic eigenstates, termed quantum many-body scars [5, 6], or extremely slow relaxation [7–9], whereas dipole conservation can prevent thermalization of large parts of the spectrum in one-dimensional fractonic systems [10–13]. A particularly generic mechanism for nonergodic behavior is hosted in lattice gauge theories (LGTs) where local constraints emerge naturally due to the local gauge symmetry, leading to an extensive number of local conserved quantities. Specifically, this can lead to the absence of ergodicity in 1D LGTs with discrete [14, 15] and continuous [16] gauge symmetries or for higher-dimensional systems in the low-energy limit [17] or when they are free [18]. It has remained, however, a key challenge to rigorously identify non-ergodic behavior in genuinely interacting quantum matter beyond one dimension.

In this work we show that the 2D $U(1)$ quantum link model features both localized and ergodic phases in the absence of disorder. Our proof relies on a mapping onto a classical correlated percolation problem providing a rigorous bound on the localization transition of the quantum model. We identify a distinguishing quantum-dynamical signature of the two phases by studying the propagation of an initial line defect, which leads to two different light cone structures. For the description of the nonequilibrium dynamics of the system we introduce variational classical networks (vCNs), which provide an efficient and perturbatively controlled representation of the quantum many-body wave function in terms of a network of classical spins akin to artificial neural networks (ANNs). The introduced methods can be applied to any LGT with finite local Hilbert spaces irrespective of dimensionality.

Quantum link model. We study the 2D $U(1)$ quantum link model (QLM) [19, 20], which has been introduced as a descendant of lattice quantum electrodynamics with spin-$1/2$ gauge degrees of freedom. In the QLM the spins $S_{r,\mu}$ reside on the links of a square lattice connecting vertices $r = (x, y)$ and $r + \mu$ (here $\mu = \hat{i}, \hat{j}$ is one of the two unit vectors of the lattice, Fig. 1(a)), with the Hamiltonian:

$$H = H_0 + V \equiv \lambda \sum_{\square}(U_{\square} + U_{\square}^\dagger)^2 - J \sum_{\square}(U_{\square} + U_{\square}^\dagger)$$

(1)

The sums run over all plaquettes $\square$, $U_{\square} = S_{r,\hat{i}}^+ S_{r+\hat{i},\hat{j}}^- S_{r+\hat{j},\hat{i}}^- S_{r,\hat{j}}^-$ induces a collective flip of all spins on plaquette $\square$, and $S_{r,\mu}^{\pm}$ denote the raising and lowering operators. The first (potential) term counts the number of flippable plaquettes and the second (kinetic) term induces coherent dynamics. For what follows, we will consider periodic boundary conditions and the case of a strong potential term with $J/\lambda = -0.1$. The QLM not only appears in the context of high-energy physics, but also shares strong connections to condensed matter systems featuring quantum spin ice phases [21, 22] or...
FIG. 1. a) Illustration of the U(1) quantum link model (QLM) with spin-1/2’s located on the links of the square lattice. Spins pointing $\rightarrow$ or $\uparrow$ correspond to $S^z = +1$ and $\leftarrow$ or $\downarrow$ to $S^z = -1$, respectively. Kinetics is introduced by plaquette-flip operators $U_\square$, $U_\square^{\dagger}$ (shown for the darkened central plaquette) whenever the spins on a plaquette are oriented clockwise or counterclockwise. Flippable plaquettes are denoted by circular arrows. Background charges with nonzero in-our outflow of electric field at a given vertex are indicated by red a blue dots. b) Spatiotemporal buildup of quantum correlations $\langle S^z_{\mu}(t)S^z_{\mu+d,y}(t) \rangle = \langle S^z_0(t)S^z_d(t) \rangle$ ($d = d\bar{d}$) starting from $|\psi(\alpha = 0)\rangle = |\rightarrow\rangle$ in the localized phase of the QLM for $J/\lambda = -0.1$ and a system of size $20 \times 20$, i.e., 800 spins.

quantum dimer models [1, 23]. On the experimental side various proposals have explored the potential realization of the QLM in quantum simulators within the last years [24, 25].

The local gauge symmetry of the QLM is generated by the operators $G_r = \sum_\mu (S^x_{r,\mu} - S^x_{r,-\mu})$ counting the total inflow of the electric field to the vertex $r$. Since $[G_r,H] = 0$ for all lattice points and $[G_r,G_{r'}] = 0$, eigenstates of $H$ can be classified by the respective eigenvalues $q_r \in \{-2,-1,0,1,2\}$ of $G_r$. The set of $q = \{q_r\}$ defines the so-called superselection sector of states $|\psi_q\rangle$ with $G_r |\psi_q\rangle = q_r |\psi_q\rangle$, so that each of the $q_r$ can be given a physical meaning in terms of static background charges located at $r$ [16]. The QLM further has global conserved quantities $\Phi_x = \sum_y S^z_{r,i,y}$, $\Phi_y = \sum_x S^z_{r,j,x}$, which define the flux sectors.

Disorder-free localization. The existence of these sectors, protected by gauge invariance, can lead to an unconventional scenario for ergodicity breaking. Consider a homogeneous superposition state $|\psi\rangle = \sum_q C_q |\psi_q\rangle$ involving many superselection sectors. As the Hamiltonian and typical observables are block-diagonal, i.e., $H |\psi_q\rangle = H_q |\psi_q\rangle$, the expectation values of an operator $O$ during dynamics become equivalent to $\langle O(t) \rangle = \sum_q |C_q|^2 \langle \psi_q | e^{itH} O e^{-itH} | \psi_q \rangle$ resembling an effective disorder average with the disorder strength determined by the random background charges in the typical superselection sectors [16]. This can, in principle, lead to non-ergodic behavior of $O(t)$, although both the initial state and the Hamiltonian are homogeneous leading to the notion of disorder-free localization [14].

Initial states for time-evolution. We now aim to characterize the nonequilibrium dynamics of the QLM for the following homogeneous initial states $|\psi_0\rangle$:

(i) $|\psi_0\rangle = |\rightarrow\rangle = \bigotimes_i |\uparrow_i\rangle + |\downarrow_i\rangle$, where $|\uparrow_i\rangle$ and $|\downarrow_i\rangle$ are the two basis states at link $i$. This state is distributed over all superselection sectors of the model.

(ii) $|\psi_0\rangle = |\rightarrow\rangle_{FF}$ which is a projection of $|\rightarrow\rangle$ to a single “fully-flippable” (FF) sector, defined as the zero-charge zero-flux sector, $|\rightarrow\rangle_{FF}$ is an equal-weighted superposition of all states from the FF sector (i.e. the Rokhsar-Kivelson state [1] for the FF sector).

While $|\rightarrow\rangle$ is a product state, $|\rightarrow\rangle_{FF}$ is entangled. Nevertheless, $|\rightarrow\rangle$ can be continuously connected to a product state from the same FF sector via

$$|\psi(\alpha)\rangle = \bigotimes_i [\cos(\alpha + \frac{\pi}{4}) |\mathbf{FF}\rangle + \sin(\alpha + \frac{\pi}{4}) |\overline{\mathbf{FF}}\rangle],$$

with $\alpha \in [0, \pi/4]$. Here, $|\mathbf{FF}\rangle$ and $|\overline{\mathbf{FF}}\rangle$ denote the local spin orientations of the two states with all plaquettes flippable and therefore with checkerboard alternating clockwise (FF) and anticlockwise (FF$_\bar{}$) orientations. By construction, $|\psi(\alpha = 0)\rangle = |\rightarrow\rangle$ and $|\psi(\alpha = \pi/4)\rangle = |\mathbf{FF}\rangle$.

Importantly, the states $|\psi(\alpha)\rangle$ are spatially uniform. The resulting dynamics for $\alpha = 0$ is displayed in Fig. 1(b), where we monitor the spatiotemporal buildup of quantum correlations. We will identify the limited spatial propagation with nonergodic behavior below.

Variational classical networks. Calculating the dynamics of interacting quantum systems in 2D is an inherently hard problem without a general-purpose computational method available to date. Representing a generic quantum many-body state as $|\psi\rangle = \sum_\mathbf{s} \psi(\mathbf{s}) |\mathbf{s}\rangle$ requires, in principle, the storage of exponentially many amplitudes $\psi(\mathbf{s})$ (here $\mathbf{s} = (s_1,...,s_N)$). Recently, it has been proposed to use networks of classical spins to solve this problem [26, 27] by avoiding to store the $\psi(\mathbf{s})$’s. The amplitudes $\psi(\mathbf{s}) \approx \exp[\mathcal{H}(\mathbf{s},W)]$ are rather generated on the fly when needed via a complex classical spin model with Hamiltonian $\mathcal{H}(\mathbf{s},W)$ determined by a set of couplings $W$ between the involved spins. Here, we construct $\mathcal{H}(\mathbf{s},W)$ using a perturbatively controlled expansion and extend the recently proposed classical networks [27] upon imposing an additional optimization principle. The resulting approach can be interpreted as encoding $|\psi\rangle$ in an ANN with a specific simplified network structure.

Within the vCNs we perform an expansion around a classical limit, which in the case of the QLM is the potential term $H_0$ in (1). By representing the evolution operator in the interaction picture $W(t) = T \exp(-\int_0^t dt' V(t'))$ we can write
ψ(\vec{s},t) = \langle \psi_0 | e^{-iHt} | \psi_0 \rangle = e^{iE_\vec{s}t} \langle \vec{s} | W(t) | \psi_0 \rangle, \quad \text{where} \quad H_0 | \vec{s} \rangle = E_\vec{s} | \vec{s} \rangle. \quad \text{For the remaining term} \quad \langle \vec{s} | W(t) | \psi_0 \rangle \text{we perform a cumulant expansion for time-ordered exponential operators} \quad [27–29], \quad \text{which, e.g., to the first order yields:} \quad H' = -iE_\vec{s}t - \langle \vec{s} | \int_0^t dt' V(t')} | \psi_0 \rangle / \langle \vec{s} | \psi_0 \rangle. \quad \text{Taking} \quad | \rightarrow \rangle \text{one obtains for the QLM} \quad H(| \vec{s} \rangle, t) = -iE_\vec{s}t - iJ \sum_{\square} \int_0^t dt' e^{i\omega_\square(t')} \quad \text{Here} \quad \sum_{\square} \int_0^t dt' \text{denotes the sum over all flippable plaquettes in the spin configuration} \quad \vec{s}, \quad \text{and} \quad \omega_\square = -4, ..., 4 \text{counts the difference between number of flippable plaquettes surrounding the given} \quad \square \text{before and after its flip (i.e.} \quad \omega_\square \text{gives the potential energy difference before and after the flip). For example, for the configuration in Fig. 1(a) we have} \quad \omega_\square = 3 - 1 = 2 \quad \text{for the central plaquette. Going beyond previous work} \quad [27] \quad \text{we promote} \quad \lambda t \text{and its functions such as} \quad \int_0^t dt' e^{i\lambda_\square(t')} \text{to ten variational parameters} \quad W_k(t) = (W_0^{(0)}, W_0^{(1)}, ..., W_4^{(1)}) \quad \text{yielding} \quad H(\vec{s}, W_0(t)) = -iE_\vec{s}W_0^{(0)} - iJ \sum_{\square} W_0^{(1)}(t) \quad \text{The local connectivity of the vCN is encoded in the function} \quad \omega_\square(\vec{s}). \quad \text{For the actual shown numerical simulations we use a second-order ansatz and more complex initial states (see} \quad [29, 30]). \quad \text{The} \quad W_k(t) \text{’s are determined by a time-dependent variational principle translating quantum dynamics into a system of coupled classical differential equations} \quad \sum_{\square} \sum_{k,k'} S_{k,k'} \omega_\square = -iF_k \text{in the space of variational parameters. Here,} \quad S_{k,k'} = (O_k^*O_{k'}) - \langle O_k^*O_{k'} \rangle, \quad \text{and} \quad F_k = (E_{\text{loc}}O_k^*) - \langle E_{\text{loc}}O_k^* \rangle, \quad \text{with} \quad O_k(\vec{s}) = \partial_{\psi_0} \psi(\vec{s}) / \partial W_k \quad \text{E}_{\text{loc}}(\vec{s}) = \langle \vec{s} | H | \psi_0 \rangle / \langle \vec{s} | \psi_0 \rangle \quad \text{The}} \quad E_{\text{loc}} \quad \text{denotes the expected energy contribution and use the character of energy propagation to distinguish between ergodic and localized dynamics. Concretely, we consider the two initial conditions} \quad | \psi_0 \rangle = | \rightarrow \rangle \text{or} \quad | \rightarrow \rangle_{\text{FF}} \text{upon applying in addition} \quad P = \prod_{\square \in C_d} \left[ 1 + (U_\square + U_\square^*) \right] ^2 \text{along all plaquettes in column} \quad d = 0; \text{here} \quad C_d \text{denotes the set of plaquettes in column} \quad d. \quad \text{In Fig. 2 we plot the (normalized) column energy} \quad \varepsilon_d(t) = \langle H_d(t) \rangle - H_{\text{av}} \quad \text{with} \quad H_d = \sum_{\square \in C_d} H_\square \text{the total energy for the plaquettes in d-th column,} \quad H_\square = \lambda(U_\square + U_\square^*)^2 - J(U_\square + U_\square^*). \quad \text{Further,} \quad H_{\text{av}} = \langle H \rangle / L \text{denotes the expected} \quad \langle H_d(t) \rangle \text{in the long-time limit when the system is thermalizing} \quad (L \text{is the number of columns).} \quad \text{Comparing Figs. 2(a) and (c) we observe that the dynamics differs qualitatively for the two initial conditions, although the Hamiltonian parameters are identical. While for} \quad | \psi_0 \rangle = | \rightarrow \rangle \text{energy transport is highly suppressed and only visible on short distances (Fig. 2(a)), the opposite happens for} \quad | \rightarrow \rangle_{\text{FF}}. \quad \text{This becomes even more apparent in Figs. 2(b,d), where} \quad \varepsilon_d(t) \text{relative to the initial value} \quad \varepsilon_d(0) \text{is shown, therefore more directly highlighting energy propagation. While for} \quad | \rightarrow \rangle_{\text{FF}} \text{we identify a linearly propagating front, for} \quad | \rightarrow \rangle \text{we observe a strong bending. We argue below that this front for} \quad | \rightarrow \rangle \text{can extend only to a finite region as a consequence of disorder-free localization.} \quad 

FIG. 2. \quad \text{Quantum energy dynamics for line defects created in the} \quad | \rightarrow \rangle \text{ (upper panel) and} \quad | \rightarrow \rangle_{\text{FF}} \text{ (lower panel) initial states, displaying the normalized plaquette energy} \quad \varepsilon_d(t) \text{of d-th column. (a,c) \varepsilon_d(t) for different columns} \quad d = 2 \text{ where darker colors refer to larger distances from the initial defect. Insets show the same data including} \quad d = 0, 1 \text{ in a color plot. (b,d) Absolute deviation of} \quad \varepsilon_d(t) \text{from the initial value} \quad \varepsilon_d(0). (b) Signal propagation for} \quad | \rightarrow \rangle \text{ showing a strong bending of the light cone indicating localized behavior and (d) for} \quad | \rightarrow \rangle_{\text{FF}} \text{ consistent with linear propagation indicative of ergodic behavior. For all the plots} \quad J/\lambda = -0.1 \text{ and system size} \quad 10 \times 10.
**Bound on quantum dynamics by unconventional percolation.** The qualitative difference in the quantum dynamics for the initial states $|→⟩$ and $|→⟩_{FP}$ originates from a dynamical transition, which one can study systematically upon tuning the parameter $\alpha$ for the initial state (2). For this purpose, we employ an unconventional correlated classical percolation problem [32] and establish a bound on the quantum localized-ergodic transition in the QLM providing a numerical proof for an extended nonergodic phase as a consequence of disorder-free localization.

We illustrate the idea for the initial state $|→⟩$, distributed over all superselection sectors. Consider a typical (random) sector from this distribution (Fig. 3(a)). Such sector exhibits many background charges $q_r$ whenever the “two-in two-out” rule at vertex $r$ is violated. Importantly, these background charges (constants of motion by gauge invariance) impose strong kinetic constraints. For instance, $q_r = ±2$ implies that neighboring spins either all point inwards or outwards, hence the adjacent plaquettes remain unflippable forever. The influence of $q_r = ±1$ charges is more subtle. They make at least 2 adjacent plaquettes unflippable, while their positions might change over time.

The question we address now is whether these constraints are so strong to fragment the square lattice into sets of kinetically disconnected islands or whether one can contain an extensive (percolating) connected cluster. For that purpose we study an unconventional percolation problem using an infinite-temperature classical MC simulation. We start from the initial condition (2), sampling a random basis state (and thus a sector) with a distribution set by the amplitudes in $|ψ(α)⟩$. Then we determine which parts of the systems are kinetically connected, using MC search with random plaquette flips. The simulation is stopped when every plaquette is flipped either 0 or more than some fixed threshold (= 100) number of times (or after $10^{11}$ MC steps if this condition is still not satisfied). As a result we find the number of performed flips for each plaquette (Fig. 3(b)). Repeating this procedure for different initial configurations at a given $\alpha$ and scanning $\alpha$, we finally obtain the percolation probability (Fig. 3(c)). Most importantly, one can observe a clear evidence for a percolation threshold $\alpha_c ≈ 0.25$. Although the simulation termination condition is chosen such as to minimize the number of potentially missed “weak connections” between flippable clusters, we cannot exclude the possibility of such misses. While we don’t expect a significant impact deep in the respective phases, this caveat might become important in the vicinity of $\alpha_c$, preventing us to obtain a precise value of $\alpha_c$ and to study the critical behavior.

Since $\alpha_c > 0$, the initial state $|ψ(α = 0)⟩ = |→⟩$ corresponds to the classically non-percolating side of the transition, while from $\alpha_c < \pi/4$ it follows that state $|ψ(α = \pi/4)⟩ = |FF⟩$ and all other states from the FF-sector (including $|→⟩_{FP}$) lie on the percolating side. This classical threshold is imprinted in the quantum dynamics and ultimately leads to the strong localization observed in propagation of correlations (Fig. 1(b)) and of the energy (Fig. 2(a),(b)) for $|→⟩$. For the FF-sector state $|→⟩_{FP}$ there is no percolation constraint, which allows propagation of the signal to long distances (Fig. 2(c),(d)). This analysis sets a lower bound onto the critical value of the quantum transition $\alpha_c^{(q)}$, since the quantum system might be still localized due to interference even on the classically percolating side.

**Summary and outlook.** We have shown that genuinely interacting 2D homogeneous LGTs can become nonergodic as a consequence of disorder-free localization. This is all the more surprising as many-body localization is predicted to be unstable in 2D at elevated energy den-
sities [33], implying that gauge invariance represents a more robust mechanism of ergodicity breaking compared to conventional disorder. The key element of our analysis is a bound on the localization-delocalization transition based on a classical correlated percolation problem implying a strong fragmentation of Hilbert space into kinetically disconnected regions. Both the percolation analysis as well as the introduced variational classical networks can be directly applied to other quantum many-body systems with finite-dimensional local Hilbert spaces independent of dimensionality, such as 3D quantum spin ice systems, which might be an interesting scope of the developed techniques in the future. Further, it might be interesting to explore how the quantum and classical percolation thresholds are related to each other as well as to determine their respective critical behaviors, and whether the disorder-free localization scenario holds also in the presence of matter degrees of freedom. Our theoretical analysis appears within reach of future experiments: significant efforts in the last years have explored percolation analysis within the Marie Skłodowska-Curie grant agreement No. 891717 of the European Union’s Horizon 2020 research and innovation program. Y.-P.H. receives funding from the European Commission via the Gottfried Wilhelm Leibniz Prize and acknowledges support by the Deutsche Forschungsgemeinschaft (grant agreement No. 853443), and M. H. further acknowledges support by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme; this project has received funding from the Marie Skłodowska-Curie grant agreement No. 701647. M.S. was supported through the Leopoldina Fellowship Programme of the German National Academy of Sciences Leopoldina (LPDS 2018-07) with additional support from the Simons Foundation.

Acknowledgments. We are grateful to H. Burau, J. Chalker, M. Dalmonte, R. Moessner, and G.-Y. Zhu for helpful discussions. P.K. acknowledges the support of the Alexander von Humboldt Foundation and the Ministry of Science and Higher Education of the Russian Federation. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 853443), and M. H. further acknowledges support by the Deutsche Forschungsgemeinschaft via the Gottfried Wilhelm Leibniz Prize program. Y.-P.H. receives funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 701647. M.S. was supported through the Leopoldina Fellowship Programme of the German National Academy of Sciences Leopoldina (LPDS 2018-07) with additional support from the Simons Foundation.

* karpov@pks.mpg.de

[1] D.S. Rokhsar and S.A. Kivelson, Phys. Rev. Lett. 61, 2376 (1988).
[2] M.J.P. Gingras and P.A. McClarty, Rep. Prog. Phys. 77, 056501 (2014).
[3] M. Stern, C. Castelnovo, R. Moessner, V. Oganesyan, and S. Gopalakrishnan, arXiv 1911.05742.
[4] Lattice Gauge Theories: An Introduction (3rd ed.), H.J. Rothe (World Scientific, 2005).
[5] C.J. Turner, A.A. Michaillidis, D.A. Abanin, M. Serbyn, and Z. Papic, Nat. Phys. 14, 745 (2018).
[6] C.-J. Lin and O.I. Motrunich, Phys. Rev. Lett. 122, 173401 (2019).
[7] M. van Horssen, E. Levi, and J.P. Garrahan, Phys. Rev. B 92, 100305(R) (2015).
[8] N. Pancotti, G. Giudice, J.I. Cirac, J.P. Garrahan, and M.C. Banuls, arXiv:1910.06616.
[9] Z. Lan, M. van Horssen, S. Powell, and J.P. Garrahan, Phys. Rev. Lett. 121, 040603 (2018).
[10] P. Sala, T. Rakovszky, R. Verresen, M. Knap, and F. Pollmann, arXiv 1904.04266.
[11] S. Pai and M. Pretko, Phys. Rev. Lett. 123, 136401 (2019).
[12] S. Pai, M. Pretko, and R.M. Nandkishore, Phys. Rev. X 9, 021003 (2019).
[13] V. Khemani and R. Nandkishore, arXiv:1904.04815.
[14] A. Smith, J. Knolle, D.L. Kovrizhin, and R. Moessner, Phys. Rev. Lett. 118, 266601 (2017).
[15] A. Smith, J. Knolle, R. Moessner, and D.L. Kovrizhin, Phys. Rev. Lett. 119, 176601 (2017).
[16] M. Brenes, M. Dalmonte, M. Heyl, and A. Scardicchio, Phys. Rev. Lett. 120, 030601 (2018).
[17] I.H. Kim and J. Haah, Phys. Rev. Lett. 116, 027202 (2016).
[18] A. Smith, J. Knolle, R. Moessner, and D.L. Kovrizhin, Phys. Rev. B 97, 245137 (2018).
[19] U.-J. Wiese, Annalen der Physik 525, 777 (2013).
[20] S. Chandrasekharan and U.-J. Wiese, Nuclear Physics B 492, 455 (1997).
[21] N. Shannon, G. Misguich, and K. Penc, Phys. Rev. B 69, 220403(R) (2004).
[22] D. Banerjee, F.-J. Jiang, P. Widmer, and U.-J. Wiese, J. Stat. Mech P12010 (2013).
[23] A. Celi, B. Vemersch, O. Viyuela, H. Pichler, M.D. Lukin, P. Zoller, arXiv 1907.03311.
[24] A. Glaetzle, M. Dalmonte, R. Nath, I. Roussochatzakis, R. Moessner, and P. Zoller, Phys. Rev. X 4, 041037 (2014).
[25] G. Carleo, M. Troyer, Science 355, 602 (2017).
[26] M. Schmitt and M. Heyl, SciPost Phys. 4, 013 (2018).
[27] R. Kubo, Generalized Cumulant Expansion Method, J. Phys. Soc. Jpn. 17, 1100 (1962).
[28] See the Supplementary Material.
[29] R. Verdel, P. Karpov, Y.-P. Huang, M. Schmitt and M. Heyl, to appear.
[30] M. Schmitt and M. Heyl, arXiv:1912.08828.
[31] A. Coniglio, A. Fierro, Correlated Percolation. In: Meyers R. (eds) Encyclopedia of Complexity and Systems Science (Springer, New York, 2009).
[32] W. De Roeck and J. Z. Imbrie, Philos. Trans. Royal Soc. A 375, 20160422 (2017).
[33] E. Martinez, C. Muschik, P. Schindler, D. Nigg, A. Erdhard, M. Heyl, P. Hauke, M. Dalmonte, T. Monz, P. Zoller, and R. Blatt, Nature 534, 516 (2016).
[34] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, et al., Nature 569, 355, (2019).
[35] F. Görg, K. Sandholzer, J. Minguzzi, R. Desbuquois, M. Messer and T. Esslinger, Nature Physics 15, 1161 (2019).
[36] C. Schweizer, F. Grusdt, M. Berngruber, L. Barbiero, E. Demler, N. Goldman, I. Bloch, and M. Aidelsburger, Nature Physics 15, 1168 (2019).
[37] A. Mil, T.V. Zache, A. Hegde, A. Xia, R.P. Bhatt, M.K. Oberthaler, P. Hauke, J. Berges, and F. Jendrzejewski, arXiv:1909.07641.
Supplemental Material for “Disorder-free localization in an interacting 2D lattice
gauge theory”
P. Karpov, R. Verdel, Y.-P. Huang, M. Schmitt, and M. Heyl
(Dated: March 10, 2020)

I. CONSTRUCTION OF VARIATIONAL CLASSICAL NETWORK ANSÄTZE

In this Supplementary Section we outline the procedure of construction of a variational classical networks for the quantum link model, which can be straightforwardly generalized to other quantum models with finite local Hilbert spaces [S1]. If the Hamiltonian of the model can be written as \( H = H_0 + V \), where \( H_0 \) is diagonal in the computational basis and \( V \ll H_0 \) is a small non-diagonal perturbation, then in order to describe the evolution of the wave function we can use a perturbative treatment based on the cumulant expansion [S2]. Moreover, we can build on top of this perturbative description and construct adequate variational Ansätze for the wavefunction.

For the studied quantum link model we have \( H = H_0 + V = \lambda \sum_{\square} (U_{\square} + U_{\square}^\dagger)^2 - J \sum_{\square} (U_{\square} + U_{\square}^\dagger) \), where the sums run over all plaquettes \( \square \). In the perturbative limit it is convenient to work in the interaction representation and rewrite the evolution operator as \( e^{-iHt} = e^{-iH_0t}e^{iH't} \), where \( W(t) = T \exp(-i \int_0^t dt' V(t')) \). Let further the initial state be \( |\psi_0\rangle = \sum_{\vec{s}} \psi_0(\vec{s}) |\vec{s}\rangle \) (i.e. \( \psi_0(\vec{s}) = \langle \vec{s}|\psi_0 \rangle \)), where \( |\vec{s}\rangle = |s_1, s_2, \ldots \rangle \); and the time-evolved state \( |\psi(t)\rangle = e^{iHt} |\psi_0\rangle = \sum_{\vec{s}} \psi(\vec{s},t) |\vec{s}\rangle \).

Below we explicitly describe the construction of the first-order ansatz and sketch an analogous derivation for the second-order ansatz.

A. First-order variational ansatz

Within the above settings the coefficients for the time-evolved state \( |\psi(t)\rangle \) in the lowest-order in the cumulant expansion [S3] in the perturbation \( V \) can be obtained as

\[
\psi(\vec{s},t) = \langle \vec{s} \rangle e^{iE_\vec{s}t} W(t) |\psi_0\rangle = e^{iE_\vec{s}t} \langle \vec{s} \rangle W(t) |\psi_0\rangle = e^{iE_\vec{s}t} \langle \vec{s} \rangle 1 - i \int_0^t dt' V(t') + O(V^2) |\psi_0\rangle = \\
= \langle \vec{s} |\psi_0 \rangle e^{iE_\vec{s}t} \left( 1 - i \frac{\langle \vec{s} \rangle \int_0^t dt' V(t') |\psi_0\rangle}{\langle \vec{s} |\psi_0 \rangle} + O(V^2) \right),
\]

(S1)

where \( E_\vec{s} \) is the eigenvalue of the unperturbed Hamiltonian: \( H_0 |\vec{s}\rangle = E_\vec{s} |\vec{s}\rangle \) and we assumed that \( \langle \vec{s} |\psi_0 \rangle \) doesn’t contain any special smallness. To the lowest order in \( V \) we can reexponentiate the result and obtain

\[
\psi(\vec{s},t) = \psi_0(\vec{s}) e^{-iE_\vec{s}t} \exp \left( -i \int_0^t dt' \frac{\langle \vec{s} |V(t')|\psi_0 \rangle}{\langle \vec{s} |\psi_0 \rangle} + O(V^2) \right)
\]

(S2)

We recast our wavefunction coefficient in the form of an effective Hamiltonian of a classical spin model \( \mathcal{H} \): \( \psi(\vec{s},t) = e^{\mathcal{H}(\vec{s},t)} \) convenient for Monte Carlo sampling:

\[
\psi(\vec{s},t) = \exp \left( \ln \psi_0(\vec{s}) - iE_\vec{s}t - i \sum_{\vec{s}_1} \frac{\psi_0(\vec{s}_1)}{\psi_0(\vec{s})} \int_0^t dt' \langle \vec{s}_1 |V(t')|\vec{s}_1 \rangle + O(V^2) \right).
\]

(S3)

In order to proceed we need to find \( V(t) \), which can be expressed using \( U_{\square}(t) \). Equation of motion for operators \( U_{\square}(t) \equiv e^{iH_0 t} U_{\square} e^{-iH_0 t} \) in the interaction representation is given by

\[
-i \frac{d}{dt} U_{\square}(t) = [H_0, U_{\square}(t)]
\]

(S4)

Only four plaquettes (indexed by letters a, b, c, d, Fig. S1) adjacent to the plaquette under consideration \( \square \) contribute to the commutator \([H_0, U_{\square}] = \Omega_{\square} U_{\square}] \), where \( \Omega_{\square} \) is a diagonal operator

\[
\Omega_{\square} = -P_{a1}^+ P_{a2}^+ P_{a4}^+ - P_{b1}^+ P_{b2}^+ P_{b3}^+ - P_{c1}^+ P_{c3}^+ P_{c4}^+ - P_{d1}^+ P_{d3}^+ P_{d4}^+ \\
+ P_{a1}^- P_{a2}^- P_{a4}^- + P_{b1}^- P_{b2}^- P_{b3}^- + P_{c1}^- P_{c3}^- P_{c4}^- + P_{d1}^- P_{d3}^- P_{d4}^- \)
\]

(S5)
and $P^+$ and $P^-$ are projectors to $S^z = +1$ and $S^z = -1$ correspondingly: $P^\pm = (1 \pm S^z)/2$. Thus the equation of motion (S4) transforms to $-i\frac{d}{dt}U^{\square}(t) = \lambda(t)U^{\square}(t)$ and can be readily solved by

$$U^{\square}(t) = e^{i\lambda(t)\sum_{\square}U^{\square}}$$
$$U^{\square\dagger}(t) = e^{-i\lambda(t)\sum_{\square}U^{\square\dagger}}$$

(note that $\Omega^{\square}$ commutes with $U^{\square}$ since they do not contain the same spins).

Now we’re ready to substitute $V(t) = J \sum_{\square}(U^{\square}(t) + U^{\square\dagger}(t))$ in (S3). In order to do so it’s convenient to introduce the diagonal operator $F^{\square} = U^{\square}U^{\square\dagger} - U^{\square\dagger}U^{\square}$; so its matrix element $F^{\square} = +1, -1, 0$ is the flippability of the plaquette $\square$ in the state $s$ for counterclockwise ($+1$) or clockwise-oriented ($-1$) flippable, or non-flippable plaquettes (0) respectively. By $s^{\square}$ we denote configuration which differs from $s$ by the flip of single plaquette: $|s^{\square}\rangle = (U^{\square} + U^{\square\dagger})|s\rangle$. It’s also convenient to denote $\omega^{\square}(s) = F^{\square}(s)\Omega^{\square}(s)$. It takes values $\omega^{\square} = -4, ..., 4$ and counts the difference between number of flippable plaquettes surrounding the given $\square$ before and after its flip (i.e. $\lambda\omega^{\square}$ gives the potential energy difference before and after the flip). Then $\langle s | (e^{i\lambda(t)\sum_{\square}U^{\square}} + e^{-i\lambda(t)\sum_{\square}U^{\square\dagger}})|s^{\square}\rangle = e^{iF^{\square}(s)\Omega^{\square}(s)t} = e^{i\omega^{\square}t}$. Finally substituting $V(t)$ to (S3) we get

$$\psi(s, t) = \exp \left( \ln \psi_0(s) - iE_s t + iJ \sum_{\square} \frac{\psi_0(s^{\square})}{\psi_0(s)} \int_0^t dt' e^{i\lambda\omega^{\square}(s)t'} + O(V^2) \right)$$

(S8)

where prime over the sum indicates summation over only flippable plaquettes. This gives us the lowest-order term of the perturbative expansion of $\psi(s, t)$.

Now we want to construct a variational ansatz that should at least reproduce this perturbative result (S8) and hopefully improve it. We use the following strategy: we promote all functions of time to the variational parameters $\lambda(t)$ and for labeling the neighboring plaquettes (a, b, c, d) for a given one (central, darkened). Recall that spin directions $→$, $↑$ correspond to $S^z = +1$, while $←$ and $↓$ correspond to $S^z = -1$ and note that $S_1 \equiv S_{a3}$, $S_2 \equiv S_{b4}$ and so on.

$$\psi(\tilde{s}, t) = \exp \left( \ln \psi_0(\tilde{s}) - iE_{\tilde{s}} t + iJ \sum_{\square} \frac{\psi_0(\tilde{s}^{\square})}{\psi_0(\tilde{s})} \omega^{\square}(\tilde{s}) + O(V^2) \right)$$.  

(S11)

FIG. S1. Convention for numbering the spins within the plaquette (1, 2, 3, 4 are the bottom, right, upper, and left spins) and for labeling the neighboring plaquettes (a, b, c, d) for a given one (central, darkened). Recall that spin directions $→$, $↑$ correspond to $S^z = +1$, while $←$ and $↓$ correspond to $S^z = -1$ and note that $S_1 \equiv S_{a3}$, $S_2 \equiv S_{b4}$ and so on.
Such ansatz works for translationally-invariant initial states. For the inhomogeneous initial state we replicate this set of variational parameters for each plaquette or each column depending on the spatial symmetry of the state.

Note that the ansatz (S11) can be explicitly recast to a classical Ising-like spin model with multiple (up to 16) spin interaction terms, $W^{(1)}_{\omega_1(\bar{s})} = \sum_{\omega_1=0}^{4} \delta_{\omega_1,\omega_1(\bar{s})} W_{\omega_1}$, where for example $\delta_{\omega_1=4,\omega_2(\bar{s})} = P_1 P_3 P_4 P_5 P_6 P_7 P_8 P_9 P_{10} P_{11} P_{12} P_{13} P_{14} P_{15} P_{16} + h.c.$, with $P_i^2 = (1 \pm S_i^z)/2$. We emphasize that functions $\omega_1(\bar{s})$ and the corresponding variational parameters $W^{(1)}_{\omega_1(\bar{s})}$ encode this rather involved spin model with a complex network of local spin connections (i.e., interaction terms) in a much simpler way, while containing exactly the same information. We refer to the ansatz (S11) as to a first-order variational classical network (vCN) for the quantum link model. In the main text we show its simplified version for the case $\psi_0(\bar{s}) = \text{const}$.

B. Second-order variational ansatz

Here we sketch the analogous derivation for the second-order variational ansatz used for the calculations in the present work.

Second-order cumulant expansion is given by

$$
\psi(\bar{s}, t) = \psi_0(\bar{s}) e^{-iE_0 t} \exp \left( -i \int_0^t dt' \frac{\langle \bar{s}| V(t') \rangle |\psi_0 \rangle}{\langle \bar{s}| \psi_0 \rangle} - \int_0^t dt' \int_0^{t'} dt'' \frac{\langle \bar{s}| V(t') V(t'') \rangle |\psi_0 \rangle}{\langle \bar{s}| \psi_0 \rangle} - \frac{\langle \bar{s}| V(t') |\psi_0 \rangle \langle \bar{s}| V(t'') |\psi_0 \rangle}{\langle \bar{s}| \psi_0 \rangle^2} \right) + O(V^3) \tag{S12}
$$

Using eqs. (S6), (S7) and the resolution of identity $I = \sum \frac{|\bar{s}\rangle \langle \bar{s}|}{\langle \bar{s}| \psi_0 \rangle}$, the first term of the expansion can be recast as

$$
\frac{\langle \bar{s}| V(t') V(t'') \rangle |\psi_0 \rangle}{\langle \bar{s}| \psi_0 \rangle} = J^2 \sum_{\bar{s}', \bar{s}''} \left\{ \langle \bar{s}| V(t') \rangle |\bar{s}'\rangle \langle \bar{s}'| V(t'') |\psi_0 \rangle \right\} e^{i\omega_1(\bar{s}')} e^{i\omega_2(\bar{s}'') t''} \tag{S13}
$$

Here $\sum''$ indicates that we sum only over such plaquette pairs that plaquette $\bar{s}'$ is flippable in configuration $\bar{s}$ and plaquette $\bar{s}''$ is flippable in configuration $\bar{s}'$; the summation is over the pairs of nearest neighboring and next-nearest neighboring plaquettes (contributions from more distant neighbors are canceled by the second term in the curly brackets in (S12)). After integration over $t'$ and $t''$ eq. (S13) transforms to

$$
- \int_0^t dt' \int_0^{t'} dt'' \frac{\langle \bar{s}| V(t') V(t'') |\psi_0 \rangle}{\langle \bar{s}| \psi_0 \rangle} = J^2 \sum_{\bar{s}', \bar{s}''} \left\{ \langle \bar{s}| V(t') \rangle |\bar{s}'\rangle \langle \bar{s}'| V(t'') |\psi_0 \rangle \right\} f(\omega_1(\bar{s}), \omega_2(\bar{s}'), t) \tag{S14}
$$

The pair $(\omega_1(\bar{s}), \omega_2(\bar{s}''))$ takes (not more than) $9^2 = 81$ possible values and we promote function $f(\omega_1(\bar{s}), \omega_2(\bar{s}''), t)$ for each of these values to a variational parameter

$$
f(\omega_1(\bar{s}), \omega_2(\bar{s}''), t) = \sum_{\omega_1, \omega_2=0}^{4} \delta_{\omega_1, \omega_1(\bar{s})} \delta_{\omega_2, \omega_2(\bar{s}'')} f(\omega_1, \omega_2, t) \rightarrow \sum_{\omega_1, \omega_2=0}^{4} \delta_{\omega_1, \omega_1(\bar{s})} \delta_{\omega_2, \omega_2(\bar{s}'')} W^{(2)}_{\omega_1(\bar{s}), \omega_2(\bar{s}'')} = W^{(2)}_{\omega_1(\bar{s}), \omega_2(\bar{s}'')} \tag{S15}
$$

Analogously the second term is

$$
- \int_0^t dt' \int_0^{t'} dt'' \frac{\langle \bar{s}| V(t') V(t'') |\psi_0 \rangle}{\langle \bar{s}| \psi_0 \rangle^2} = J^2 \sum_{\bar{s}', \bar{s}''} \left\{ \langle \bar{s}| V(t') \rangle |\bar{s}'\rangle \langle \bar{s}'| V(t'') |\psi_0 \rangle \right\} \tilde{f}(\omega_1(\bar{s}), \omega_2(\bar{s}'), t) \tag{S16}
$$

Here the prime denotes sum over all pairs of flippable plaquettes. We promote $\tilde{f}(\omega_1(\bar{s}), \omega_2(\bar{s}''), t)$ to 81 variational parameters $\tilde{W}^{(2)}$. Finally, the second-order vCN ansatz reads

$$
\psi(\bar{s}, t) = \exp \left( \ln \psi_0(\bar{s}) - i E_0 t W^{(0)} - i J \sum_{\bar{s}, \bar{s}'} \frac{\psi_0(\bar{s})}{\psi_0(\bar{s}')} W^{(1)}_{\omega_1(\bar{s}), \omega_1(\bar{s}')}(\bar{s}) + J^2 \sum_{\bar{s}', \bar{s}''} \frac{\psi_0(\bar{s})}{\psi_0(\bar{s}')} W^{(2)}_{\omega_1(\bar{s}), \omega_2(\bar{s}'')}(\bar{s}) - J^2 \sum_{\bar{s}', \bar{s}''} \frac{\psi_0(\bar{s})}{\psi_0(\bar{s}')^2} \tilde{W}^{(2)}_{\omega_1(\bar{s}), \omega_2(\bar{s}'')}(\bar{s}) \right) \tag{S17}
$$
II. TIME-DEPENDENT VARIATIONAL PRINCIPLE

The time-dependent variational principle (TDVP) is an approximate Monte Carlo (MC) scheme of evolving the wavefunction provided some suitable ansatz [S4]. Here we briefly review its main steps.

First, we choose a suitable ansatz \( \psi(\mathbf{s}, \mathbf{W}) \) for a given problem (such as the classical network introduced in the previous supplementary section and used in the present work). Here \( \mathbf{W} \) denotes a set of complex time-dependent variational parameters \( \mathbf{W}(t) = \{W_0(t), W_1(t), \ldots, W_K(t)\} \), where we include an additional “dummy” parameter \( W_0 \) that fixes the normalization and the global phase of the wavefunction.

Given a state \( |\psi_W\rangle = \sum |\psi(\mathbf{s}, \mathbf{W})\rangle |\mathbf{s}\rangle \) at time \( t \) we consider an infinitesimally evolved exact state \( |\Phi\rangle = |\psi_W\rangle - i\delta t H |\psi_W\rangle \) and an approximately evolved variational state \( |\Psi\rangle = |\psi_W\rangle + \delta t \partial_t |\psi_W\rangle \). Let us introduce the total rate of departure of the variational state \( |\Psi\rangle \) with respect to the exactly evolved state \( |\Phi\rangle \):

\[
\Delta^2(t) = \sum_{\{\mathbf{s}\}} |\dot{\psi}(\mathbf{s}, t) - \dot{\Phi}(\mathbf{s}, t)|^2 = \sum_{\{\mathbf{s}\}} |\psi(\mathbf{s}, \mathbf{W})|^2 \sum_{k=0}^K O_k(\mathbf{s}) \dot{W}_k + i E_{\text{loc}}(\mathbf{s})|^2, \tag{S18}
\]

where \( E_{\text{loc}}(\mathbf{s}) = \langle \mathbf{s} | H | \psi_W\rangle / \langle \mathbf{s} | \psi_W\rangle \) is the “local energy”, \( O_k(\mathbf{s}) = \partial \ln |\psi(\mathbf{s}, \mathbf{W})|^2 / \partial W_k \). By minimizing \( \Delta^2 \) with respect to \( \mathbf{W}^* \) we obtain a system of linear first-order differential equations for evolution of the variational parameters \( \mathbf{W}(t) \):

\[
\sum_{k'=1}^K S_{k,k'} \dot{W}_{k'} = -i F_k, \tag{S19}
\]

where \( S \) is the so-called covariance matrix with the matrix elements \( S_{k,k'} = \langle O'^*_k O_k \rangle - \langle O'^*_k \rangle \langle O_k \rangle \), and \( F_k = \langle E_{\text{loc}} O'^*_k \rangle - \langle E_{\text{loc}} \rangle \langle O'^*_k \rangle \). Note that \( S \), \( F \) and other observables can be efficiently measured by performing a Monte Carlo sampling of \( |\psi(\mathbf{s}, \mathbf{W})|^2 \). Solving the system of linear equations (S19) with respect to \( \mathbf{W} \) we obtain the time-derivatives \( \dot{\mathbf{W}}(t) \), which can be integrated and give us the approximately evolved state \( \psi(\mathbf{s}, \mathbf{W}(t)) \). Importantly, the error production rate function (S18) gives us a self-contained method of quantifying the quality of our approximate solution.
III. BENCHMARKS

In this Supplementary Section we provide benchmarks for the variational classical network method. The approximate time-evolved state is constructed by using the 2nd order perturbational ansatz (introduced in Suppl. Sec. I), which is evolved using TDVP (Suppl. Sec. II). During the evolution we sample various observables using the Monte Carlo method. We compare the vCN results with the results of exact diagonalization for the fully flippable sector of $10 \times 2$ system (the dimension of the Hilbert space is 17906). Below we present a) single-plaquette observables, b) correlation functions, and c) error production rate for translationally-invariant initial state $\mid \rightarrow \rangle_{FF}$ and a non-uniform initial state $P \mid \rightarrow \rangle_{FF}$ defined in the main text.

Figures S2-S4 show that up to time $\lambda t \approx |\lambda/J| \approx 10$, vCN method works very well for both the single-plaquette observables (Figs. S2-S3) and for the long-distance correlators (Fig. S4), which is also manifested in the fact that the error production rate remains small till $\lambda t \approx 10$ and starts significantly grow after that (Fig. S5). Nevertheless we see that vCN can capture the single plaquette observables up to much longer times, especially the average behavior with smoothed out high-frequency oscillations (note much smaller vertical scale for $d = 1 - 5$ compared to $d = 0$ in Fig. S3).

A. Single-plaquette observables

FIG. S2. Mean kinetic energy per plaquette $-J(U_{\Box} + U_{\Box}^\dagger)$ for the uniform initial state $\mid \rightarrow \rangle_{FF}$. Comparison of the exact diagonalization (ED, green) and vCN (dashed red).

FIG. S3. Energy of $d$-th column for $d = 0, \ldots, 5$ for the non-uniform initial state $P \mid \rightarrow \rangle_{FF}$ (defined in the main text) with the excess of energy around $d = 0$. Comparison of ED (green) and vCN (dashed red).
B. Correlation functions

FIG. S4. Correlation function $\langle S_z^0(t)S_z^d(t) \rangle$ for $d = 1, \ldots, 5$ for the uniform initial state $|\rightarrow\rangle_{FF}$. Comparison of ED (green) and vCN (dashed red).

C. Error production rate

FIG. S5. Error production rate defined by eq. (S18) for uniform state $|\rightarrow\rangle_{FF}$ and non-uniform state $P|\rightarrow\rangle_{FF}$.

[S1] R. Verdel, P. Karpov, Y.-P. Huang, M. Schmitt and M. Heyl, to appear.
[S2] M. Schmitt and M. Heyl, SciPost Phys. 4, 013 (2018).
[S3] R. Kubo, Generalized Cumulant Expansion Method, J. Phys. Soc. Jpn. 17, 1100 (1962).
[S4] Quantum Monte Carlo approaches for correlated systems, F. Becca and S. Sorello (Cambridge University Press, 2017).