Emergent quantum mechanics at the boundary of a local classical lattice model

Kevin Slagle [1,2,3] and John Preskill[3,4]

[1]*Department of Electrical and Computer Engineering, Rice University, Houston, Texas 77005, USA*
[2]*Department of Physics, California Institute of Technology, Pasadena, California 91125, USA*
[3]*Institute for Quantum Information and Matter and Walter Burke Institute for Theoretical Physics,*
California Institute of Technology, Pasadena, California 91125, USA
[4]*AWS Center for Quantum Computing, Pasadena, California 91125, USA*

We formulate a model in which quantum mechanics emerges from classical mechanics. Given a local Hamiltonian H acting on n qubits, we define a local classical model with an additional spatial dimension whose boundary dynamics is approximately—but to arbitrary precision—described by Schrödinger's equation and H. The bulk consists of a lattice of classical bits that propagate towards the boundary through a circuit of stochastic matrices. The bits reaching the boundary are governed by a probability distribution whose deviation from the uniform distribution can be interpreted as the quantum-mechanical wave function. Bell nonlocality is achieved because information can move through the bulk much faster than the boundary speed of light. We analytically estimate how much the model deviates from quantum mechanics, and we validate these estimates using computer simulations.

I. INTRODUCTION

General. relativity and the Standard Model of particle physics are not exact descriptions of reality; rather they emerge as low-energy effective field theory descriptions of some underlying theory (e.g., string theory). A characteristic signature of emergence at low energy or long time and distance scales is that the resulting physics is typically well described by remarkably simple equations, which are often linear (e.g., the harmonic oscillator) or only consist of lowest-order terms in an effective Lagrangian.

In principle, it is possible that quantum mechanics is also only an approximate description of reality. Indeed, Schrödinger's equation is a simple linear differential equation, suggesting that it might arise as the leading approximation to a more complete model. The advent of quantum computing opens opportunities to probe and test quantum mechanics in an unprecedented regime. Much evidence indicates that if standard quantum theory is exactly correct, then the cost of simulating a quantum computation with a classical computer must grow exponentially with the size of the quantum computer [1]. To the extent possible, this extraordinary hypothesis about the quantum world should be tested in the laboratory. Indeed, if quantum theory actually emerges from an underlying classical model, then this exponential scaling must eventually fail for real devices. Therefore, aside from verifying Bell nonlocality [2] and studying the behavior of macroscopic superpositions [3], we should also conceive and perform experiments that characterize the computational power of nature [4–6].

Numerous experiments [2,7–13] and theoretical observations [14–19] significantly constrain, but do not completely rule out, possible deviations from standard quantum theory. For example, measurements of the anomalous magnetic dipole moment of the electron [20] agree with quantum predictions up to roughly ten digits of precision. However, these experiments, and most other current tests of quantum theory, probe properties of matter with relatively low computational complexity, and so might be insensitive to deviations from quantum theory that become evident only for more complex states [21]. Quantum theory also successfully predicts properties of ground states and of low-energy dynamics for many materials and molecules, but here too the detailed agreement between theory and experiment has mostly been limited to quantum states that are not profoundly entangled [22,23], and so the successful predictions do not rule out departures from quantum predictions for states of high complexity. To probe the high-complexity regime convincingly, highly excited matter should be carefully studied. It may also be necessary to measure many observables, since classically tractable models of thermalization [24,25] and emergent hydrodynamics [26–28] may suffice for explaining the observed data when only a few degrees of freedom are measured. In contrast to more conventional experimental tools, future quantum computers that prepare highly entangled states and perform intricate measurements will be well equipped for probing the behavior of matter in the regime far beyond the reach of efficient classical simulation [11–13].

Though models in which quantum dynamics emerges from underlying classical dynamics should be testable in the high-complexity regime, such tests need not be applicable to other proposed modifications of standard quantum theory. For example, in models with intrinsic wave function collapse [29], quantum error correction [30] might overcome the damaging effects of the intrinsic noise, restoring the full computational power of quantum theory. Furthermore, generic nonlinear corrections to Schrödinger's equation may well enhance rather than diminish the computational power of quantum systems

[31], while our goal is to explore whether nature could be computationally weaker, not stronger, than standard quantum theory predicts.

Such considerations motivate the quest for testable models in which quantum mechanics emerges from classical mechanics and for which deviations from standard quantum theory are detectable in the high-complexity regime [6,32–41]. In pursuit of this quest, we aim to construct a local classical lattice model that exhibits emergent quantum mechanics (EmQM) [42]. We define a local classical model to consist of a lattice, where the state at each lattice site is defined by a finite list of numbers (which does not grow with the system size), and the time evolution of each lattice site only depends on the state of nearby sites. The time evolution is allowed to be stochastic and either continuous or discrete. Cellular automata and local classical lattice Hamiltonians (and Lagrangians) are examples of local classical models.

We consider a model to exhibit EmQM if its slowly varying and long-distance physics is well described by Schrödinger's equation:

$$\partial_t \Psi = -iH\Psi. \tag{1}$$

Ψ is the wave function, which encodes the state of the system, while H is the Hamiltonian,[1] which defines the dynamics.

In a sense, the classical model is performing an approximate simulation of suitably encoded Schrödinger dynamics. We want more than just that—to the extent possible we want the classical dynamics to be a reasonable model of how nature might really behave. Typical classical simulation algorithms will not fit the bill, for example because the classical dynamics is spatially nonlocal or because the required number of local degrees of freedom increases exponentially with system size (e.g., tensor network methods [43,44] require an exponentially large bond dimension).[2] Drawing inspiration from neutral network algorithms [46–50], we seek models without such shortcomings.

We partially succeed in the following sense. Given any local Hamiltonian and initial value wave function for n qubits in D spatial dimensions, we can define a local classical lattice model in $D + 1$ spatial dimensions whose D-dimensional boundary dynamics can be well-approximated by Schrödinger's equation if the extra spatial dimension has a length S that is exponentially large in n, i.e., if $S \gg 2^n$.

We view this as only a partial success because if $S \gg 2^n$ does not hold, then the boundary dynamics instead obeys a Schrödinger's equation with a highly nonlocal Hamiltonian (i.e., its terms are geometrically nonlocal and also act on many qubits at once). In order to be a model of EmQM in our universe, we would like to view n as the number of (possibly Planckian-sized) qubits in our universe (with an effective low-energy Lagranian or Hamiltonian consistent with the standard model [51,52]). More conservatively, we would like to take

---

[1]In this work, we will consider only local Hamiltonians for a lattice of qubits, for which H is a Hermitian matrix and Ψ is a complex-valued vector. A quantum Hamiltonian is local if it is a sum of terms that each act only on nearby qubits.

[2]Spatial locality is also a challenge for tensor networks, although a spatially local algorithm has been derived in one dimension [45].

n to at least be as large as the number of qubits needed to describe a macroscopic region of space, e.g., certainly larger than Avogadro's number: $n > 10^{23}$. Thus, in order to be consistent with local quantum dynamics, the extra dimension would have to be tremendously long, e.g., $S \gg 2^{10^{23}}$. Future work is necessary to determine if it is possible to alleviate the $S \gg 2^n$ requirement. For example, one might instead demand that an EmQM model with fixed $S \gg 2^{\tilde{n}}$ be consistent with any experiment that probes only $\tilde{n}$ highly entangled qubits with high fidelity, e.g., the logical qubits in a quantum computer. This would be desirable because only $\tilde{n} \sim \log_2 S$ highly entangled qubits would be needed to experimentally test such a model of EmQM, which would be experimentally relevant in the near term if, e.g., $S \sim 2^{1000}$.

Various challenges had to be overcome while constructing our model of EmQM. In Sec. II we recount these challenges as guiding principles that intuitively motivate necessary ingredients for EmQM. In Sec. III we promote this intuition to an explicit model. In Sec. IV we estimate how much our EmQM model deviates from quantum mechanics, and we numerically validate these estimates in Sec. V. In Sec. VI we discuss possible experimental tests of EmQM models similar to the model we study. In Sec. VII we mention future directions, such as how our EmQM might be modified to possibly alleviate the $S \gg 2^n$ requirement.

II. INGREDIENTS FOR EmQM

A. Bell nonlocality from fast variables

Bell inequality [53,54] experiments have shown that the outcomes of spacelike separated quantum measurements are incompatible with local hidden variable theories unless information can travel faster than light. Therefore, we posit the existence of hidden *fast* degrees of freedom that move much faster than the speed of light and change much more rapidly than the wave function. Although we make no assumptions regarding local realism (another assumption used to derive Bell inequalities), we will be led to an EmQM model without local realism. For simplicity, we consider a wave function for qubits. Therefore, it is natural to take the fast variables to be classical bits.

We note that in order for a theory with faster-than-light degrees of freedom to be consistent with previous tests of Lorentz invariance [55], we likely also need to posit that the observed Lorentz invariance in our universe is emergent (rather than exact). See Appendix C for further discussion regarding the feasibility of this possibility.

B. Linearity from perturbative expansion

Another notable feature of Schrödinger's equation is that it is linear in the wave function, while classical systems generically exhibit nonlinear behavior. However, linearity is a generic result of leading-order perturbative expansions. For example, the linear harmonic oscillator describes small oscillations of a pendulum. The gravitational force in Newtonian gravity is a linear superposition of forces, which can be derived from general relativity in a certain limit where the gravitational force is weak. Even the training dynamics of wide neural networks (for which there are many neurons per

layer) can be reduced to a linear equation after a perturbative expansion about small deviations from the initial conditions [46,56], which was a significant inspiration for the EmQM model that we introduce.

We therefore posit that our model contains degrees of freedom that change so slowly with time that their dynamics can be treated in a linear approximation. If we attempt to follow the evolution for a very long time, terms nonlinear in the wave function may become significant, resulting in nonlinear corrections [57] to the emergent Schrödinger equation. These higher-order corrections might depend on details of the underlying classical dynamics, rather than being expressible in terms of the emergent wave function alone.

C. Wave function from probability vector

An EmQM model should also explain how the quantum wave function is related to the classical model. Specifying the wave function for n qubits requires 2^n numbers. This feature is reminiscent of classical probability distributions, which also require 2^n numbers for n bits. Therefore, we posit that the wave function is mathematically related to a probability vector $\boldsymbol{P}$.

In order to incorporate the previous two guiding principles, we further posit that $\boldsymbol{P}$ is a probability distribution for the fast degrees of freedom (i.e., the classical bits) and that $\boldsymbol{P}$ is determined by the slow degrees of freedom. We further assume that the wave function Ψ for n qubits describes perturbations from a uniform probability vector for n classical bits:

$$\boldsymbol{P} = \frac{\mathbf{1}}{N} + \epsilon_\Psi \Psi, \tag{2}$$

where ϵ_Ψ is perturbatively small. $\boldsymbol{P}$ is a vector of probabilities for the fast degrees of freedom. $\mathbf{1}$ is a vector of $N = 2^n$ ones so that $\frac{\mathbf{1}}{N}$ is a uniform probability vector for n bits.

We will be content to describe the evolution of the emergent wave function in our EmQM model and will not discuss measurement as a separate phenomenon. To accommodate measurements, one could adopt the Everett interpretation [58] by including the observer and measurement apparatus as part of the physical system described by the wave function. See Appendix A for more details.

Constraints

In order for $\boldsymbol{P}$ to be a valid probability vector, Ψ must be a real vector with elements that sum to zero:

$$\sum_i \Psi_i = 0. \tag{3}$$

In order to preserve these constraints, the Hamiltonian H in Schrödinger's equation (1) must be an imaginary-valued and antisymmetric matrix with rows and columns that sum to zero:

$$\sum_i H_{ij} = 0,$$
$$\sum_j H_{ij} = 0. \tag{4}$$

In order to obtain a local EmQM model, we also require that every term of the Hamiltonian is local and satisfies Eq. (4).

In Appendix B, we show that these constraints on H and Ψ do not result in any significant loss of generality. In particular, given any Hamiltonian H and wave function $\Psi(t)$ that satisfy Schrödinger's equation, we find a linear mapping to a dual Hamiltonian $\widetilde{H}$ and wave function $\widetilde{\Psi}(t)$ that satisfy Schrödinger's equation and the above constraints. Furthermore, if H is local, then $\widetilde{H}$ can also be chosen to be local. Every term of $\widetilde{H}$ will also satisfy Eq. (4).

D. Quantum complexity from a large extra dimension

Simulating Schrödinger's equation with a classical computer generically requires CPU time that increases exponentially with system size. Therefore, simulating an underlying classical EmQM model should also have a high cost. To ensure that the EmQM model is costly to simulate, we posit a large extra spatial dimension of length S. In effect, this large dimension enables the EmQM model to describe quantum mechanics in an exponentially large Hilbert space.

In our model, the stochastic classical bits are fast in the sense that they are frequently sampled. But the emergent wave function is related to the probability distribution $\boldsymbol{P}$ from which these bits are sampled, where $\boldsymbol{P}$ itself evolves quite slowly in comparison. A general probability distribution on n bits is parameterized by $2^n - 1$ nonnegative real numbers. If we want to allow general n-qubit quantum pure states in our EmQM model, then many parameters must be needed to specify the stochastic process from which $\boldsymbol{P}$ arises. For this reason we assume that $\boldsymbol{P}$ is obtained by composing stochastic matrices in a very deep circuit. The number of parameters needed to parametrize the circuit is linear in its depth S, which we therefore assume to be exponential in n. We envision this circuit as extending into an auxiliary spatial dimension, not to be confused with the spatial dimensions of the emergent quantum system.

Consider an EmQM model consisting of an $S + 1$ by n grid of bits $a_{s,x}$, which are depicted as red dots in Fig. 1. For simplicity of exposition, we focus on the EmQM of a one-dimensional chain of n qubits. Generalizations to higher spatial dimensions are straightforward. Let $\boldsymbol{P}$ be the probability distribution for the n bits at the $s = S$ boundary (circled in red) of the extra spatial dimension. We then suppose that at the $s = 0$ boundary, the n bits are generated uniformly at random, and that the bulk dynamics interpolate between the uniform distribution $\frac{1}{N}$ and $\boldsymbol{P} = \frac{1}{N} + \epsilon_\Psi \Psi$.

Next, we take inspiration from unitary circuits, which can generate entangled wave functions from direct product states by repeatedly acting on pairs of qubits with unitary matrices. Since the EmQM model involves classical bits instead of qubits, we instead consider a circuit of stochastic matrices $M_{s,x}$ (yellow in Fig. 1). A stochastic matrix is a matrix with columns that are probability vectors, i.e., vectors with positive entries that sum to one. Therefore, given two input bits, a 4×4 stochastic matrix maps those bits to a probability distribution, from which a new pair of bits can be sampled. A 4×4 stochastic matrix can also be used to linearly map a probability vector for two bits to a new probability vector. The EmQM model utilizes a circuit of stochastic matrices to sample bits from the probability vector $\boldsymbol{P}$. We emphasize that
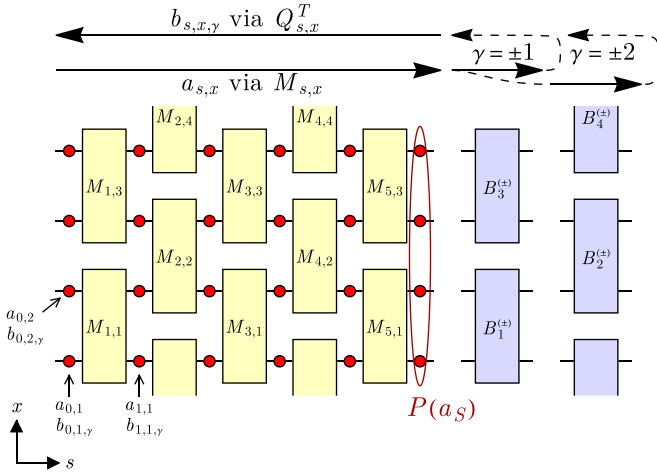
FIG. 1. Square lattice picture of our model (with $S = 5$ and $n = 4$), which exhibits EmQM on a one-dimensional boundary (circled in red) of a two-dimensional bulk. Classical bits $a_{s,x}$ (red dots) propagate forward through a slowly time-evolving circuit of stochastic matrices $M_{s,x}$ (yellow), and classical bits $b_{s,x,\gamma}$ propagate backwards through a time-independent circuit of permutation matrices $Q_{s,x}^T$. The emergent Schrödinger equation describes the time evolution of the probability distribution $P(a_S)$ governing the boundary bits. The back-propagating bits determine how the stochastic matrices are updated in each time step. See Sec. III A for an overview.

P is not a physical degree of freedom; P is only implicitly defined by the circuit of stochastic matrices.[3]

The stochastic matrices $M_{s,x}$ vary slowly with the time step τ, and can be decomposed as

$$M_{s,x}(\tau) = Q_{s,x} + m_{s,x}(\tau), \tag{5}$$

where $Q_{s,x}$ is time independent and $m_{s,x}(\tau)$ is a time-dependent perturbation. We take the $Q_{s,x}$ to be permutation matrices, which have the useful property that their inverse is also a stochastic matrix.[4] We will assume that at all times, the perturbation $m_{s,x}(\tau)$ is sufficiently small that we can accurately account for the evolution of P by expanding to linear order. As a result, the time evolution of the emergent wave function will also be described by a linear equation.

E. Unitarity from destructive interference

A final challenge is to obtain dynamics for P such that Ψ in Eq. (2) undergoes a unitary evolution described by Schrödinger's equation.[5] One route to realizing unitarity is

---

[3]The classical bits and stochastic matrices are ontic in our model. There are thus an infinite number of possible physical states, as implied by Hardy's excess baggage theorem [59]. However, P and the wave function are both fully determined by the stochastic matrices, which implies that our model is ψ-ontic [60] in the sense that distinct wave functions always correspond to distinct physical states (of classical bits and stochastic matrices).

[4]More generic choices for $Q_{s,x}$ could be a useful direction for future work, which we briefly discuss in Sec. VII A.

[5]Unitary dynamics implies many other useful properties. For example, Tsirelson's bound is an upper bound for how much quantum

to suppose that after the forward-propagating bits $a_{s,x}$ reach the $s = S$ boundary, the bits are transformed by a shallow stochastic circuit B that encodes a small unitary time evolution generated by the Hamiltonian H. The resulting bits could then back-propagate through the circuit (via $Q_{s,x}^T$) while dictating how the stochastic matrices slowly evolve such that P undergoes the desired dynamics.

But how could a stochastic circuit encode a time evolution by a generic imaginary-valued Hamiltonian satisfying Eq. (4)? One possibility is that a pair of stochastic circuits $B^{(\pm)}$ outputs two sets of bits with probability vectors $P_{\pm}$ that "destructively interfere" with each other:

$$P_+ - P_- = (-iH\delta_t)P. \tag{6}$$

This could be achieved by stochastic circuits $B^{(\pm)}$ defined such that

$$B^{(+)} - B^{(-)} = -iH\delta_t. \tag{7}$$

Such a decomposition is always possible for sufficiently small δ_t. For example, if
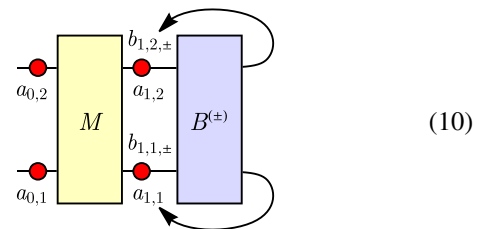
$$-iH = \begin{pmatrix} 0 & -1 & +1 & 0 \\ +1 & 0 & -1 & 0 \\ -1 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{8}$$

for $n = 2$ qubits, then we can choose

$$B^{(+)} = \begin{pmatrix} 1-\delta_t & 0 & \delta_t & 0 \\ \delta_t & 1-\delta_t & 0 & 0 \\ 0 & \delta_t & 1-\delta_t & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$
$$B^{(-)} = \begin{pmatrix} 1-\delta_t & \delta_t & 0 & 0 \\ 0 & 1-\delta_t & \delta_t & 0 \\ \delta_t & 0 & 1-\delta_t & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{9}$$

To illustrate how this works in a simple setting, let us continue this $n = 2$ example and further suppose that $S = 1$ such that there is only a single stochastic matrix M, as depicted below:

$$(10)$$

For each discrete time step, all of the bits $a_{s,x}$ and $b_{s,x,\pm}$ are updated. The two input bits $a_{0,1}$ and $a_{0,2}$ are chosen uniformly at random. The boundary bits at $s = 1$ are randomly chosen from the conditional probability distributions $p(a_1|a_0) = M(a_1, a_0)$

---

theory can violate Bell's inequality [54]. Tsirelson's bound is saturated by Schrödinger's equation, but Tsirelson's bound is exceeded by some alternatives of quantum theory [61]. If Ψ in our model obeys a unitary evolution corresponding to a local Hamiltonian, then in addition to violating Bell's inequality, our model would saturate Tsirelson's bound (in agreement with quantum theory).

and $p(b_\pm|a_1) = B^{(\pm)}(b_\pm, a_1)$, where (for now) a_s and $b_\pm$ denote pairs of bits $(a_{s,1}, a_{s,2})$ and $(b_{1,1,\pm}, b_{1,2,\pm})$, which are used to index the 4×4 matrices M and $B^{(\pm)}$.

In order to get the desired time evolution for $\boldsymbol{P}$, for each discrete time step, m in $M = Q + m$ is updated according to

$$m' = m + \Delta_m(\hat{\boldsymbol{b}}_+ - \hat{\boldsymbol{b}}_-) \otimes \hat{\boldsymbol{e}}, \tag{11}$$

where Δ_m is a small constant. $\hat{\boldsymbol{b}}_+$ denotes the basis 4-vector indexed by the two bits $b_{1,1,+}$ and $b_{1,2,+}$, and similar for $\hat{\boldsymbol{b}}_-$. For example, $\hat{\boldsymbol{b}}_+ = (0, 1, 0, 0)$ if $b_{1,1,+} = 0$ and $b_{1,2,+} = 1$. $\hat{\boldsymbol{e}}$ is chosen uniformly at random from the four basis 4-vectors [(1,0,0,0), (0,1,0,0), etc.] with the constraint that $M = Q + m$ remains nonnegative. Such a choice always exists as long as the elements of m (which are assumed to be small) remain smaller than 1. The choice of Q does not play an important role in this $S = 1$ example.

$\boldsymbol{P}$ is the probability distribution that governs the sampled boundary bits $a_{1,1}$ and $a_{1,2}$. In addition, $\boldsymbol{P}$ itself has statistical fluctuations, because in each time step, the stochastic matrix M is updated according to Eq. (11), where $\hat{\boldsymbol{b}}_+$ and $\hat{\boldsymbol{b}}_-$ are also stochastic variables. To see that Eq. (11) results in an emergent Schrödinger equation, we first calculate the expectation values (denoted by a bar) of the boundary bits:

$$\boldsymbol{P} = \overline{\hat{\boldsymbol{a}}_1} = M\frac{\mathbf{1}_4}{4}, \tag{12}$$

$$\overline{\hat{\boldsymbol{b}}_\pm} = B^{(\pm)}\boldsymbol{P}, \tag{13}$$

where $\mathbf{1}_4/4$ denotes a uniform probability vector with four elements. Equation (7) then implies that

$$\overline{\hat{\boldsymbol{b}}_+ - \hat{\boldsymbol{b}}_-} = -i\delta_t H \boldsymbol{P}. \tag{14}$$

Thus, the average change in $\boldsymbol{P}$ after one time step evolves according to a discrete Schrödinger's equation:

$$\begin{aligned} \overline{\boldsymbol{P}' - \boldsymbol{P}} &= \overline{(m' - m)}\frac{\mathbf{1}_4}{4} \\ &= \frac{\Delta_m}{4}\overline{(\hat{\boldsymbol{b}}_+ - \hat{\boldsymbol{b}}_-)} \\ &= -i\Delta_t H \boldsymbol{P}, \end{aligned} \tag{15}$$

where $\Delta_t = \frac{1}{4}\delta_t \Delta_m$. The above three equalities follow from Eqs. (12), (11), and (14), respectively. In the second equality, we see that the random choice of $\hat{\boldsymbol{e}}$ in Eq. (11) does not matter because $\hat{\boldsymbol{e}}$ is multiplied by $\mathbf{1}_4$ in the first line. Since $\boldsymbol{P}$ and Ψ are linearly related by Eq. (2), Ψ obeys Schrödinger's equation on average. Statistical fluctuations about this average are negligible in the small Δ_m limit.

To quantify statistical fluctuations, let τ denote the discrete time step of $\boldsymbol{P}^{(\tau)}$. After τ steps, the elapsed time in the emergent quantum mechanics is $t \approx \Delta_t \tau$. Statistical fluctuations of $\boldsymbol{P}^{(\tau)}$ grow as[6] $O(\Delta_m\sqrt{\delta_t\tau}) = O(\sqrt{\Delta_m t})$. Therefore, statistical fluctuations become arbitrarily small as Δ_m decreases.

F. More qubits

Finally, we must scale up the previous $S = 1$ and $n = 2$ example to large S and n, as depicted in Fig. 1. We will now assume that the time-independent permutation matrices $Q_{s,x}$ in Eq. (5) are chosen uniformly at random and independently for each s and x. These random permutation matrices play an important role as they randomize the subspace in which each $m_{s,x}$ affects $\boldsymbol{P}$. Therefore when $Sn \gg N = 2^n$, the span of these subspaces covers all N dimensions of $\boldsymbol{P}$.

When $S > 1$, the bits $b_{S,x,\gamma}$ at the boundary will have to back-propagate through the circuit before affecting the time evolution of $M_{s,x}$. In order to back-propagate the maximal amount of information from the boundary to the bulk, it would be ideal to use the inverse of the stochastic matrices $M_{s,x}^{-1}$. But this is not possible since these inverses are generically not stochastic matrices. However, since $m_{s,x}$ is assumed to be small, $M_{s,x}^{-1} \approx Q_{s,x}^T$ and therefore $Q_{s,x}^T$ can be used to deterministically back-propagate the bits.[7]

Finally, we must split $B^{(\pm)}$ into local pieces, as depicted in Fig. 1. To be concrete, we assume that the Hamiltonian $H = \sum_x H_x$ is local such that each H_x only acts on two neighboring qubits. Generalizing Eq. (7), $B_x^{(\pm)}$ are chosen such that

$$B_x^{(+)} - B_x^{(-)} = -iH_x\delta_t \tag{16}$$

with the constraint that the stochastic matrix

$$B_x^{(\pm)} = \mathbb{1}_4 + O(\delta_t) \tag{17}$$

is the 4×4 identity matrix $\mathbb{1}_4$ up to order δ_t corrections. We then define four different depth-1 stochastic circuits:

$$\begin{aligned} B^{(\pm1)} &= \otimes_{\text{odd }x}B_x^{(\pm)}, \\ B^{(\pm2)} &= \otimes_{\text{even }x}B_x^{(\pm)}, \end{aligned} \tag{18}$$

which are now guaranteed to satisfy a generalization of Eq. (7):

$$+B^{(+1)} - B^{(-1)} + B^{(+2)} - B^{(-2)} = -iH\delta_t + O(\delta_t^2). \tag{19}$$

Since there are now four different flavors of $B^{(\gamma)}$, we also take four different flavors of back-propagating bits $b_{s,x,\gamma}$ with $\gamma = \pm1$ and $\gamma = \pm2$. With these ingredients combined, we arrive at a local model of emergent quantum mechanics.

III. EmQM MODEL

A. Overview

In summary, we introduce a two-dimensional local classical model for which a wave function Ψ of qubits is encoded in a probability distribution $\boldsymbol{P} = \frac{1}{N} + \epsilon_\Psi \Psi$ [Eq. (2)] for n classical bits on a one-dimensional boundary of the model.

---

[6] $\delta_t\tau$ is roughly the number of time steps for which $\hat{\boldsymbol{b}}_+ - \hat{\boldsymbol{b}}_- \neq 0$. Each of these time steps changes m by Δ_m.

[7] $M_{s,x}^T$ can not be used for back-propagation since $M_{s,x}^T$ is not guaranteed to be a stochastic matrix. We can not simply require $M_{s,x}^T$ to be a stochastic matrix in our model (which would imply that $M_{s,x}$ is doubly stochastic) since we require that $M_{s,x}$ maps a uniform probability distribution to a different distribution. But doubly stochastic matrices always map the uniform distribution to a uniform distribution since the rows of a doubly stochastic matrix must sum to one (by definition).

Generalizing to higher dimensions or Z_k qudits (rather than Z_2 qubits) is straightforward. For each discrete time step, a bit string a_S (e.g., $a_S = 011010$) is generated with probability $P(a_S)$ at the $s = S$ boundary. Our model exhibits EmQM in the sense that $\boldsymbol{P}$ [i.e., $P(a_S)$ viewed as a vector with $N = 2^n$ components] evolves according to

$$\partial_t \boldsymbol{P} \approx -iH\boldsymbol{P}, \tag{20}$$

where H is imaginary valued and obeys Eq. (4). Schrödinger's equation (1) for Ψ then follows from $\boldsymbol{P} = \frac{1}{N} + \epsilon_\Psi \Psi$ [Eq. (2)]. We will derive Eq. (20) from a well-controlled perturbation theory, which we verify numerically. Furthermore, the approximate equality in Eq. (20) becomes exact in a well-defined limit.
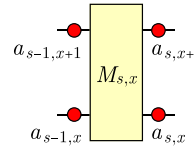
Figure 1 summarizes the basic structure of our EmQM model. The model consists of a square lattice of classical bits $a_{s,x}$ and $b_{s,x,\gamma}$ (red dots) that propagate through a brick circuit of slowly varying stochastic matrices $M_{s,x}$ (yellow). For each discrete time step, pairs of classical forward-propagating bits ($a_{s,x}$ and $a_{s,x+1}$) are randomly sampled (21) from probability distributions conditioned on the pair of bits to the left ($a_{s-1,x}$ and $a_{s-1,x-1}$). The conditional probabilities are encoded in the stochastic matrices $M_{s,x}$, which are perturbatively close to time-independent permutation matrices $Q_{s,x} \approx M_{s,x}$. Also at each time step, pairs of backward-propagating bits ($b_{s,x,\gamma}$ and $b_{s,x+1,\gamma}$) are deterministically replaced (30) by the permutation $Q_{s,x}^T$ of bits to their right ($b_{s+1,x}$ and $b_{s+1,x+1}$). The bits $a_{0,x}$ at the left $s = 0$ boundary are uniformly initialized at random. The bits $b_{S,x,\gamma}$ at the right $s = S$ boundary result from applying (29) a shallow stochastic circuit (18) to the bits $a_{S,x}$ at $s = S$. The shallow circuit is related (19) to the Hamiltonian and consists of a layer of time-independent stochastic matrices $B_x^{(\pm)}$ (blue). The bits $a_{S,x}$ on the $s = S$ boundary (circled in red) follow a probability distribution $\boldsymbol{P}$ defined by a tensor network product (23) of the matrices $M_{s,x}$. The time evolution of $M_{s,x}$ depends (31) on back-propagating bits $b_{s,x,\gamma}$, which effectively back-propagate the result of a small Hamiltonian time evolution on $\boldsymbol{P}$, such that $\boldsymbol{P}$ and Ψ approximately obey Schrödinger's equation. See Table I for a notational reference.

B. Stochastic circuit

We now define and study the EmQM model in greater detail. Each lattice site hosts a classical bit $a_{s,x} = 0, 1$, where $s = 0, 1, \ldots, S$, and $x = 1, \ldots, n$ index the different lattice sites. s is the coordinate for an extra dimension, while x is the coordinate for the spatial dimension on which the emergent qubits live. Our model evolves in discrete time steps indexed by integer-valued τ, while t is reserved for the emergent time variable of the emergent quantum mechanics. $a_{s,x}^{(\tau)}$ denotes the value of the bit $a_{s,x}$ at time step τ, and similar for other time-dependent variables. In contexts where the time step is not important, we often omit the (τ) superscript to avoid clutter.

To avoid clutter, we denote a pair of bits $(a_{s,x}, a_{s,x+1})$ as $a_{s,\mathbb{x}}$, where $\mathbb{x}$ is shorthand for $(x, x+1)$. For each time step τ, the classical bits $a_{s,x}^{(\tau-1)}$ with $1 \leqslant s \leqslant S$ are stochastically updated to $a_{s,x}^{(\tau)}$ with conditional probabilities that are conditioned on the bits $a_{s-1,\mathbb{x}}^{(\tau-1)}$. These conditional probabilities are

given by time-dependent stochastic matrices $M_{s,x}$:

$$p(a_{s,\mathbb{x}}^{(\tau)}|a_{s-1,\mathbb{x}}^{(\tau-1)}) = M_{s,x}^{(\tau-1)}(a_{s,\mathbb{x}}^{(\tau)}, a_{s-1,\mathbb{x}}^{(\tau-1)})$$

$$\tag{21}$$

For each (s, x) with $s = 1, 2, \ldots, S$ and $s - x$ even, $M_{s,x}$ is a 4×4 stochastic matrix. That is, the columns of $M_{s,x}$ are probability vectors; i.e., $M_{s,x}$ has positive elements and columns that sum to 1. We index the 4×4 matrices using pairs of bits, such that for fixed $a_{s-1,\mathbb{x}}^{(\tau-1)}$, $M_{s,x}^{(\tau-1)}(a_{s,\mathbb{x}}^{(\tau)}, a_{s-1,\mathbb{x}}^{(\tau-1)})$ is a probability distribution for $a_{s,\mathbb{x}}^{(\tau)}$. The bits along the $s = 0$ line at the beginning of the circuit are randomly sampled with equal probability:

$$p(a_{0,x}) = \tfrac{1}{2}. \tag{22}$$

Let $a_s \equiv (a_{s,1}, a_{s,2}, \ldots, a_{s,n})$ denote the string of bits along a column of fixed s. For each s, let $\boldsymbol{P}_s$ denote the vector of probabilities for the different bit strings a_s. Then $\boldsymbol{P}_s^{(\tau)}$ can be expressed recursively as

$$\boldsymbol{P}_s^{(\tau)} = M_s^{(\tau-1)} \boldsymbol{P}_{s-1}^{(\tau-1)},$$
$$\boldsymbol{P}_0^{(\tau)} = \frac{\mathbf{1}}{N}, \tag{23}$$

where M_s denotes the Kronecker product of stochastic matrices with fixed s:

$$M_s^{(\tau)} = \otimes_x^{\overset{\text{even}}{s-x}} M_{s,x}^{(\tau)}. \tag{24}$$

The notation on the right-hand side means we take the Kronecker product of all $M_{s,x}^{(\tau)}$ for even x when s is even (or for odd x when s is odd), producing the brickwork circuit shown in Fig. 1. The bit string probabilities at the end of the circuit are given by $\boldsymbol{P} = \boldsymbol{P}_S$ (circled in red in Fig. 1).

C. Perturbative expansion

To gain analytical tractability, we assume that the stochastic matrices are very close to permutation matrices:

$$M_{s,x}^{(\tau)} = Q_{s,x} + m_{s,x}^{(\tau)}. \tag{25}$$

Each $Q_{s,x}$ is a randomly chosen (for each s and x) but time-independent 4×4 permutation matrix, while $m_{s,x}^{(\tau)}$ is a small time-dependent perturbation. A permutation matrix is a stochastic matrix where all elements are either 0 or 1. A matrix is a permutation matrix if and only if it is both stochastic and orthogonal. The dynamics of $m_{s,x}$ will be constrained such that $M_{s,x}$ remains a stochastic matrix (with nonnegative components).

Since they are orthogonal, the permutation matrices have the effect of a basis transformation. Similar to Eq. (24), let Q_s denote the Kronecker product of permutation matrices with fixed s:

$$Q_s = \overset{\overset{\text{even}}{s-x}}{\underset{x}{\otimes}} Q_{s,x}. \tag{26}$$

Let $Q_{S \leftarrow s}$ denote the product

$$Q_{S \leftarrow s} = Q_S Q_{S-1} \cdots Q_{s+1} \tag{27}$$

TABLE I. Notation used throughout the main text (with the exception of Sec. II E, which uses some simplified notations). Notations are grouped as follows: model parameters, dependent variables, dynamical variables, randomly initialized constants, and other constants.

Variable	Definition	Reference
n	Number of bits output by circuit at each time step	Sec. III A
$N = 2^n$	Number of bit strings for n bits; $x = 1, 2, \ldots, n$	Sec. III A
S	Length of extra dimension; $s = 1, 2, \ldots, S$	Sec. III A
δ_t	Raw time step (not to be confused with Δ_t)	Sec. II E
m_0	Approximate Hilbert-Schmidt norm of each $m_{s,x}$	Eq. (28)
Δ_m	Step size for $m_{s,x}$ for each time step	Eq. (31)
ϵ_0	Small parameter used to parametrize the above four constants	Eq. (79)
H, H_x	Effective local Hamiltonian $H = \sum_x H_x$	Sec. III A
G, G_x	Real-valued: $G = -iH$, $G_x = -iH_x$	Eq. (36)
Δ_t	Effective EmQM time step: $t \approx \Delta_t \tau$	Eq. (50)
ϵ_Ψ	Small constant relating $\boldsymbol{P} - \frac{1}{N} \approx \epsilon_\Psi \Psi$	Eqs. (2) and (54)
$B_x^{(\pm)}$	4×4 stochastic matrix used to sample $b_{s,x,\pm\gamma}$	Eqs. (16), (29), and (37)
$B^{(\gamma)}$	Depth-1 stochastic circuit; related to Hamiltonian	Eqs. (18) and (19)
Δ_P	Rough 2-norm of $\overline{\boldsymbol{P}^{(\tau)} - \boldsymbol{P}^{(\tau-1)}}$	Eq. (53)
$P(a_S)$	Probability the circuit outputs bit string a_S	Sec. III A
$\boldsymbol{P} = \boldsymbol{P}_S$	Expectation value of $\hat{\boldsymbol{a}}_S$, or $P(a_S)$ viewed as a probability vector	Eqs. (23), and (28)
$\varepsilon(t)$	EmQM deviation $\|\Psi(t) - \Psi_{QM}(t)\|$ from quantum mechanics (QM)	Eqs. (52) and (75)
$\varepsilon_m(t), \varepsilon_t(t)$	Deviations due to finite m_0 and δ_t	Eqs. (63) and (65)
$\varepsilon_S(t), \varepsilon_{\text{stat}}(t)$	Deviations due to finite S and statistical fluctuations	Eqs. (70) and (74)
τ	Integer-valued time step	Sec. III B
$a_{s,x}$	Forward-propagating stochastic bit	Sec. III A
$a_{s,\mathbb{x}}$	Pair of bits $(a_{s,x}, a_{s,x+1})$	Above Eq. (21)
a_s	Bit string $(a_{s,1}, a_{s,2}, \ldots, a_{s,n})$	Below Eq. (22)
$\hat{\boldsymbol{a}}_s$	Length-n basis vector indexed by the bit string a_s	Below Eq. (39)
$b_{s,x,\gamma}$	Backward-propagating stochastic bit	Sec. III A
$b_{s,\mathbb{x},\gamma}$	Pair of bits $(b_{s,x,\gamma}, a_{s,x+1,\gamma})$	Above Eq. (29)
$\hat{\boldsymbol{b}}_{s,\mathbb{x},\gamma}$	(1,0,0,0), (0,1,0,0), (0,0,1,0), or (0,0,0,1) when $b_{s,\mathbb{x},\gamma} = 00, 01, 10,$ or 11	Below Eq. (34)
$b_{s,\gamma}$	Bit string $(b_{s,1,\gamma}, b_{s,2,\gamma}, \ldots, b_{s,n,\gamma})$	Below Eq. (34)
$\hat{\boldsymbol{b}}_{s,\gamma}$	Length-n basis vector indexed by the bit string $b_{s,\gamma}$	Below Eq. (34)
$M_{s,x}$	4×4 stochastic matrices, which define the classical circuit	Sec. III A
M_s	Kronecker product of $M_{s,x}$ with fixed s	Eq. (24)
$M_{S \leftarrow s}$	$= M_S M_{S-1} \cdots M_{s+1}$	Eq. (56)
$m_{s,x}$	Perturbations to $M_{s,x} = Q_{s,x} + m_{s,x}$	Eq. (25)
$Q_{s,x}$	Time-independent 4×4 permutation matrices	Below Eq. (25)
Q_s	Kronecker product of $Q_{s,x}$ with fixed s	Eq. (26)
$Q_{S \leftarrow s}$	$= Q_S Q_{S-1} \cdots Q_{s+1}$	Eq. (27)
$\boldsymbol{1}$ or $\boldsymbol{1}_n$	Vector of n ones	
$\mathbb{1}$ or $\mathbb{1}_n$	$n \times n$ identity matrix	
$\mathcal{P}_x$	Projects out all bits except x and $x+1$	Eq. (35)
$\mathcal{W}$	Sum of conjugated projectors $\mathcal{P}_x$	Eqs. (66) and (44)

such that it encodes the change of basis from the column of bits at s to the end at $s = S$.

Expanding $\boldsymbol{P}$ to first order in the perturbations $m_{s,x}$ results in

$$\boldsymbol{P}^{(\tau)} = \frac{1}{N} + \sum_{s=1}^{S} \sum_{x}^{\overset{\text{even}}{s-x}} Q_{S \leftarrow s}\, m_{s,x}^{(\tau_s)}\, \frac{1}{N} + O(m_0^2)$$

$$(28)$$

where $\tau_s = \tau - 1 - (S - s)$ results from a time delay, and $O(m_0^2)$ denotes terms that are quadratic in $m_{s,x}$. For simplicity, we assume that each $m_{s,x}$ has Hilbert-Schmidt norm roughly equal to m_0. In Sec. IV we will address more precisely the effects of the nonlinear $O(m_0^2)$ terms and other approximations that we make.

As elucidated in the above picture, $m_{s,x}^{(\tau_s)}$ affects the probabilities for two bits at s and $(x, x+1)$; these probabilities span a four-dimensional subspace of the N-dimensional vector space of probabilities $\boldsymbol{P}_s$ (after averaging over the uniform distribution of input bits). This subspace is then scrambled into a different basis by the product $Q_{S \leftarrow s}$ of many random (but time-independent) permutation matrices. If $nS \gg N = 2^n$, then the linear combination of these subspaces will span

the entire N-dimensional vector space. Thus, the sum of contributions from all $m_{s,x}$ can encode any wave function Ψ [that is real valued and satisfies Eq. (3)], where $\boldsymbol{P} = \frac{1}{N} + \epsilon_\Psi \Psi$ [Eq. (2)].

D. Time evolution

We now want to give the perturbations $m_{s,x}$ a time evolution that leads to an emergent Schrödinger equation; i.e., $\boldsymbol{P}^{(\tau)} - \boldsymbol{P}^{(\tau-1)} \propto -iHP$. To accomplish this, we introduce a set of backward-propagating bits $b_{s,x,\gamma}$ with $\gamma = \pm 1, \pm 2$. At the $s = S$ boundary, the pair of bits $b_{S,\mathbb{x},\pm\gamma}^{(\tau)}$ are randomly sampled with probabilities that are conditioned on the bits $a_{S,\mathbb{x}}^{(\tau-1)}$. These conditional probabilities are encoded in time-independent 4×4 stochastic matrices $B_x^{(\pm)}$ (blue in Fig. 1):

$$p\big(b_{S,\mathbb{x},\pm\gamma}^{(\tau)} \big| a_{S,\mathbb{x}}^{(\tau-1)}\big) = B_x^{(\pm)}\big(b_{S,\mathbb{x},\pm\gamma}^{(\tau)}, a_{S,\mathbb{x}}^{(\tau-1)}\big), \quad (29)$$

where $\gamma = 1$ if x is odd, else $\gamma = 2$.

The bits $b_{s,\mathbb{x},\gamma}$ deterministically propagate backwards through the circuit via the transposed permutation matrices $Q_{s,x}^T$:

$$\hat{\boldsymbol{b}}_{s,\mathbb{x},\gamma}^{(\tau)} = Q_{s,x}^T \hat{\boldsymbol{b}}_{s+1,\mathbb{x},\gamma}^{(\tau-1)}. \quad (30)$$

The hat over $\hat{\boldsymbol{b}}_{s,\mathbb{x},\gamma}^{(\tau)}$ denotes the basis 4-vector indexed by the two bits $b_{s,\mathbb{x},\gamma}^{(\tau)}$; i.e., $\hat{\boldsymbol{b}}_{s,\mathbb{x},\gamma}^{(\tau)} = (1,0,0,0)$ if $b_{s,\mathbb{x},\gamma}^{(\tau)} = 00$, and $\hat{\boldsymbol{b}}_{s,\mathbb{x},\gamma}^{(\tau)} = (0,1,0,0)$ if $b_{s,\mathbb{x},\gamma}^{(\tau)} = 01$, etc.

The time evolution of the perturbations $m_{s,x}$ is stochastic and depends on the back-propagating bits as follows [extending Eq. (11)]:

$$m_{s,x}^{(\tau)} = m_{s,x}^{(\tau-1)} + \Delta_{\mathrm{m}} \sum_{\gamma=1,2} \big(\hat{\boldsymbol{b}}_{s,\mathbb{x},+\gamma}^{(\tau)} - \hat{\boldsymbol{b}}_{s,\mathbb{x},-\gamma}^{(\tau)}\big) \otimes \hat{\boldsymbol{e}}_{s-1,x,\gamma}^{(\tau)}, \quad (31)$$

where Δ_{m} is a small positive constant. Similar to $\hat{\boldsymbol{b}}_{s,\mathbb{x},\gamma}^{(\tau)}$ (defined in the previous paragraph), $\hat{\boldsymbol{e}}_{s-1,x,\gamma}^{(\tau)}$ is also one of the four basis 4-vectors, except it is chosen uniformly at random from the set of basis 4-vectors that keep $M_{s,x}^{(\tau)} = Q_{s,x} + m_{s,x}^{(\tau)}$ nonnegative. Such a choice always exists as long as the elements of $m_{s,x}^{(\tau)}$ (which are assumed to be small) remain smaller than 1. The random choice of $\hat{\boldsymbol{e}}_{s-1,x,\gamma}^{(\tau)}$ has little effect since $m_{s,x}^{(\tau)}$ enters Eq. (28) after right-multiplication by $\mathbf{1}$.

The following quantity will play an important role:

$$\big(m_{s,x}^{(\tau)} - m_{s,x}^{(\tau-1)}\big)\frac{\mathbf{1}_4}{4} = \frac{\Delta_{\mathrm{m}}}{4} \sum_{\gamma=1,2} \big(\hat{\boldsymbol{b}}_{s,\mathbb{x},+\gamma}^{(\tau)} - \hat{\boldsymbol{b}}_{s,\mathbb{x},-\gamma}^{(\tau)}\big), \quad (32)$$

where $\frac{\mathbf{1}_4}{4}$ is a length-4 uniform probability vector. When $m_{s,x}^{(\tau)} - m_{s,x}^{(\tau-1)}$ acts on two bits of a uniform probability vector of length $N = 2^n$, the result is similar:

$$\big(m_{s,x}^{(\tau)} - m_{s,x}^{(\tau-1)}\big)\frac{\mathbf{1}_N}{N}$$
$$= \frac{\mathbf{1}_{2^{x-1}}}{2^{x-1}} \otimes \left[\big(m_{s,x}^{(\tau)} - m_{s,x}^{(\tau-1)}\big)\frac{\mathbf{1}_4}{4}\right] \otimes \frac{\mathbf{1}_{2^{n-x-1}}}{2^{n-x-1}} \quad (33)$$

$$= \frac{\Delta_{\mathrm{m}}}{4} \sum_{\gamma=1,2} \mathcal{P}_x\big(\hat{\boldsymbol{b}}_{s,+\gamma}^{(\tau)} - \hat{\boldsymbol{b}}_{s,-\gamma}^{(\tau)}\big). \quad (34)$$

In the last line, $\hat{\boldsymbol{b}}_{s,\gamma}$ denotes a length-n basis vector indexed by the bit string $b_{s,\gamma} = (b_{s,1,\gamma}, b_{s,2,\gamma}, \ldots, b_{s,n,\gamma})$. For example if $n = 3$, then $\hat{\boldsymbol{b}}_{s,\gamma} = (1,0,0,0,0,0,0,0)$ if $b_{s,\gamma} = 000$, and $\hat{\boldsymbol{b}}_{s,\gamma} = (0,1,0,0,0,0,0,0)$ if $b_{s,\gamma} = 001$, etc. $m_{s,x}^{(\tau)} - m_{s,x}^{(\tau-1)}$ acts only on bits x and $x + 1$. Therefore the other bits are projected out by the projection matrix

$$\mathcal{P}_x = \otimes_{x'=1}^n \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{array}{c} x' = x \\ \text{or} \\ x' = x+1 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} & \text{otherwise} \end{cases} . \quad (35)$$

Expressing Eq. (34) in this way will be useful later on.

E. $B_x^{(\pm)}$ encode the Hamiltonian

Given a Hamiltonian $H = \sum_x H_x$, we now wish to choose $B_x^{(\pm)}$ such that inserting Eq. (34) into (28) yields a discrete Schrödinger equation, $\boldsymbol{P}^{(\tau)} - \boldsymbol{P}^{(\tau-1)} \propto GP$, where we define

$$G = -iH \quad (36)$$

and $G_x = -iH_x$.

We assume H_x is Hermitian, imaginary-valued, and acts only on two neighboring qubits, which implies that G_x is a real antisymmetric 4×4 matrix. In accordance with Eq. (4), we further assume that G_x has rows and columns that sum to zero. We also decompose $G_x = G_x^{(+)} - G_x^{(-)}$ such that $G_x^{(+)}$ and $G_x^{(-)}$ only have nonnegative elements.

We require that $B_x^{(\pm)}$ satisfy Eqs. (16) and (17) so that these matrices encode the Hamiltonian, as in Eq. (19). To achieve this for a general (geometrically two-local) Hamiltonian, we choose $B_x^{(\pm)}$ to be the following 4×4 matrix:

$$B_x^{(\pm)}(b_\pm, a) = \delta_{\mathrm{t}} G_x^{(\pm)}(b_\pm, a) + \begin{cases} 1 - \delta_{\mathrm{t}}\, g_x(a) & b_\pm = a \\ 0 & b_\pm \neq a \end{cases}, \quad (37)$$

where $g_x(a)$ is a column sum of $G_x^{(\pm)}$:

$$g_x(a) = \sum_{b=00,01,10,11} G_x^{(\pm)}(b, a); \quad (38)$$

"b_+," "b_-," and "a" each denote a pair of bits, which respectively correspond to $b_{S,\mathbb{x},+\gamma}$, $b_{S,\mathbb{x},-\gamma}$, and $a_{S,\mathbb{x}}$ in Eq. (29). $B_x^{(\pm)}$ are 4×4 stochastic matrices, which we view as a probability vector for 2 bits ($b_\pm$) given two bits (a). Either $G_x^{(+)}$ or $G_x^{(-)}$ can be used in the right-hand side of Eq. (38); both give the same result since the column sum of $G_x = G_x^{(+)} - G_x^{(-)}$ is zero. $B_x^{(\pm)}$ is a stochastic matrix as long as δ_{t} is sufficiently small. See Eq. (9) for an example.

Importantly, since this choice of $B_x^{(\pm)}$ implies Eq. (19), the bits $b_{S,\gamma}$ encode a short time evolution by the Hamiltonian:

$$\sum_{\gamma=1,2} \overline{\hat{\boldsymbol{b}}_{S,+\gamma}^{(\tau)} - \hat{\boldsymbol{b}}_{S,-\gamma}^{(\tau)}}$$
$$= (+B^{(+1)} - B^{(-1)} + B^{(+2)} - B^{(-2)})\overline{\hat{\boldsymbol{a}}_S^{(\tau-1)}}$$
$$= (-i\delta_{\mathrm{t}} H)\boldsymbol{P}^{(\tau-1)} + O(\delta_{\mathrm{t}}^2). \quad (39)$$

The first equality follows from Eq. (29), where $B^{(\gamma)}$ was defined in Eq. (18). Similar to the $\hat{\boldsymbol{b}}_{s,\gamma}$ notation, $\hat{\boldsymbol{a}}_s$ denotes

a length-n basis vector indexed by the bit string a_s. The final line follows from Eq. (19) and the definition $\boldsymbol{P}^{(\tau)} = \overline{\hat{\boldsymbol{a}}_S^{(\tau)}}$.

F. Emergent quantum mechanics

Now that we have specified the EmQM model, we show that it can exhibit emergent quantum mechanics. We begin by evaluating the expectation value of $\boldsymbol{P}^{(\tau)} - \boldsymbol{P}^{(\tau-1)}$ using Eq. (28):

$$\overline{\boldsymbol{P}^{(\tau)} - \boldsymbol{P}^{(\tau-1)}} = \sum_{s=1}^{S} \sum_{x}^{\overset{\text{even}}{s-x}} Q_{S \leftarrow s} \overline{\left(m_{s,x}^{(\tau_s)} - m_{s,x}^{(\tau_s-1)}\right)} \frac{\mathbf{1}}{N} + O(m_0^2). \tag{40}$$

We can simplify the summand as follows:

$$\overline{\left(m_{s,x}^{(\tau_s)} - m_{s,x}^{(\tau_s-1)}\right)} \frac{\mathbf{1}}{N}$$

$$= \frac{\Delta_{\mathrm{m}}}{4} \sum_{\gamma=1,2} \mathcal{P}_x Q_{S \leftarrow s}^T \overline{\left(\hat{\boldsymbol{b}}_{S,+\gamma}^{(\tau_s')} - \hat{\boldsymbol{b}}_{S,-\gamma}^{(\tau_s')}\right)} \tag{41}$$

$$= \frac{\Delta_{\mathrm{m}}}{4} \mathcal{P}_x Q_{S \leftarrow s}^T (-i\delta_t H) \overline{\boldsymbol{P}^{(\tau_s'-1)}} + O(\delta_t^2). \tag{42}$$

Equation (41) is obtained from Eq. (34) after replacing $\hat{\boldsymbol{b}}_{s,\gamma}^{(\tau_s)} = Q_{S \leftarrow s}^T \hat{\boldsymbol{b}}_{S,\gamma}^{(\tau_s')}$, where $\tau_s' = \tau_s - (S - s) = \tau - 1 - 2(S - s)$. Equation (42) follows from Eq. (39).

We assume $\Delta_{\mathrm{m}} \ll S^{-1}$ such that the slow variables $m_{s,x}$ change very little over S time steps so that we can approximate $\boldsymbol{P}^{(\tau_s'-1)} = \boldsymbol{P}^{(\tau-1)} + O(S\Delta_{\mathrm{m}})$. Inserting Eq. (42) into (40) results in

$$\overline{\boldsymbol{P}^{(\tau)} - \boldsymbol{P}^{(\tau-1)}} = \frac{\Delta_{\mathrm{m}}}{4} \delta_t \mathcal{W}(-iH) \overline{\boldsymbol{P}^{(\tau-1)}} + O(m_0^2, \delta_t^2, \Delta_{\mathrm{m}}^2). \tag{43}$$

where $\mathcal{W}$ is a sum of conjugated projectors $\mathcal{P}_x$ [Eq. (35)]:

$$\mathcal{W} = \sum_{s=1}^{S} \sum_{x}^{\overset{\text{even}}{s-x}} Q_{S \leftarrow s} \mathcal{P}_x Q_{S \leftarrow s}^T. \tag{44}$$

The error estimates $O(m_0^2, \delta_t^2, \Delta_{\mathrm{m}}^2)$ in Eq. (43) neglect factors of S and N, which will be accounted for in Sec. IV.

If $Sn \gg N$, then $\mathcal{W}$ approaches a simple form:

$$\mathcal{W} = \underbrace{\frac{Sn}{2}\mathcal{P}_{\mathbf{1}} + \frac{Sn}{2}\frac{3}{N-1}\mathcal{P}_{\mathbf{1}}^{\perp}}_{\mathcal{W}_0} + O\left(\sqrt{\frac{Sn}{N}}\right). \tag{45}$$

where the matrix $\mathcal{P}_{\mathbf{1}} = \frac{\mathbf{1} \otimes \mathbf{1}}{N}$ projects onto the one-dimensional subspace spanned by $\mathbf{1}$ (i.e., the vector of N ones) and $\mathcal{P}_{\mathbf{1}}^{\perp} = \mathbb{1} - \mathcal{P}_{\mathbf{1}}$ projects into the orthogonal subspace. $\mathbb{1}$ is an $N \times N$ identity matrix.

The first term in Eq. (45) follows from multiplying Eq. (44) by $\mathbf{1}$, and noting that $Sn/2$ is the number of terms summed in Eq. (44). To understand the following two terms, note that within the $(N-1)$-dimensional subspace orthogonal to $\mathbf{1}$, the set of permutation matrices is a unitary 1-design [62] up to a constant factor. That is,

$$\mathbb{E}_{Q} Q_{\perp} \otimes Q_{\perp}^T = \frac{1}{1 - N^{-1}} \mathbb{E}_{U} U_{\perp} \otimes U_{\perp}^{\dagger}, \tag{46}$$

where $Q_{\perp} = \mathcal{P}_{\mathbf{1}}^{\perp} Q \mathcal{P}_{\mathbf{1}}^{\perp}$ and $U_{\perp} = \mathcal{P}_{\mathbf{1}}^{\perp} U \mathcal{P}_{\mathbf{1}}^{\perp}$, and $\mathbb{E}_{Q}$ averages over all permutation matrices Q while $\mathbb{E}_{U}$ averages over all unitary matrices.[8] Our numerical experiments show that the permutation matrices $Q_{S \leftarrow s}$ with $S - s \gg n$ summed in Eq. (44) are an approximate 1-design in the same sense [65–67]:

$$\mathbb{E}_{Q_{s,x}} (Q_{S \leftarrow s})_{\perp} \otimes (Q_{S \leftarrow s})_{\perp}^T \approx \frac{1}{1 - N^{-1}} \mathbb{E}_{U} U_{\perp} \otimes U_{\perp}^{\dagger}, \tag{47}$$

where $(Q_{S \leftarrow s})_{\perp} = \mathcal{P}_{\mathbf{1}}^{\perp} Q_{S \leftarrow s} \mathcal{P}_{\mathbf{1}}^{\perp}$. The above equation allows us to approximate the average of $\mathcal{W}$ (in the subspace orthogonal to $\mathbf{1}$) as

$$\mathcal{P}_{\mathbf{1}}^{\perp} \overline{\mathcal{W}} \mathcal{P}_{\mathbf{1}}^{\perp} \approx \frac{S}{2} \frac{1}{1 - N^{-1}} \sum_{x} \mathbb{E}_{U} U_{\perp} \mathcal{P}_x U_{\perp}^{\dagger}$$

$$= \frac{S}{2} \frac{1}{N - 1} \sum_{x} \mathcal{P}_{\mathbf{1}}^{\perp} \mathcal{P}_x \mathcal{P}_{\mathbf{1}}^{\perp}$$

$$= \frac{Sn}{2} \frac{3}{N - 1} \mathcal{P}_{\mathbf{1}}^{\perp}. \tag{48}$$

Although $Q_{S \leftarrow s}$ near the boundary (i.e., $S - s \ll n$) will not obey the approximate 1-design property (used to obtain the first equality), the sum of conjugated projection operators near the boundary in Eq. (44) only contributes at order $O(n^2)$, which is negligible compared to the other terms in Eq. (45). We used $\mathbb{E}_{U} U_{ij} \otimes U_{kl}^* = \frac{1}{N} \delta_{ik} \delta_{jl}$ (where δ_{ij} denotes a Kronecker delta) to obtain the second line above. The final line follows from the definitions of $\mathcal{P}_{\mathbf{1}}^{\perp}$ and $\mathcal{P}_x$ [Eq. (35)]. This is the second term in Eq. (45). To quantify statistical fluctuations about the above mean, we numerically find that the standard deviation of the eigenvalues of $\mathcal{W}$ within the subspace orthogonal to $\mathbf{1}$ is approximately $4\sqrt{Sn/2N}$ [the third term in Eq. (45)], which is much smaller than the $O(Sn/N)$ mean.

Note that $\mathbf{1}H = 0$ since the columns of H sum to zero [Eq. (4)]. Therefore inserting Eq. (45) into (43) yields

$$\overline{\boldsymbol{P}^{(\tau)} - \boldsymbol{P}^{(\tau-1)}} \approx -\Delta_t iH \overline{\boldsymbol{P}^{(\tau-1)}} \tag{49}$$

up to $O(m_0^2, \delta_t^2, \Delta_{\mathrm{m}}^2, \sqrt{N/Sn})$ corrections, where

$$\Delta_t = \frac{\Delta_{\mathrm{m}}}{4} \frac{Sn}{2N} \frac{3}{1 - N^{-1}} \delta_t \tag{50}$$

is the effective EmQM time step. Equation (49) is precisely the discrete-time analog of the emergent Schrödinger's equation (20) that we wanted.

The emergent wave function $\Psi(t)$ can be extracted from $\boldsymbol{P}^{(\tau)}$ in Eq. (2) as follows:

$$\Psi(t = \Delta_t \tau) \propto \boldsymbol{P}^{(\tau)} - \frac{\mathbf{1}}{N}. \tag{51}$$

---

[8]Equation (46) also holds if $\mathbb{E}_{U}$ only averages over orthogonal matrices (since the set of orthogonal matrices is also a unitary 1-design) or if $\mathbb{E}_{Q}$ averages only over permutation matrices that are affine transformations (which is equivalent to the set of matrices generated by permutation matrices that act only on two bits [63,64]). Equation (46) follows after evaluating $\mathbb{E}_{Q} Q_{ij} Q_{kl} = \frac{1}{N} \mathbb{1}_{ik} \mathbb{1}_{jl} + \frac{1}{N(N-1)} (\mathbf{1} \otimes \mathbf{1} - \mathbb{1})_{ik} (\mathbf{1} \otimes \mathbf{1} - \mathbb{1})_{jl}$ and $\mathbb{E}_{U} U_{ij} \otimes U_{kl}^* = \frac{1}{N} \mathbb{1}_{ik} \mathbb{1}_{jl}$. The components $\mathbb{1}_{ij}$ of the identity matrix $\mathbb{1}$ are equivalent to a Kronecker delta: $\delta_{ij} = \mathbb{1}_{ij}$.

Ψ obeys Schrödinger's equation exactly in a $m_0, \delta_t, S^{-1}, \Delta_m \to 0$ limit, which we derive below.

IV. DEVIATIONS FROM QUANTUM MECHANICS

We now study how much the EmQM model deviates from Schrödinger's equation. In particular, we will estimate how

$$\varepsilon(t) = ||\Psi(t) - \Psi_{QM}(t)|| \quad (52)$$

grows with time, where Ψ is the EmQM wave function defined by Eq. (51), Ψ_{QM} is calculated using Schrödinger's equation (1), and $||\cdots||$ denotes a Euclidean 2-norm. At time $t = 0$, we take $\Psi_{QM}(0) = \Psi(0)$. The deviation $\varepsilon(t)$ from quantum mechanics increases due to four separate contributions resulting from finite m_0, δ_t, S^{-1}, and Δ_m. We find that small values of these parameters result in linearity, unitarity, locality, and small statistical fluctuations, respectively.

Calculating the contributions to $\varepsilon(t)$ is rather technical. For brevity, we will only sketch the derivation, which we verify numerically in Sec. V. While estimating $\varepsilon(t)$, we also ignore constant factors (e.g., factors of 2), which we emphasize by using "$\sim$" symbols (instead of "$\approx$").

A. Preliminaries

It is essential to differentiate systematic and statistical errors, which respectively add coherently and incoherently. That is, if one adds a series of n errors following a normal distribution with mean μ (systematic error) and standard deviation σ (statistical error), then the total error is also normally distributed with mean $n\mu$ and standard deviation $\sqrt{n}\sigma$. We say that the means add coherently ($\propto n$), while the standard deviation adds incoherently ($\propto \sqrt{n}$).

It is instructive to first roughly estimate the 2-norm of $\overline{P^{(\tau)} - P^{(\tau-1)}} \approx -\Delta_t i H P^{(\tau-1)}$ in Eq. (49):

$$\Delta_P \approx ||\Delta_t\, iHP|| \sim \Delta_t \sqrt{n}\, \epsilon_\Psi. \quad (53)$$

This follows from $HP = \epsilon_\Psi H\Psi$ via the constraint (4) that $\sum_j H_{ij} = 0$. And $||H\Psi|| \sim \sqrt{n}$ due to the n terms in $H = \sum_x H_x$ which add up incoherently for generically random wave functions Ψ (which we consider in our numerical validation)[9] due to approximate orthogonality. This results in the $\sqrt{n}$ factor above. For simplicity, we assume that each H_x (viewed as a 4×4 matrix) has norm roughly equal to 1.

We will also require an estimate for ϵ_Ψ, which is defined by $P = \frac{1}{N} + \epsilon_\Psi\Psi$ in Eq. (2) with $||\Psi|| = 1$:

$$\epsilon_\Psi \sim \sqrt{\frac{Sn}{N}}\, m_0. \quad (54)$$

This expression results from Eq. (28), which expresses $P - \frac{1}{N}$ as a sum of roughly $Sn/2$ terms of the form $Q_{S\leftarrow s}m_{s,x}\frac{1}{N}$. These $Sn/2$ terms add up incoherently, and each term has a norm of roughly $m_0/\sqrt{N}$ (since $\mathbf{1}$ has norm $\sqrt{N}$), which results in the above expression.

B. Small m_0 controls linearity

As noted in Sec. II B, we expect linearity to result only if $M_{s,x}$ changes very little over time, which is controlled by the smallness of m_0. This limit also allowed us to expand Eq. (28) to first order in m_0. Keeping all higher-order terms yields a more accurate version of Eq. (40):

$$\overline{P^{(\tau)} - P^{(\tau-1)}} = \sum_{s=1}^{S} \sum_{x}^{\substack{\text{even} \\ s-x}} M_{S\leftarrow s}\overline{\left(m_{s,x}^{(\tau_s)} - m_{s,x}^{(\tau_s-1)}\right)}\frac{\mathbf{1}}{N} + O(\Delta_m^2), \quad (55)$$

where we only drop terms that are second order in $\Delta_m \sim ||m_{s,x}^{(\tau_s)} - m_{s,x}^{(\tau_s-1)}||_{op} \ll m_0$. We use $||M||_{op}$ to denote the operator norm of a matrix M, which is equivalent to the largest singular value of M. Analogous to $Q_{S\leftarrow s}$, we define $M_{S\leftarrow s}$ as a product of stochastic matrices $M_{s',x}$ with $S \geqslant s' > s$:

$$M_{S\leftarrow s} = M_S M_{S-1} \cdots M_{s+1}. \quad (56)$$

This leads to a nonlinear Schrödinger's equation

$$\partial_t \Psi \approx -i\widetilde{\mathcal{W}}H\Psi, \quad (57)$$

where $\widetilde{\mathcal{W}}$ is similar to $\mathcal{W}$ in Eq. (44) but with $Q_{S\leftarrow s}$ replaced with $M_{S\leftarrow s}$:

$$\widetilde{\mathcal{W}} = \sum_{s=1}^{S} \sum_{x}^{\substack{\text{even} \\ s-x}} M_{S\leftarrow s}\mathcal{P}_x Q_{S\leftarrow s}^T. \quad (58)$$

Equation (57) is highly nonlinear because the right-hand-side depends on the product of many dynamical variables $m_{s,x}$ (via $\widetilde{\mathcal{W}}$) in addition to Ψ. Furthermore, unlike Schrödinger's equation, time-evolving Ψ using Eq. (57) also requires keeping track of the time evolution of $m_{s,x}$.

To estimate how much nonlinear corrections to the emergent Schrödinger equation contribute to the deviation $\varepsilon(t)$ in Eq. (52), consider the $O(Sn)$ many $O(m_0)$ terms in $\widetilde{\mathcal{W}}$ that were neglected in $\mathcal{W}$. These terms add up coherently[10] by subtracting weight from the identity matrix component of $\widetilde{\mathcal{W}}$, such that

$$||\widetilde{\mathcal{W}} - \mathcal{W}||_{op} \sim \min(Snm_0, 1). \quad (59)$$

The $\min(\cdots, 1)$ results because once $Snm_0 \gtrsim 1$, most of the weight has been removed from the identity matrix component. This correction to $\mathcal{W}$ adds coherently to $\varepsilon(t)$ over many time steps. Indeed, we find that $\partial_t\varepsilon(t) \propto ||\widetilde{\mathcal{W}} - \mathcal{W}||_{op}$:

$$\partial_t \varepsilon(t) = \partial_t ||\Psi - \Psi_{QM}||$$

$$= \text{Re}\frac{(\Psi - \Psi_{QM})^*}{||\Psi - \Psi_{QM}||} \cdot \partial_t(\Psi - \Psi_{QM})$$

$$\lesssim \left|\left|\text{Im}\frac{(\Psi - \Psi_{QM})^*}{||\Psi - \Psi_{QM}||} \cdot (\widetilde{\mathcal{W}}H\Psi - \mathcal{W}H\Psi_{QM})\right|\right| \quad (60)$$

$$\sim \left|\left|\text{Im}\frac{(\Psi - \Psi_{QM})^*}{||\Psi - \Psi_{QM}||} \cdot (\widetilde{\mathcal{W}} - \mathcal{W})H\Psi\right|\right| \quad (61)$$

---

[9]$||H\Psi|| \sim n$ for low-energy states. Such states thus require an extra factor of $\sqrt{n}$ in Eq. (53), but this is relatively negligible compared e.g., to factors of N.

[10]There is also an incoherent contribution. However, the coherent contribution dominates for large S since it adds up more rapidly.

$$\leqslant ||\widetilde{\mathcal{W}} - \mathcal{W}||_{\mathrm{op}} ||H\Psi||$$

$$\sim \min(Snm_0, 1)\sqrt{n}. \qquad (62)$$

In Eq. (60) we wish to bound how quickly the deviation $\varepsilon(t)$ increases just due to finite m_0. Thus, we inserted the nonlinear Eq. (57) in for $\partial_t \Psi$, and we used $\partial_t \Psi_{\mathrm{QM}} \approx -i\mathcal{W}H\Psi_{\mathrm{QM}}$ since here we want to ignore the $O(\sqrt{Sn/N})$ corrections to $\mathcal{W}$ in Eq. (45). Equation (61) also follows from ignoring the $O(\sqrt{Sn/N})$ corrections to $\mathcal{W}$ since $\mathrm{Im}(\Psi - \Psi_{\mathrm{QM}})^* H(\Psi - \Psi_{\mathrm{QM}}) = 0$. Equation (62) follows from Eq. (59). The $\sqrt{n}$ factor results from the n terms in the Hamiltonian, which add incoherently as in Eq. (53).

We therefore find that the small m_0 approximation contributes $\varepsilon_{\mathrm{m}}(t)$ to $\varepsilon(t)$ where

$$\varepsilon_{\mathrm{m}}(t) \sim \min(Snm_0, 1)\sqrt{n}\, t. \qquad (63)$$

C. Small δ_{t} controls unitarity

To obtain Eq. (39), which is inserted into Eq. (42), we kept only terms with a single factor of $\delta_t H_x$, which is justified for small δ_t. Terms that are higher-order in δ_t lead to a nonunitary evolution in Ψ. A more accurate version of Eq. (49) would replace

$$-i\delta_t H \to \overset{\mathrm{odd}}{\underset{x}{\otimes}} B_x^{(+)} - \overset{\mathrm{odd}}{\underset{x}{\otimes}} B_x^{(-)} + \overset{\mathrm{even}}{\underset{x}{\otimes}} B_x^{(+)} - \overset{\mathrm{even}}{\underset{x}{\otimes}} B_x^{(-)}, \quad (64)$$

which leads to an effective Hamiltonian that has $O(\delta_t)$ non-Hermitian terms.

To estimate the contribution of finite δ_t to $\varepsilon(t)$, note that the right-hand side of Eq. (64) contains roughly n^2 many $O(\delta_t^2)$ terms. When Eq. (64) is inserted into Schrödinger's equation, these n^2 terms add up incoherently, leading to an $O(n\delta_t)$ correction. Similar to Eq. (62), after many time steps, this correction adds coherently and contributes $\varepsilon_{\mathrm{t}}(t)$ to $\varepsilon(t)$, where

$$\varepsilon_{\mathrm{t}}(t) \sim n\delta_t t. \qquad (65)$$

D. Large S controls emergent locality

Equation (43) leads to a modified Schrödinger's equation:

$$\partial_t \Psi \approx -i\mathcal{W}H\Psi. \qquad (66)$$

$\mathcal{W}$ was defined in Eq. (44) and consists of a sum of $Sn/2$ projection matrices $Q_{S\leftarrow s}\mathcal{P}_x Q_{S\leftarrow s}^T$, which each project onto a four-dimensional subspace. These are the four-dimensional subspaces for which the $Sn/2$ stochastic matrices $M_{s,x}$ can affect $\boldsymbol{P}$. Therefore, $\mathcal{W}$ is a real symmetric matrix that has at most $2Sn$ nonzero eigenvalues. If $2Sn \ll N$, then $\mathcal{W}$ reduces the dynamics of Ψ to a random $(2Sn)$-dimensional subspace.[11] If $Sn \gg N$, then $\mathcal{W}$ has full rank; however, recall from Eq. (45) that the eigenvalues of $\mathcal{W}$ are randomly distributed (due to the random $Q_{s,x}$) with a standard deviation that is $\sqrt{N/Sn}$ times smaller than the mean. The eigenvectors

of $\mathcal{W}$ do not have any special property that preserves locality. Therefore, the dynamics of Ψ are local only to the extent that $Sn \gg N$. That is, far-away terms in the effective Hamiltonian $\mathcal{W}H = \sum_x \mathcal{W}H_x$ commute only in the $Sn \gg N$ limit (which we have checked numerically); when $Sn < N$, the noise in the randomness of $\mathcal{W}$ leads to nonlocal dynamics.

To intuitively understand the last point, once could consider approximating the randomness in $\mathcal{W} - \mathcal{W}_0$ as a symmetric matrix of small Gaussian random entries. $\mathcal{W}_0$, defined in Eq. (45), is the nonrandom part of $\mathcal{W}$. Since $\mathcal{W} - \mathcal{W}_0$ is highly nonlocal, it makes $\mathcal{W}H$ also become nonlocal. Technically, this nonlocality occurs only for circuit depths S between $n \ll S \ll N/n$. If the circuit depth is very small $S \ll n$, then the dynamics of the emergent wave function [Eq. (51)] will actually be local. This results because the "light cone" of back-propagating bits can only spatially extend to $\Delta x \sim S$ when $S \ll n$. However, although the dynamics are local when $S \ll n$, $\mathcal{W}$ will have very few nonzero eigenvalues, which causes the EmQM model to be a very bad approximation to quantum mechanics.

The $Sn \gg N$ requirement for locality is problematic because if there are many $n \gg 1$ qubits, then the length S of the extra dimension would have to be tremendously large ($Sn \gg 2^n$) in order for Ψ to have local dynamics. Mitigating this nonlocality is an important future direction that we elaborate on in Sec. VII A.

Interestingly, $\mathcal{W}H$ at least has real eigenvalues. This occurs because $\mathcal{W}H$ is similar[12] to

$$\widetilde{H} = \mathcal{W}^{1/2} H \mathcal{W}^{1/2} \qquad (67)$$

(when $\mathcal{W}$ is invertible), which implies that $\mathcal{W}H$ and $\widetilde{H}$ have the same eigenvalues. $\mathcal{W}^{1/2}$ is Hermitian, which implies that $\widetilde{H}$ is also Hermitian. Therefore $\widetilde{H}$ and $\mathcal{W}H$ both have real eigenvalues. Furthermore, by absorbing $\mathcal{W}$ into the Hamiltonian as above, we obtain an effective Schrödinger's equation:

$$\partial_t \widetilde{\Psi} = -i\widetilde{H}\widetilde{\Psi}, \qquad (68)$$

where the wave function is

$$\widetilde{\Psi} = \mathcal{W}^{-1/2}\Psi. \qquad (69)$$

However, the issue of nonlocality remains in the sense that $\widetilde{H}$ contains geometrically nonlocal and high-weight terms.

To estimate the contribution of finite S to $\varepsilon(t)$, recall from Eq. (45) that the eigenvalues of $\mathcal{W}$ are random with standard deviation $\sqrt{N/Sn}$. This randomness induces an $O(\sqrt{N/Sn})$ correction to $\mathcal{W}$ in Eq. (66). Similar to Eq. (62) [but with $\widetilde{\mathcal{W}}$ and $\mathcal{W}$ respectively replaced by $\mathcal{W}$ and $\mathcal{W}_0$ from Eq. (45)], this correction adds coherently over many time steps and contributes $\varepsilon_S(t)$ to $\varepsilon(t)$, where

$$\varepsilon_S(t) \sim \sqrt{\frac{N}{S}}\, t. \qquad (70)$$

Similar to Eqs. (53) and (62), the extra factor of $\sqrt{n}$ results from the n terms in the Hamiltonian.

---

[11]The subspace is spanned by states that are each a symmetric superposition of a random quarter of the states in the classical basis, e.g., states like $\frac{1}{2}|0011\rangle + \frac{1}{2}|0110\rangle + \frac{1}{2}|1010\rangle + \frac{1}{2}|1111\rangle$.

[12]Matrices A and B are similar if $B = PAP^{-1}$ for some invertible matrix P.

E. Small Δ_m controls statistical fluctuations

So far, we have focused on only the mean of $\boldsymbol{P}^{(\tau)}$. But $\boldsymbol{P}$ is defined by the stochastic matrices $M_{s,x}$, which have stochastic dynamics that induce statistical fluctuations on the time evolution of $M_{s,x}$ and $\boldsymbol{P}$. However, these statistical fluctuations will be small if the perturbations $m_{s,x}$ change very slowly with time and aren't too small, which will give the output bits a_S enough time to thoroughly sample the wave function $\Psi \propto \boldsymbol{P} - \mathbf{1}/N$.

The contribution of statistical fluctuations to $\varepsilon(t)$ can be estimated by considering how much $\boldsymbol{P}$ will be affected by statistical fluctuations after τ time steps. $\boldsymbol{P}$ [defined in Eq. (23)] will change only after time steps for which an $M_{s,x}$ matrix changes, which occurs only if the $\pm$ back-propagating bits are different [see Eq. (31)], i.e., when $b_{S,+\gamma} \neq b_{S,-\gamma}$ [defined below Eq. (34)]. For small $n\delta_t \ll 1$, this occurs for roughly a $n\delta_t$ fraction of time steps since $B_x^{(\pm)} = \mathbb{1}_4 + O(\delta_t)$ [Eq. (17)].

We can then think of the statistical fluctuations as a random-walk in an N-dimensional space. Recall that after n_steps steps with typical step length ℓ_step, a random walker will have moved a Euclidean distance of roughly $\ell_\mathrm{step}\sqrt{n_\mathrm{steps}}$. After τ time steps, our random walker will move $n_\mathrm{steps} \sim n\delta_t\tau$ times due to the arguments above. Below we argue that the length of each step will be roughly

$$\ell_\mathrm{step} \sim \frac{\Delta_\mathrm{m}}{N}Sn. \tag{71}$$

This implies that the statistical fluctuations to $\boldsymbol{P}$ will grow as

$$\frac{\Delta_\mathrm{m}}{N}Sn\sqrt{n\delta_t\tau}. \tag{72}$$

Consider a time step for which $b_{S,+\gamma} \neq b_{S,-\gamma}$. Most of the $Sn/2$ perturbations $m_{s,x}$ will be modified by an amount proportional to Δ_m due to the resulting back-propagating bits. Each such modification shifts $\boldsymbol{P}$ by roughly a distance $\Delta_\mathrm{m}/\sqrt{N}$, since $N^{-1/2}$ is the norm of the uniform probability vector $\mathbf{1}/N$ that multiplies $m_{s,x}$ in Eq. (28). However, one component of the shift to $\boldsymbol{P}$ adds up coherently (over the $Sn/2$ many perturbations $m_{s,x}$), while the other $N-1$ basis components add incoherently. [The coherent component is spanned by $\hat{\boldsymbol{b}}_{S,+\gamma} - \hat{\boldsymbol{b}}_{S,-\gamma}$, which appears in Eq. (41).] The incoherent components are negligible when $Sn \gg N$. Projecting onto the coherent contribution reduces $\Delta_\mathrm{m}/\sqrt{N}$ by a factor[13] of $N^{-1/2}$. This explains the Δ_m/N factor in Eq. (71). The Sn factor occurs because there are $Sn/2$ many perturbations $m_{s,x}$.

These statistical fluctuations bound $\varepsilon(t) \geqslant \varepsilon_\mathrm{stat}(t)$ where

$$\varepsilon_\mathrm{stat}(t) \sim \epsilon_\Psi^{-1}\frac{\Delta_\mathrm{m}}{N}Sn\sqrt{n\delta_t\tau}. \tag{73}$$

The factor of ϵ_Ψ^{-1} [Eq. (54)] over Eq. (72) results from solving for Ψ in Eq. (2): $\boldsymbol{P} = \frac{1}{N} + \epsilon_\Psi\Psi$. The above simplifies to

$$\varepsilon_\mathrm{stat}(t) \sim m_0^{-1}\sqrt{\Delta_\mathrm{m}nt} \tag{74}$$

after replacing $\tau \to t/\Delta_t$ using Eq. (50).

---

[13]To understand the $N^{-1/2}$ factor, consider the signed sum $\boldsymbol{S} = \sum_{k=1}^{K}\mathrm{sign}(v_1^{(k)})\boldsymbol{v}^{(k)}$ of random unit-normalized length-N vectors $\boldsymbol{v}^{(k)}$. Only the first component S_1 adds up coherently, while the other $N-1$ components add incoherently, resulting in $||\boldsymbol{S}|| \sim K/\sqrt{N} + \sqrt{K}$ for large K and N. When $K \gg N$, the coherent contribution (first term) dominates.

F. Negligible delay

There is an $O(S)$ discrete time delay between when a string of output bits a_S is sampled to when the perturbations $m_{s,x}$ are updated. However, if the perturbations $m_{s,x}$ change sufficiently slowly with time due to small Δ_m, then this delay has a negligible effect on the wave function [which we will find to be the case in Eq. (80)].

To estimate the contribution to $\varepsilon(t)$, recall that in Eq. (43) we assumed that $\boldsymbol{P}^{(\tau)}$ varies slowly over $2S$ time steps such that $\boldsymbol{P}^{(\tau_s'-1)} = \boldsymbol{P}^{(\tau-1)} + O(S\Delta_\mathrm{P})$ where $\tau_s' = \tau - 1 - 2(S - s)$, and Δ_P is defined in Eq. (53). This approximation induces an $O(\Delta_t\sqrt{n}S\Delta_\mathrm{P})$ correction in Eq. (49), where the $\Delta_t\sqrt{n}$ follows for the same reason as in Eq. (53). After τ time steps, this correction adds coherently and contributes $\varepsilon_\mathrm{delay}(t)$ to $\varepsilon(t)$, where $\varepsilon_\mathrm{delay}(t) \sim \Delta_t\sqrt{n}S\Delta_\mathrm{P}\tau/\epsilon_\Psi$, which simplifies to

$$\varepsilon_\mathrm{delay}(t) \sim \Delta_t nS t$$

$$\sim \Delta_\mathrm{m}\frac{n^2S^2}{N}\delta_t t. \tag{75}$$

G. Convenient limit

The above contributions to $\varepsilon(t) = ||\Psi(t) - \Psi_\mathrm{QM}(t)||$ add up incoherently, such that the total deviation from Schrödinger's equation is roughly

$$\varepsilon(t) \sim \sqrt{\varepsilon_\mathrm{m}(t)^2 + \varepsilon_\mathrm{t}(t)^2 + \varepsilon_\mathrm{S}(t)^2 + \varepsilon_\mathrm{stat}(t)^2} \tag{76}$$

until saturation near orthogonality at $\varepsilon(t) \approx \sqrt{2}$. $\varepsilon_\mathrm{delay}(t)$ also contributes, but we neglect it here since it contributes negligibly in the limit that we consider.

It is convenient to consider a limit of S, δ_t, m_0, and Δ_m as a function of a single small parameter ϵ_0 such that $\varepsilon(t) \to 0$ in the $\epsilon_0 \to 0$ limit. We shall consider the parametrization such that all errors are roughly equal at time $t = \epsilon_0^{-1}$:

$$\varepsilon_\mathrm{m}(t) \sim \varepsilon_\mathrm{t}(t) \sim \varepsilon_\mathrm{S}(t) \sim \epsilon_0 t,$$
$$\varepsilon_\mathrm{stat}(t) \sim \sqrt{\epsilon_0 t}. \tag{77}$$

This parametrization results in a deviation

$$\varepsilon(t) \sim \sqrt{\epsilon_0 t + 3(\epsilon_0 t)^2}$$
$$\sim \sqrt{\epsilon_0 t} \quad \text{when } t \lesssim t_0^{-1}, \tag{78}$$

which is dominated by statistical fluctuations $\varepsilon_\mathrm{stat}(t) \sim \sqrt{\epsilon_0 t}$ until $t \geqslant \epsilon_0^{-1}$. Solving Eq. (77) for S, δ_t, m_0, and Δ_m results in the following parametrization:

$$S \approx N\epsilon_0^{-2},$$
$$\delta_t = \frac{\epsilon_0}{n},$$
$$m_0 = \frac{\epsilon_0}{Sn^{3/2}} \approx \frac{\epsilon_0^3}{n^{3/2}N},$$
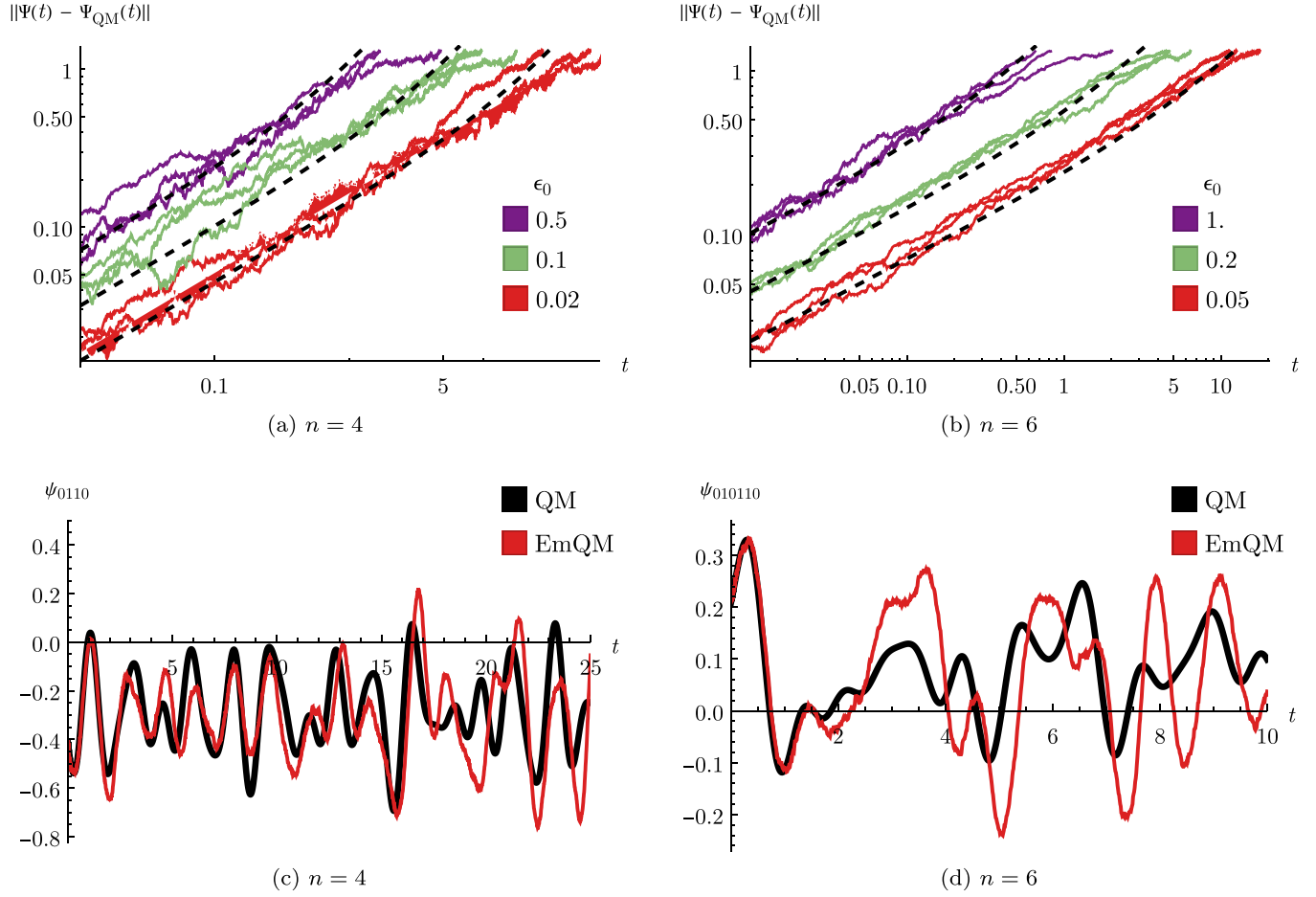$$\Delta_\mathrm{m} = \frac{m_0^2\epsilon_0}{n} \approx \frac{\epsilon_0^7}{n^4N^2}. \tag{79}$$

FIG. 2. (a, b) To validate our error estimates, we plot the deviation $\varepsilon(t) = ||\Psi(t) - \Psi_{\mathrm{QM}}(t)||$ vs time t between the quantum mechanics (QM) wave function Ψ_{QM} and emergent QM (EmQM) wave function Ψ [Eq. (51)] for (a) $n = 4$ and (b) $n = 6$ qubits using three different random initializations (colored lines) for each choice of the control parameter ϵ_0 [see Eq. (79)]. As ϵ_0 decreases, the EmQM model becomes increasingly more accurate, roughly agreeing with QM out to time $t \sim \epsilon_0^{-1}$. Dashed black lines show the estimated deviation $\varepsilon(t) \sim \sqrt{\epsilon_0 t + 3(\epsilon_0 t)^2}$ [Eq. (78)] for the three different values of ϵ_0, which match the simulated data remarkably well. (c, d) To demonstrate that the EmQM model is reproducing nontrivial dynamics, we also plot the 0110 (and 010110) component of the $n = 4$ (and $n = 6$) wave functions vs time for QM and the EmQM model with $\epsilon_0 = 0.02$ (and $\epsilon_0 = 0.05$), for which $S = 800$ (and $S = 1280$) [in accordance with Eq. (79)].

Plugging the above parameters into $\varepsilon_{\mathrm{delay}}(t)$ in Eq. (75) shows that deviations due to $\varepsilon_{\mathrm{delay}}(t)$ are negligibly small:

$$\varepsilon_{\mathrm{delay}}(t) \sim \frac{\epsilon_0^4}{n^3 N} t \ll \varepsilon(t). \tag{80}$$

V. SIMULATION

To numerically verify our theoretical results, we simulate the Hamiltonian $H = \sum_x H_x$ within the EmQM model with

$$H_x = Y_x X_{x+1} - Y_x, \tag{81}$$

which is a simple choice that satisfies the Hamiltonian constraints (4). To do this, we pick a ϵ_0 to define the model parameters according to Eq. (79). We then initialize the $m_{s,x}$ matrices using Gaussian random numbers with standard deviation m_0 and then subtract a constant from each column such that all columns sum to zero. This random initialization implicitly defines a random wave function [via (51)]. We then

time evolve the circuit for many steps. Ψ is normalized and extracted from P using Eq. (51).

However, simulating the EmQM model is extremely expensive for small ϵ_0 or many qubits since the CPU time required to simulate out to time t scales as

$$\mathrm{CPU\ time} \sim S n\, t/\Delta_{\mathrm{t}} \sim \frac{N}{\delta_{\mathrm{t}} \Delta_{\mathrm{m}}} t \approx \frac{n^5 N^3}{\epsilon_0^8} t. \tag{82}$$

The first relation follows since t/Δ_{t} time steps are required on an $S \times n$ lattice. The second relation is obtained by inserting Eq. (50) for Δ_{t}. The final expression is valid for the ϵ_0 parametrization (79). However, we can approximately simulate the EmQM model with high accuracy using the significantly faster method defined and verified in Appendix D. We validate our theory for the EmQM model in Fig. 2 by plotting the deviation $\varepsilon(t)$ of the emergent wave function Ψ from the quantum mechanics prediction. We find that the deviation is in agreement with the estimated Eq. (78). This verifies that as the control parameter ϵ_0 decreases, the EmQM model
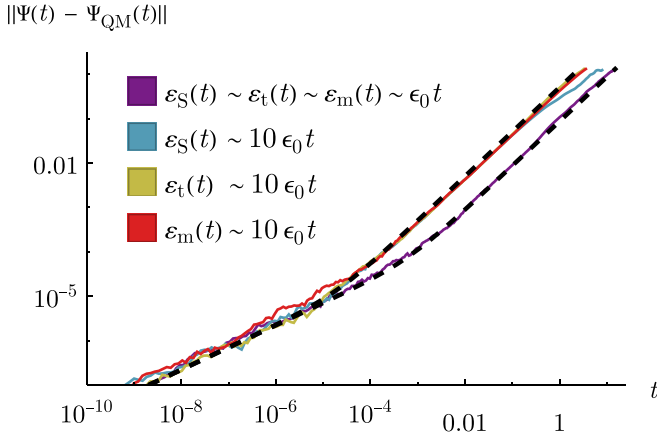
$||\Psi(t) - \Psi_{QM}(t)||$

FIG. 3. Deviation $\varepsilon(t) = ||\Psi(t) - \Psi_{QM}(t)||$ from quantum mechanics (QM) vs time t. The deviation is dominated by statistical fluctuations $\varepsilon_{stat}(t)$ until the change in slope, after which the deviation is dominated by other contributions. Blue, yellow, and red lines show dominant contributions from $\varepsilon_{S}(t)$, $\varepsilon_{t}(t)$, and $\varepsilon_{S}(m)$, respectively, defined in the legend (with $\varepsilon_{stat}(t) \sim 10^{-2}\sqrt{\epsilon_0 t}$ and $\varepsilon_{S}(t) \sim \varepsilon_{t}(t) \sim \varepsilon_{m}(t) \sim \epsilon_0 t$ whenever unspecified). Simulations are for $n = 4$ qubits and $\epsilon_0 = 0.05$ with parameters S, δ_t, m_0, and Δ_m chosen to target the previously mentioned contributions [using Eqs. (63), (65), and (70)]. The dashed lines plot the expected deviations $\varepsilon(t)$ using Eq. (76), which agree remarkably well with the simulation data (colored lines).

deviates less from quantum mechanics. All data are generated using the approximate simulation method (Appendix D) except for the $n = 4$ data with $\epsilon_0 = 0.5$ [purple in Fig. 2(a)], for which we could directly simulate the EmQM model.

The deviations from quantum mechanics shown in Fig. 2 are dominated by the statistical deviation $\varepsilon_{stat}(t)$. To verify the other contributions to the deviation $\varepsilon(t)$, we solve for parameters S, δ_t, m_0, and Δ_m such that the statistical deviation is much smaller and such that only one of the other contributions is expected to dominate. This allows us to verify each contribution individually in Fig. 3.

VI. EXPERIMENTAL SIGNATURES

A. Many entangled high-fidelity qubits

Evidence that an EmQM model really describes nature might be found by measuring deviations from Schrödinger's equation, such as the $\varepsilon(t)$ [Eq. (52)] studied in the previous section. In Sec. IV we calculated several contributions to $\varepsilon(t)$. All contributions to $\varepsilon(t)$ can be made extremely small for very long times by taking S^{-1}, δ_t, m_0, and Δ_m to be very small. $\varepsilon_{S}(t) \sim t\sqrt{N/S}$ in Eq. (70) is the only deviation that increases polynomially with N. This increase with N is significant because $N = 2^n$ is exponentially large in the number of qubits n. Therefore, even if $S \sim 10^{100}$ or 10^{1000}, only $n \sim 350$ or 3500 qubits would be needed to obtain a large deviation $\varepsilon_{S}(t) \sim 1$ after a short time $t \lesssim 1$.

But what should be the value of n? If n is taken to be the number of qubits needed to describe just a mesoscopic region of space, e.g., Avogadro's number $n \sim 10^{23}$, then a large deviation from quantum mechanics due to $\varepsilon_{S}(t)$ is already predicted for very short times unless S is extremely large

$S \gg 2^{10^{23}}$. Indeed, in Sec. IV D we found that our model exhibits nonlocal EmQM dynamics unless $Sn \gg 2^n$. Therefore, it seems implausible that the particular EmQM model studied in this work accurately describes a possible EmQM for our universe.

However, we speculate that modifications to our model, such as those discussed in Sec. VII A, could alleviate the $Sn \gg 2^n$ requirement for local dynamics of the emergent wave function and thus yield a more useful toy model for EmQM. For example, perhaps deviations from quantum mechanics might only be detectable if $Sn \gg 2^{\widetilde{n}}$, where $\widetilde{n}$ is the number of highly entangled qubits that are measured with high fidelity. This hypothesis has not yet been tested experimentally beyond very modest values of $\widetilde{n}$, but might be tested in the future for gradually increasing values of $\widetilde{n}$ by executing deep quantum circuits using quantum computers. Such experiments would significantly constrain the length S of the extra dimension because of the requirement that S be exponential in $\widetilde{n}$. This idea motivated Ref. [5], which proposed to test the validity of quantum mechanics using a Loschmidt echo circuit on many qubits.

B. Bell inequality tests

Any attempt to describe quantum reality in terms of an underlying local classical model faces the potential obstacle that locally realistic classical models conform to Bell inequalities which are known to be experimentally violated. Yet our model of EmQM agrees with quantum mechanics to high accuracy and so can be expected to pass such tests.

One way to understand why Bell inequality violation cannot easily exclude our model is to note that the speed of information propagation among the underlying classical bits, though finite, is much faster than the emergent speed of light on the boundary. This feature makes it exceedingly hard to close the "locality loophole," that is, to rule out communication between Alice's and Bob's labs during the test.

The classical bits carry information at the speed

$$v_{fast} \approx \frac{l_0}{\Delta_t} \approx \frac{l_0}{\Delta_m \delta_t} \frac{N}{Sn}, \tag{83}$$

where l_0 is the spatial distance between bits and where Δ_t [Eq. (50)] is how much time elapses in the EmQM for each discrete time step. If, for example, we insert the ϵ_0 parametrization from Eq. (79) and neglect negligible factors of n, we obtain

$$v_{fast} \sim \frac{l_0}{t_0} N^2 \epsilon_0^{-6}, \tag{84}$$

where $t_0^{-1} \sim ||H_x||_{op}$ is the norm of the local terms in the Hamiltonian (which we previously set to be roughly equal to 1). However, according to local quantum mechanics, all particle velocities (e.g., the speed of light) should be upper bounded by $v_{QM} \sim l_0/t_0$. Therefore, since $v_{fast} \gg v_{QM} \geqslant c$, Bell tests cannot easily detect signatures of our model.

VII. OUTLOOK

In future work, it will be interesting to investigate how generic emergent quantum mechanics (EmQM) is. That is, if

our model is changed slightly, will EmQM still be exhibited? Or do additional ingredients need to be added to our model such that EmQM is a generic result? Or from another point of view, can EmQM be thought of as a highly exotic phase of classical matter? In a sense, research showing that quantum mechanics is "an island in theory space" [68] that can be derived axiomatically [69–71] suggests that EmQM might indeed be a stable fixed point under coarse graining in a broad class of local classical models. It would also be useful to determine if EmQM can result without relying on very small parameters (e.g., δ_t, m_0, and Δ_m).

A. Mitigating nonlocality

As emphasized in Sec. VI A, a crucial remaining future direction is to determine if modifications of our EmQM model could alleviate the $Sn \gg 2^n$ requirement for local EmQM dynamics. We would prefer to have an EmQM model such that $S \gg 2^{\tilde{n}}$ implies consistency with any experiment that only probes $\tilde{n}$ highly entangled qubits with high fidelity, e.g., the logical qubits in a quantum computer. This would be desirable because only $\tilde{n} \sim \log_2 S$ highly entangled qubits would be needed to experimentally test such a model of EmQM, which would be experimentally relevant in the near term if e.g., $S \sim 2^{1000}$.

Nonlocal dynamics when $Sn \ll 2^n$ in our model may result because the stochastic circuits we consider are not very efficient at encoding the wave function. In particular, the emergent wave function is encoded using random permutation matrices. This inefficient encoding could be contrasted with MERA tensor networks [72] or deep neural networks, where each layer can perform a more useful entanglement renormalization [73] or coarse graining [74].

One possible approach to achieve more efficient circuits could be to relax the requirement (25) that the stochastic matrices $M_{s,x}$ are perturbatively close to permutation matrices. But then the subleading (i.e., all but the largest) singular values of $M_{S \leftarrow s}$ [Eq. (56)] will generically be exponentially small in $S - s$. If that occurs, then the overwhelming majority of singular values of $\widetilde{\mathcal{W}}$ [Eq. (58)] will also be extremely small, which will lead to emergent dynamics [Eq. (57)] that do not approximate quantum mechanics well. In this scenario, deep circuits of stochastic matrices are not useful because each layer destroys too much information.

To mitigate this problem, we could consider promoting the classical bits to real numbers. Then the permutation matrices of two bits are promoted to invertible functions from $\mathbb{R}^2$ to $\mathbb{R}^2$, which can be viewed as a permutation of $\mathbb{R}^2$. But unlike permutation matrices, such functions can map the uniform distribution to a different probability distribution. Furthermore, this map can be perfectly inverted. Therefore, unlike deep circuits of generic stochastic matrices, deep circuits of functions do not destroy information. Composing a deep circuit of functions in this way can produce arbitrary probability distributions (and thus arbitrary wave functions for EmQM), an observation which has been utilized within the deep learning community [75,76].

B. Quantum computation and fundamental physics

If quantum mechanics does emerge from classical mechanics, then the computational power of quantum computers could be severely limited [5,6]. For example, BQP-hard problems may only be tractable in actual devices for limited problem sizes. On the other hand, it is possible that deviations from quantum mechanics (such as nonlinear corrections to the Schrödinger equation) could enhance the power of quantum computers [31,77] for some problems of (possibly) limited size.

Even more speculatively, discovering that quantum mechanics emerges from an underlying local classical model might open new directions for understanding dark matter, dark energy, early universe cosmology, and the black hole information paradox [78]. Finally, that our EmQM model encodes the quantum wave function on the boundary of an extra spatial dimension suggests possible connections to holographic duality in quantum gravity [79].

APPENDIX A: MEASUREMENTS

In this Appendix we clarify how measurements could be interpreted in our model within the Everett interpretation of quantum theory. We also speculate how possible improvements upon our EmQM model might lead to a resolution of the measurement problem.

1. Everett interpretation review

In the Everett interpretation, the observer is included in the wave function. In principle, the measurement process can then be formalized as a Hamiltonian evolution via Schrödinger's equation. After the measurement, Schrödinger's equation predicts that the observer becomes entangled with the measured system. That is, the resulting wave function is in a superposition of states, where each state describes one of the measurement outcomes (from the observer's perspective).

For example, consider an observer who measures whether the state of a spin is $|\uparrow\rangle$ or $|\downarrow\rangle$. Before the measurement, suppose that wave function of the spin is

$$|\psi_{\text{before}}\rangle = a|\uparrow\rangle + b|\downarrow\rangle. \tag{A1}$$

In the Everett interpretation, we imagine describing the measurement process using a wave function $|\Psi\rangle$ for the entire universe. Before the measurement, we schematically write the universe's wave function as

$$|\Psi_{\text{before}}\rangle = |\text{observer}\rangle \otimes (a|\uparrow\rangle + b|\downarrow\rangle), \tag{A2}$$

where |observer⟩ is the wave function for the observer (and the rest of the universe other than the spin). After the measurement (assuming it is performed perfectly), the wave function of the universe should be

$$|\Psi_{\text{after}}\rangle = a\,|\text{observer sees }\uparrow\rangle \otimes |\uparrow\rangle$$
$$+\; b\,|\text{observer sees }\downarrow\rangle \otimes |\downarrow\rangle, \quad (\text{A3})$$

where |observer sees ↑⟩ ⊗ |↑⟩ is the wave function of the universe where the observer has observed the |↑⟩ state, and similar for |observer sees ↓⟩ ⊗ |↓⟩.

This exemplifies that in the Everett interpretation, the wave function consists of a superposition of all measurement outcomes. The different states in the superposition are macroscopically different and are consequently extremely unlikely to significantly interfere with each other. These different states effectively behave as different "worlds." Indeed, this interpretation is also often referred to as the *many-worlds interpretation*.

But then how does one predict the probability that an observer will measure a given outcome (from the observer's perspective)? Born's rule implies that the probabilities for the two states, or worlds, in Eq. (A3) are $|a|^2$ and $|b|^2$. But in the Everett interpretation, there are no measurement axioms or applications of Born's rule; only Schrödinger's equation is used to evolve the wave function. In Schrödinger's equation, a and b are just coefficients in a linear expansion, and it is not clear why or how these coefficients should be assigned to probabilities associated with the observer's experience.

Nevertheless, many authors have argued [80–86] that Born's rule is the only reasonable or consistent choice, under various reasonable assumptions. For example, Ref. [85] gives a brief argument that merely assumes unitary invariance, continuity, and system size invariance. Yet it remains controversial whether the Everett interpretation (without measurement axioms) provides a complete description of quantum theory [87].

2. EmQM and the Everett interpretation

In our EmQM model, the slow time dynamics of the boundary degrees of freedom can be accurately predicted using Schrödinger's equation. That is, there is an emergent wave function [Ψ in Eq. (2)] whose time dynamics can be well approximated by Schrödinger's equation. Similar to Everett's interpretation, neither measurement axioms nor Born's rule explicitly appears in our EmQM model. Thus we are left with a similar challenge as in Everett's interpretation: does Born's rule somehow follow from the equations of motion alone?

But for a model of EmQM from classical mechanics, we would not expect the emergent wave function to be able to store very many different worlds in superposition. Instead, there must eventually be some sort of "collapse" of the emergent wave function to only a subset of the superposed measurement outcomes. Importantly, this collapse of states should be (at least approximately) consistent with Born's rule if the EmQM model closely approximates standard quantum theory. Perhaps an EmQM model with this property could provide a satisfying solution to the measurement problem.

Unfortunately, our EmQM model does not appear to be successful enough to study this hypothetical wave function collapse. As explained in Sec. VI A, in a certain limit with an extremely large extra dimension, our model simply reproduces Schrödinger's equation. Therefore this limit is similar to the Everett interpretation, and it is not obvious how the measurement probabilities for an observer should be assigned, although some of the arguments in Refs. [80–86] might still be applicable. If the extra dimension isn't sufficiently large, then there are nonlocal violations of Schrödinger's equation, but we have no reason to expect these violations to imply an approximate version of Born's rule (although we have not checked thoroughly). It seems that better models are needed to assess whether the Born rule, in addition to Schrödinger's equation, could arise from a sensible classical model of EmQM.

APPENDIX B: REAL-VALUED QUANTUM MECHANICS

In this Appendix we show that any quantum Hamiltonian and wave function in Schrödinger's equation can be linearly mapped to real-valued analogs with zero row and column sums that satisfy Eqs. (3) and (4) while preserving locality. We do this by first mapping to real-valued quantum mechanics [88–91] in Appendix B 1 and then focus on zero sums in Appendix B 2.

Both mappings generically require adding additional qubits. Preserving locality requires multiplying the qubit count by a constant factor (when using the systematic mapping). As a result, the possible real-valued wave functions that result from this mapping are highly constrained in the sense that these wave functions only span a subset of the Hilbert space. Therefore, although there are well-known fundamental differences between complex and real-valued quantum mechanics [68,69,92–96], this mapping shows that complex-valued quantum mechanics is equivalent to real-valued quantum mechanics constrained to a *subspace* of the Hilbert space. However in Appendix B 1c, we emphasize that in real-valued quantum mechanics, it is not correct to assume that "the state representing two independent preparations of the two systems is the tensor product of the two preparations" [92].

1. Mapping from complex to real QM

a. Geometrically nonlocal mapping

If we do not require that the mapping is geometrically local, then mapping to real values can be achieved simply by splitting complex numbers into their real and imaginary parts [88–90]. This can be achieved for operators via the replacement

$$i \rightarrow -i\sigma^2, \quad (\text{B1})$$

where

$$\sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (\text{B2})$$

is a Pauli operator acting on an additional qubit.

We note that

$$|\pm i\rangle = \tfrac{1}{\sqrt{2}}(|\uparrow\rangle \pm i\,|\downarrow\rangle) \quad (\text{B3})$$

are eigenstates of σ^2 with eigenvalues ± 1, and that

$$P(\pm) = \tfrac{1}{2}(\mathbb{1} \pm \sigma^2) \qquad (B4)$$

are orthogonal projectors onto these eigenstates. Furthermore, $|+i\rangle$ and $|-i\rangle$ are complex conjugates of one another, as are $P(+)$ and $P(-)$. We map an n-qubit wave function $|\psi\rangle$ to a real $(n+1)$-qubit wave function $|\widetilde{\psi}\rangle$ according to

$$|\psi\rangle \rightarrow |\widetilde{\psi}\rangle = \tfrac{1}{\sqrt{2}}(|\psi\rangle \otimes |-i\rangle + |\psi\rangle^* \otimes |+i\rangle), \qquad (B5)$$

and map an n-qubit operator Q to a real $(n+1)$-qubit operator $\widetilde{Q}$ according to

$$Q \rightarrow \widetilde{Q} = Q \otimes P(-) + Q^* \otimes P(+), \qquad (B6)$$

where the asterisk (*) denotes complex conjugation; thus

$$\widetilde{Q}|\widetilde{\psi}\rangle = \tfrac{1}{\sqrt{2}}(Q|\psi\rangle \otimes |-i\rangle + Q^*|\psi\rangle^* \otimes |+i\rangle). \qquad (B7)$$

Under this mapping, Schrödinger's equation $\partial_t |\psi(t)\rangle = -iH |\psi(t)\rangle$ is equivalent to

$$\partial_t |\widetilde{\psi}(t)\rangle = -(\widetilde{iH}) |\widetilde{\psi}(t)\rangle. \qquad (B8)$$

That is, it is mapped to Schrödinger's equation for the real wave function $|\widetilde{\psi}\rangle$ with imaginary Hamiltonian $-i(\widetilde{iH})$, where $\widetilde{iH}$ denotes the result of mapping $Q = iH$ using Eq. (B6).

Another way to express the mapping is sometimes convenient. Note that if Q is real, then

$$\widetilde{Q} = Q \otimes \mathbb{1}, \qquad (B9)$$

while if Q is imaginary, then

$$\widetilde{Q} = (-iQ) \otimes (-i\sigma^2); \qquad (B10)$$

more generally

$$\widetilde{Q} = \mathrm{Re}Q \otimes \mathbb{1} + \mathrm{Im}Q \otimes (-i\sigma^2). \qquad (B11)$$

Similarly, if $|\psi\rangle$ is real in some basis, then

$$|\widetilde{\psi}\rangle = |\psi\rangle \otimes |\uparrow\rangle, \qquad (B12)$$

while if $|\psi\rangle$ is imaginary, then

$$|\widetilde{\psi}\rangle = -i|\psi\rangle \otimes |\downarrow\rangle; \qquad (B13)$$

and more generally

$$|\widetilde{\psi}\rangle = |\mathrm{Re}\psi\rangle \otimes |\uparrow\rangle + |\mathrm{Im}\psi\rangle \otimes |\downarrow\rangle. \qquad (B14)$$

b. Local mapping

Now suppose that $H = \sum_x H_x$ is a local Hamiltonian, where each H_x acts only on qubits near the spatial point x. If H_x and H_y act on sets of qubits that are distantly separated from one another, then it's not possible to put the additional qubit close to both sets. Therefore, the above operator map does not preserve the geometric locality of the Hamiltonian.

In order to promote the nonlocal mapping to a mapping that preserves locality for local operators Q_x, we can instead add a new qubit adjacent to each lattice site and map

$$i \rightarrow -i\sigma_x^2 \qquad (B15)$$

at each site. Wave functions are then mapped according to

$$|\psi\rangle \rightarrow |\widetilde{\psi}\rangle = \tfrac{1}{\sqrt{2}}(|\psi\rangle \underset{x}{\otimes} |-i\rangle_x + |\psi\rangle^* \underset{x}{\otimes} |+i\rangle_x). \qquad (B16)$$

Now suppose that Q_x is a k-local operator (e.g., a term in the Hamiltonian) acting on k qubits near x. After the mapping, the site x is accompanied by an adjacent auxiliary qubit, and Q_x is mapped to

$$\widetilde{Q}_x = Q_x \otimes P(-)_x + Q_x^* \otimes P(+)_x, \qquad (B17)$$

a $(k+1)$-local operator acting on the k qubits together with the auxiliary qubit adjacent to site x. Another way to express the mapping is

$$\widetilde{Q}_x = \mathrm{Re}Q_x \otimes \mathbb{1}_x + \mathrm{Im}Q_x \otimes (-i\sigma_x^2). \qquad (B18)$$

Thus the k-local Hamiltonian $H = \sum_x H_x$ can be mapped to the $(k+1)$-local Hamiltonian $\widetilde{H} = \sum_x \widetilde{H}_x$.

As we found for the nonlocal mapping, Schrödinger's equation $\partial_t |\psi(t)\rangle = -iH |\psi(t)\rangle$ is equivalent to to Schrödinger's equation for the real wave function $|\widetilde{\psi}\rangle$ with imaginary Hamiltonian

$$-i(\widetilde{iH}) = \sum_x \left[i\mathrm{Im}(H_x) \otimes \mathbb{1}_x - \mathrm{Re}(H_x) \otimes \sigma_x^2 \right]. \qquad (B19)$$

c. Dynamics

Here we briefly review some dynamical properties of real-valued quantum mechanics. We want iH to be real such that no imaginary values appear in Schrödinger's equation, which implies that H must be imaginary and antisymmetric. Thus for each eigenvector $|E\rangle$, taking the complex conjugate of $H|E\rangle = E|E\rangle$ (in any basis) implies that the complex-conjugated $|E\rangle^*$ has eigenvalue $-E$. Therefore, eigenvectors with nonzero eigenvalues come in complex-conjugate pairs with negated eigenvalues. It is convenient to consider real-valued linear combinations $|E\rangle_\pm = \frac{1}{\sqrt{\pm 2}}(|E\rangle \pm |-E\rangle)$ for each $E > 0$. The coefficients $a_{E,\pm}(t)$ of a wave function $|\psi(t)\rangle = \sum_{E,\pm} a_{E,\pm}(t) |E\rangle_\pm$ obey the time evolution $\partial_t a_{E,\pm}(t) = \mp E a_{E,\mp}(t)$ for each $E > 0$.

A real-valued wave function can not be in an eigenstate $|E\rangle$ with nonzero E since $|E\rangle$ is complex valued when $E \neq 0$. In complex-valued quantum mechanics, eigenstates are steady states, i.e., states that do not change with time. The analog of steady states in real-valued quantum mechanics are oscillating superpositions of $|E\rangle_\pm$:

$$|\psi(t)\rangle = \cos(Et) |E\rangle_+ - \sin(Et) |E\rangle_-$$
$$= \tfrac{1}{\sqrt{2}} e^{-iEt} |E\rangle + \tfrac{1}{\sqrt{2}} e^{+iEt} |-E\rangle. \qquad (B20)$$

The analog of the lowest-energy ground state is the above state with the largest possible E. Notably, this "steady state" looks like a cat state from the perspective of complex-valued quantum mechanics. Indeed, the steady state of two disconnected subsystems will share an entangled qubit from the perspective of complex-valued quantum mechanics. For example, if $H = -Y_1 - Y_2$ is the Hamiltonian for two qubits, then the lowest-energy steady state is not a tensor product state; instead, it is maximally entangled:

$$|\psi(t)\rangle = [\cos(t) |\uparrow\rangle - \sin(t) |\downarrow\rangle] \otimes [\cos(t) |\uparrow\rangle - \sin(t) |\downarrow\rangle]$$
$$- [\sin(t) |\uparrow\rangle + \cos(t) |\downarrow\rangle] \otimes [\sin(t) |\uparrow\rangle + \cos(t) |\downarrow\rangle]. \qquad (B21)$$

In complex-valued quantum mechanics, the following tensor product axiom holds for combining quantum states of two

systems: "the state representing two independent preparations of the two systems is the tensor product of the two preparations" [92]. However, this axiom does not hold in real-valued quantum mechanics. Indeed, Refs. [93,94] have verified that real-valued quantum mechanics with this tensor product axiom is not consistent with experiment. Nevertheless, our analysis above shows that real-valued quantum mechanics is consistent with complex-valued quantum mechanics if the tensor product axiom is dropped. Indeed the tensor product axiom is not consistent with our local mapping (B16), which maps complex-valued tensor product wave functions to a sum of two tensor products. Furthermore, Eq. (B21) shows that a sum of two tensor products arises naturally for ground-state wave functions of two disconnected subsystems in real-valued quantum mechanics.

2. Zero sum

To satisfy the zero-sum conditions (3) and (4), we once again introduce an additional qubit. We map the wave function $|\psi\rangle$ to

$$|\widetilde{\psi}\rangle = |\psi\rangle \otimes |-\rangle, \tag{B22}$$

where

$$|-\rangle = \tfrac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle). \tag{B23}$$

Similarly, any operator Q can be linearly mapped to a new operator

$$\widetilde{Q} = Q \otimes |-\rangle\langle-|, \tag{B24}$$

where

$$|-\rangle\langle-| = \tfrac{1}{2}(|\uparrow\rangle\langle\uparrow| - |\uparrow\rangle\langle\downarrow| - |\downarrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|). \tag{B25}$$

By construction, $|\widetilde{\psi}\rangle$ and $\widetilde{Q}$ obey the zero-sum conditions:

$$\sum_i \widetilde{\psi}_i = 0,$$
$$\sum_i \widetilde{Q}_{ij} = 0 = \sum_j \widetilde{Q}_{ij} \tag{B26}$$

in any basis where the additional qubit has basis vectors $|\uparrow\rangle$ and $|\downarrow\rangle$. Furthermore, the zero-sum constraints are preserved under evolution under the Schrödinger equation for a Hamiltonian that obeys the zero-sum constraints.

In order to preserve geometric locality, we add another set of new qubits adjacent to each x. The mapping for wave functions and local operators (including terms in the Hamiltonian) is

$$|\psi\rangle \to |\widetilde{\psi}\rangle = |\psi\rangle \otimes_x |-\rangle_x, \tag{B27}$$

$$Q_x \to \widetilde{Q}_x = Q_x \otimes |-\rangle_x \langle-|_x. \tag{B28}$$

3. Examples
a. Ising (systematic mapping)

As an example, we can consider applying these systematic mappings to a transverse-field Ising model Hamiltonian

$H = \sum_x H_x$:

$$\begin{aligned} H_x &= -JZ_xZ_{x+1} - hY_x \\ &\to JZ_xZ_{x+1} \otimes \sigma_x^2 - hY_x \\ &\to \left(JZ_xZ_{x+1} \otimes \sigma_x^2 - hY_x\right) \otimes |-\rangle_x\langle-|_x. \end{aligned} \tag{B29}$$

X_x, Y_x, Z_x are Pauli operators in the Ising model Hamiltonian, while σ_x^2 is a Pauli operator acting on additional qubits. $|-\rangle_x\langle-|_x$ acts on an additional set of qubits. A transverse hY_x term was considered in the first line instead of the traditional hX_x term to make the example more useful. The second line is the result of applying Eq. (B19) to obtain an imaginary-valued Hamiltonian. Eq. (B28) is applied to obtain the third line, for which the resulting Hamiltonian has zero row and column sums.

b. Ising (clever duality)

However, adding additional qubits is not always necessary. For example, the XY Hamiltonian is dual to an imaginary-valued Hamiltonian with zero row and column sums:

$$H_x = -J_{XY}(X_xX_{x+1} + Y_xY_{x+1}) \tag{B30}$$

$$\leftrightarrow J_{XY}(X_{x-1}Y_xX_{x+1} - Y_x). \tag{B31}$$

The second Hamiltonian describes the phase transition between a SPT cluster state [97] and the trivial disordered phase.

The duality mapping used above is

$$\begin{aligned} Y_xY_{x+1} &\leftrightarrow Y_x \\ Z_x &\leftrightarrow X_{x-1}X_x, \end{aligned} \tag{B32}$$

which implies that $X_xX_{x+1} \leftrightarrow -X_{x-1}Y_xX_{x+1}$. This duality only maps symmetric operators to local operators. Symmetric operators are operators that commute with certain Z_2 symmetries, which are $\prod_x Z_x$ and $\prod_x Y_x$ for the respective left and right sides of the duality. The duality maps the Z_2 symmetries to the identity (ignoring boundary conditions):

$$\begin{aligned} \prod_x Z_x &\leftrightarrow 1, \\ 1 &\leftrightarrow \prod_x Y_x. \end{aligned} \tag{B33}$$

These properties are common in duality mappings, such as the self-duality [98] of the transverse-field Ising model.

APPENDIX C: EMERGENT LORENTZ INVARIANCE

In Sec. II A we had to posit that the observed Lorentz invariance in our universe is emergent (rather than exact). This may not be a major hurdle since emergent Lorentz invariance has been shown to be a stable fixed point under the renormalization group (RG) in several strongly coupled models [99–101]. Indeed, emergent Lorentz invariance is rather ubiquitous in low-energy physics. For example, emergent Lorentz invariance occurs in materials such as graphene, which exhibits an electron band structure with a Lorentz-invariant Dirac cone at low energy. As a result, the electrons in graphene experience a Lorentz-invariant speed limit that is much smaller than the speed of photons.

But with multiple species of fermions, emergent Lorentz invariance would require that all Dirac fermions have the same velocity. Although the velocities for different species flow to the same value under RG, the flow is very slow; hence emergent Lorentz invariance seems to require fine tuning in order to be consistent with certain very precise experiments. [102,103] Furthermore, general relativity would likely also need to be emergent in this scenario [104,105].

It is not clear if including the effects of a fully quantized theory of emergent quantum gravity could significantly affect the RG flow such that Lorentz invariance emerges more rapidly. In particular, we note that when Lorentz invariance is broken, the local diffeomorphism "gauge symmetry" of general relativity is broken. This reminds us of $(3 + 1)$-dimensional $U(1)$ gauge theory, in which we could imagine adding an $A^2 = A_\mu A^\mu$ term to the Lagrangian, which breaks the gauge symmetry. However, although A^2 naively appears to be relevant in $(3 + 1)$ dimensions since A^2 has energy dimension 2, A^2 is actually an irrelevant perturbation. In fact, all perturbations to $(3 + 1)$-dimensional $U(1)$ gauge theory are irrelevant, including the perturbations that break gauge invariance [106].

We emphasize that although these gauge-symmetry-breaking terms are irrelevant for the compact gauge group $U(1)$, such terms are *relevant* (as naively expected) for the noncompact gauge group $\mathbb{R}$. Note that both of these gauge groups lead to the same classical equations of motion since they have the same Lie algebra. It is remarkable that it is only after quantizing these two different gauge theories, with gauge groups $U(1)$ or $\mathbb{R}$, that we discover that gauge invariance can be emergent in $U(1)$ gauge theory but not in $\mathbb{R}$. With this in mind, we speculate that emergent Lorentz invariance may occur sufficiently rapidly in some theories of emergent gravity such that it could be consistent with experiments. [107]

One approach to intuitively understanding why A^2 is irrelevant in $U(1)$ gauge theory is to consider the corresponding lattice gauge theory (without constraining the Hilbert space to the gauge-invariant states). The lattice gauge theory is similar to the toric code [108], except the Z_2 qubits are replaced by integer-valued degrees of freedom $E_e \in \mathbb{Z}$ on each edge e of the lattice. The Hamiltonian is $H = \sum_i (\nabla_i n)^2 - \sum_p \cos(\nabla_p \times A)$. $\nabla_i n$ denotes the lattice divergence of n_e centered at the vertex i, while $\nabla_p \times A$ is the lattice curl of A_e around the plaquette p. The commutation relations are $[e^{iA_e}, n_{e'}] = \delta_{e,e'}$. The states with $\nabla_i n = 0$ obey Gauss's law and are the gauge-invariant states.

Gauge invariance is often imposed by taking $\nabla_i n = 0$ to be a constraint on the Hilbert space. Here we do not constrain the Hilbert space; instead we use the first term of the Hamiltonian to impose an energy penalty on states that are not gauge invariant. Since n_e is integer valued, all energy excitations of $\sum_i (\nabla_i n)^2$ cost *finite energy*, which makes the gauge invariance stable to perturbations. For example, a ϵA^2 term in the $U(1)$ gauge theory Lagrangian is analogous to a $\epsilon \sum_e \cos(A_e)$ term in the lattice gauge theory Hamiltonian. If ϵ is sufficiently small, this term does not does not lead to confinement since its excitations cost finite energy due to the $(\nabla_i n)^2$ term. This can be shown more formally using degenerate perturbation theory [109]. Similarly, in the toric

code (i.e., Z_2 lattice gauge theory where Gauss's law enters the Hamiltonian as an energy penalty rather than a Hilbert space constraint), arbitrary perturbations are irrelevant and do not destabilize the emergent gauge invariance or topological order. [108]

We emphasize that even if there are faster-than-light degrees of freedom, it could still be very difficult for observers to send information faster than light. Suppose a Hamiltonian H is Lorentz invariant (either exactly or approximately at low energies) with velocity c, and suppose an observer living within the wave function Ψ wants to use the fast classical degrees of freedom to send signals faster than c (without exceeding low energies if the Lorentz-invariance is approximate). If $\Psi(t)$ is well described by Schrödinger's equation for a local Hamiltonian, then the observer will only be able to send signals much faster than c by taking advantage of very small possible violations of Schrödinger's equation [15,16].

APPENDIX D: APPROXIMATE SIMULATION

As noted in Eq. (82), simulating the EmQM model is very CPU intensive. Fortunately, we can approximately simulate the EmQM model significantly faster by "integrating out" the fast degrees of freedom (i.e., the forward and backward-propagating bits) so that we only have to directly simulate the slow degrees of freedom (i.e., the stochastic matrices). This will allow us to simulate $\Delta_{\text{jump}} \gg S$ many time steps all at once with negligible error.

To do this, we approximate the stochastic matrices $M_{s,x}$ as constant over Δ_{jump} many time steps. We then estimate how much the stochastic matrices might change after these Δ_{jump} time steps. If $\Delta_{\text{jump}} \gg N = 2^n$, then each bit string will occur many times over this many time steps. Therefore, simulating the propagation of Δ_{jump} many bit strings involves a lot of duplication of effort. To speed up the simulations, we can instead just calculate how much each stochastic matrix will be affected by the propagation of all N possible bit strings, each weighted by its probability times Δ_{jump}.

1. Approximate algorithm

The approximate algorithm to compute how Δ_{jump} time steps could affect the stochastic matrices is as follows.

We first compute $\boldsymbol{P}_S^{(\tau)}$ using Eq. (23) to get the probability vector for the output bits a_S of the stochastic circuit.

Then for each $\gamma = 1, 2$, we sample an $N \times N$ matrix $\beta_{S,\gamma}^{(\tau)}$ from the multinomial distribution of Δ_{jump} trials with a matrix of probabilities $(B^{(+\gamma)} \boldsymbol{P}_S^{(\tau)}) \otimes (B^{(-\gamma)} \boldsymbol{P}_S^{(\tau)})$, where $B^{(\gamma)}$ is the $N \times N$ stochastic matrix defined in Eq. (18). Therefore, $\beta_{S,\gamma}^{(\tau)}(b_{S,+\gamma}, b_{S,-\gamma})$ counts how many times the pair $(b_{S,+\gamma}, b_{S,-\gamma})$ of bit strings could have occurred after Δ_{jump} time steps. Note that we approximated $M_{s,x}^{(\tau)}$ as constant over these Δ_{jump} time steps.

Next we back-propagate [analogous to Eq. (30)] $\beta_{S,\gamma}^{(\tau)}$ using the permutation matrices Q_s defined in Eq. (26):

$$\beta_{s,\gamma}^{(\tau)} = Q_s^T \beta_{s+1,\gamma}^{(\tau)} Q_s. \tag{D1}$$

Finally, we wish to update the perturbations $m_{s,x}$ in accordance with Eq. (31) using $\beta_{S,\gamma}^{(\tau)}$. To do this, we modify

Eq. (31) to

$$m_{s,x}^{(\tau+\Delta_{\text{jump}})} = m_{s,x}^{(\tau)} + \Delta_{\text{m}} \sum_{b_{\mathbb{x}}^+} \sum_{b_{\mathbb{x}}^-} \sum_{e_{\mathbb{x}}} \left[\sum_{\gamma=1,2} \beta_{s,x,\gamma}^{(\tau)}(b_{\mathbb{x}}^+, b_{\mathbb{x}}^-) \, p_{s,x}^{(\tau)}(e_{\mathbb{x}}|b_{\mathbb{x}}^-)(\hat{\boldsymbol{b}}_{\mathbb{x}}^+ - \hat{\boldsymbol{b}}_{\mathbb{x}}^-) \otimes \hat{\boldsymbol{e}}_{\mathbb{x}} \right], \tag{D2}$$

where $b_{\mathbb{x}}^+$, $b_{\mathbb{x}}^-$, and $e_{\mathbb{x}}$ are each a pair of bits (00, 01, 10, or 11). These bit pairs index the 4×4 matrices $m_{s,x}^{(\tau)}$ and $\beta_{s,x,\gamma}^{(\tau)}$. The bit pairs also determine the basis 4-vectors $\hat{\boldsymbol{b}}_{\mathbb{x}}^{\pm}$ and $\hat{\boldsymbol{e}}_{\mathbb{x}}$; e.g., $\hat{\boldsymbol{b}}_{\mathbb{x}}^+ = (1, 0, 0, 0)$ if $b_{\mathbb{x}}^+ = 00$. Above, $\beta_{s,x,\gamma}^{(\tau)}(b_{\mathbb{x}}^+, b_{\mathbb{x}}^-)$ counts how many times the pair $(b_{\mathbb{x}}^+, b_{\mathbb{x}}^-)$ could have occurred over the Δ_{jump} time steps. That is,

$$\beta_{s,x,\gamma}^{(\tau)} = \sum_{b^+} \sum_{b^-} \hat{\boldsymbol{b}}_{\mathbb{x}}^+ \otimes \hat{\boldsymbol{b}}_{\mathbb{x}}^- \, \beta_{s,\gamma}^{(\tau)}(b^+, b^-), \tag{D3}$$

where $\sum_{b^{\pm}}$ sums over the $N = 2^n$ different bit strings of n bits, and $\hat{\boldsymbol{b}}_{\mathbb{x}}^{\pm}$ is the basis 4-vector that depends on bits x and $x+1$ of the bit string $b^{\pm}$; e.g., $\hat{\boldsymbol{b}}_{\mathbb{x}}^+ = (1, 0, 0, 0)$ [or $(0,1,0,0)$] if bits x and $x+1$ of $b^{\pm}$ are 00 [or 01]. In Eq. (D2), $p_{s,x}^{(\tau)}(e_{\mathbb{x}}|b_{\mathbb{x}}^-)$ is a conditional probability distribution for $e_{\mathbb{x}}$ given $b_{\mathbb{x}}^-$.

We need to calculate $p_{s,x}^{(\tau)}(e_{\mathbb{x}}|b_{\mathbb{x}}^-)$ such that it is approximately equal to the probability that $\hat{\boldsymbol{e}}_{s-1,x,\gamma}^{(\tau)} = \hat{\boldsymbol{e}}_{\mathbb{x}}$ given $\hat{\boldsymbol{b}}_{s,\mathbb{x},-\gamma}^{(\tau)} = \hat{\boldsymbol{b}}_{\mathbb{x}}^-$ in Eq. (31) over the Δ_{jump} time steps. Recall that $\hat{\boldsymbol{e}}_{s-1,x,\gamma}^{(\tau)}$ is chosen uniformly at random from the set of basis 4-vectors that keep $M_{s,x}^{(\tau)} = Q_{s,x} + m_{s,x}^{(\tau)}$ nonnegative. Therefore we can take $p_{s,x}^{(\tau)}(e_{\mathbb{x}}|b_{\mathbb{x}}^-) = 1/4$ to be uniform probabilities as long as this results in a stochastic matrix $M_{s,x}^{(\tau+\Delta_{\text{jump}})}$ with nonnegative entries.

However, if $M_{s,x}^{(\tau+\Delta_{\text{jump}})}$ has negative entries, then $p_{s,x}^{(\tau)}$ can not be uniform. Instead, for each negative $M_{s,x}^{(\tau+\Delta_{\text{jump}})}(b_{\mathbb{x}}^-, e_{\mathbb{x}}) < 0$ entry that we find, we must make

$p_{s,x}^{(\tau)}(e_{\mathbb{x}}|b_{\mathbb{x}}^-)$ a free parameter [while uniformly adjusting the other entries such that $p_{s,x}^{(\tau)}(e_{\mathbb{x}}|b_{\mathbb{x}}^-)$ is a probability distribution for $e_{\mathbb{x}}$ given $b_{\mathbb{x}}^-$]. We then solve for these free parameters such that the previously negative entries of $M_{s,x}^{(\tau+\Delta_{\text{jump}})}$ are zero. If $M_{s,x}^{(\tau+\Delta_{\text{jump}})}$ for the new $p_{s,x}^{(\tau)}$ has additional negative entries, then we repeat the procedure until all entries are nonnegative. That is, we make more entries in $p_{s,x}^{(\tau)}$ free parameters and resolve for the old and new free parameters such that any previously negative entry of $M_{s,x}^{(\tau+\Delta_{\text{jump}})}$ is zero. At most, $p_{s,x}^{(\tau)}$ can only have 4×3 free parameters, since it describes four different probability distributions, each with up to three free parameters. Therefore, the above procedure can not repeat more than 12 times. As desired, this procedure results in a $p_{s,x}^{(\tau)}$ that is approximately equal to the probability that $\hat{\boldsymbol{e}}_{s-1,x,\gamma}^{(\tau)} = \hat{\boldsymbol{e}}_{\mathbb{x}}$ given $\hat{\boldsymbol{b}}_{s,\mathbb{x},-\gamma}^{(\tau)} = \hat{\boldsymbol{b}}_{\mathbb{x}}^-$ in Eq. (31) over the Δ_{jump} time steps.

2. Approximation error

The approximate algorithm makes the approximation that $M_{s,x}^{(\tau)}$ is constant over Δ_{jump} many time steps. This approximation is very similar to the cause of the deviation $\varepsilon_{\text{delay}}(t)$ [Eq. (75)] that results from the $O(S)$ discrete time delay for the classical bits to move through the circuit. The primary difference is that the delay is Δ_{jump} instead of $O(S)$ for the approximate algorithm. Therefore, the approximate algorithm results in an error similar to Eq. (75), except a factor of S is
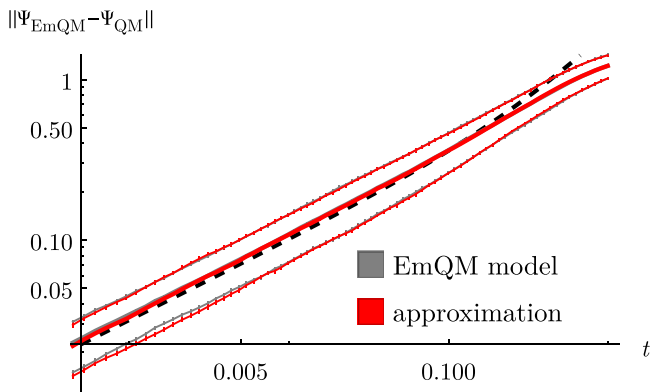

FIG. 4. Deviation $||\Psi(t) - \Psi_{\text{QM}}(t)||$ vs time t for $n = 4$ qubits with $\epsilon_0 = 1$, showing agreement between the EmQM model (gray), approximate algorithm with $\epsilon_{\text{j}} = 0.02$ (red), and estimated $\varepsilon(t)$ from Eq. (78) (dashed black). For the EmQM model and approximate algorithm, we average over 1000 random realizations and plot the mean (thick lines) along with the mean plus or minus the standard deviation (thin lines) to demonstrate that both the mean and statistical fluctuations agree well. Error bars denote one standard deviation of statistical error resulting from the finite number of 1000 samples.
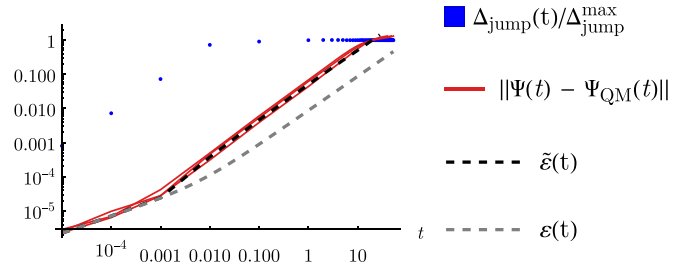

FIG. 5. Deviation $||\Psi(t) - \Psi_{\text{QM}}(t)||$ (red) vs time t calculated using the approximate algorithm with a large $\varepsilon_{\text{jump}} = 10$ so that we can study errors due to the approximate simulation algorithm. Simulations use $n = 4$ qubits and three different random initializations with $\varepsilon = 0.01$ [in Eq. (79)]. In all simulations in this work, Δ_{jump} is chosen to be as large as possible without exceeding $\Delta_{\text{jump}}^{\text{max}}$. Δ_{jump} is sometimes limited (shown by blue dots) when the time step is very small (due to the logarithmic time axis). By inserting the actual time-dependent $\Delta_{\text{jump}}(t)$ into Eq. (D4), our estimate $\tilde{\varepsilon}(t)$ [(D7), dashed black line] for the deviation (under the influence of the approximation) matches the $\varepsilon_{\text{jump}}(t)$-dominated simulation data (red) remarkably well. In contrast, the dashed gray curve shows the expected deviation $\varepsilon(t)$ of the EmQM model without approximation, which is significantly smaller when $0.01 \lesssim t \lesssim 10$. This validates Eq. (D7).

replaced by Δ_{jump}:

$$\varepsilon_{\text{jump}}(t) \sim \frac{\Delta_{\text{jump}}}{S}\varepsilon_{\text{delay}}(t)$$
$$\sim n\Delta_t\Delta_{\text{jump}}\,t. \tag{D4}$$

In our simulations, we want to choose Δ_{jump} to be as large as possible without introducing noticeable errors in our plots. To achieve this, we choose Δ_{jump} such that $\varepsilon_{\text{jump}}(t)$ in Eq. (D4) is parameterized by a new control parameter ϵ_{j}:

$$\varepsilon_{\text{jump}}(t) \sim \epsilon_0\,\epsilon_{\text{j}}\,t. \tag{D5}$$

Small ϵ_{j} will then ensure that $\varepsilon_{\text{jump}}(t)$ contributes negligibly to $\varepsilon(t) = ||\Psi - \Psi_{\text{QM}}||$ in comparison to the estimated $\varepsilon(t) \sim \sqrt{\epsilon_0 t}$ in Eq. (78). Therefore, Eq. (D4) implies that we must limit Δ_{jump} to be no larger than

$$\Delta_{\text{jump}}^{\max} \approx \frac{\epsilon_0\,\epsilon_{\text{j}}}{\Delta_t n}$$
$$\sim \frac{n^3 N^2}{\epsilon_0^5}\epsilon_{\text{j}}. \tag{D6}$$

Due to this additional source of error, the approximate algorithm will produce slightly larger deviations from quantum mechanics, which we estimate to be

$$\widetilde{\varepsilon}(t) \sim \sqrt{\varepsilon(t)^2 + \varepsilon_{\text{jump}}(t)^2}. \tag{D7}$$

To ensure that the error $\varepsilon_{\text{jump}}(t)$ from the approximate algorithm remains negligible, we use small $\epsilon_{\text{j}} = 0.02$ for most simulations (excluding Fig. 5). The only exception is the $\epsilon_0 = 0.05$ data in Figs. 2(b) and 2(d), which used $\epsilon_{\text{j}} = 0.1$ in order to keep the simulation time under one month.

To validate that our approximate algorithm does not significantly affect the deviation $||\Psi - \Psi_{\text{QM}}||$ or its statistical fluctuations when $\varepsilon_{\text{jump}}$ is small, in Fig. 4 we compare these quantities for the EmQM model and the approximate algorithm. To validate Eq. (D7), in Fig. 5 we plot the deviation $||\Psi(t) - \Psi_{\text{QM}}(t)||$ from quantum mechanics using the approximate model with large ϵ_{j} such that the error from approximation dominates.

[1] S. Aaronson and L. Chen, Complexity-theoretic foundations of quantum supremacy experiments, arXiv:1612.05903 (2016).

[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[3] M. Arndt and K. Hornberger, Testing the limits of quantum mechanical superpositions, Nat. Phys. **10**, 271 (2014).

[4] D. Aharonov and U. Vazirani, Is quantum mechanics falsifiable? A computational perspective on the foundations of quantum mechanics, arXiv:1206.3686 (2012).

[5] K. Slagle, Testing quantum mechanics using noisy quantum computers, arXiv:2108.02201 (2021).

[6] G. 't Hooft, The cellular automaton interpretation of quantum mechanics, arXiv:1405.1548 (2014).

[7] S. K. Lamoreaux, A review of the experimental tests of quantum mechanics, Int. J. Mod. Phys. A **07**, 6691 (1992).

[8] E. S. Fry and T. Walther, Fundamental tests of quantum mechanics, Adv. Atomic Mol. Opt. Phys. **42**, 1 (2000).

[9] A. Keshavarzi, K. S. Khaw, and T. Yoshioka, Muon g–2: A review, Nucl. Phys. B **975**, 115675 (2022).

[10] P. Shadbolt, J. C. F. Mathews, A. Laing, and J. L. O'Brien, Testing foundations of quantum mechanics with photons, Nat. Phys. **10**, 278 (2014).

[11] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[12] Q. Zhu *et al.*, Quantum computational advantage via 60-qubit 24-cycle random circuit sampling, Science Bulletin **67**, 240 (2022).

[13] H.-S. Zhong, J.-W. Pan *et al.*, Quantum computational advantage using photons, Science **370**, 1460 (2020).

[14] M. F. Pusey, J. Barrett, and T. Rudolph, On the reality of the quantum state, Nat. Phys. **8**, 475 (2012).

[15] J. Polchinski, Weinberg's Nonlinear Quantum Mechanics and the Einstein-Podolsky-Rosen Paradox, Phys. Rev. Lett. **66**, 397 (1991).

[16] N. Gisin, Weinberg's non-linear quantum mechanics and supraluminal communications, Phys. Lett. A **143**, 1 (1990).

[17] M. J. Bremner, R. Jozsa, and D. J. Shepherd, Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy, Proc. R. Soc. London A **467**, 459 (2011).

[18] S. Aaronson, A. Bouland, L. Chua, and G. Lowther, ψ-epistemic theories: The role of symmetry, Phys. Rev. A **88**, 032111 (2013).

[19] B. Mielnik, Generalized quantum mechanics, Commun. Math. Phys. **37**, 221 (1974).

[20] D. Hanneke, S. Fogwell, and G. Gabrielse, New Measurement of the Electron Magnetic Moment and the Fine Structure Constant, Phys. Rev. Lett. **100**, 120801 (2008).

[21] S. Aaronson, Multilinear formulas and skepticism of quantum computing, in *Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '04 (Association for Computing Machinery, New York, 2004), pp. 118–127.

[22] S. Lee, J. Lee, H. Zhai, Y. Tong, A. M. Dalzell, A. Kumar, P. Helms, J. Gray, Z.-H. Cui, W. Liu *et al.*, Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry, Nat Commun **14**, 1952 (2023).

[23] A. Milsted, J. Liu, J. Preskill, and G. Vidal, Collisions of false-vacuum bubble walls in a quantum spin chain, PRX Quantum **3**, 020316 (2022).

[24] J. M. Deutsch, Eigenstate thermalization hypothesis, Rep. Prog. Phys. **81**, 082001 (2018).

[25] M. Srednicki, Chaos and quantum thermalization, Phys. Rev. E **50**, 888 (1994).

[26] C. D. White, M. Zaletel, R. S. K. Mong, and G. Refael, Quantum dynamics of thermalizing systems, Phys. Rev. B **97**, 035127 (2018).

[27] E. Leviatan, F. Pollmann, J. H. Bardarson, D. A. Huse, and E. Altman, Quantum thermalization dynamics with matrix-product states, arXiv:1702.08894 (2017).

[28] T. Rakovszky, C. W. von Keyserlingk, and F. Pollmann, Dissipation-assisted operator evolution method for capturing hydrodynamic transport, Phys. Rev. B **105**, 075131 (2022).

[29] A. Bassi, K. Lochan, S. Satin, T. P. Singh, and H. Ulbricht, Models of wave-function collapse, underlying theories, and experimental tests, Rev. Mod. Phys. **85**, 471 (2013).

[30] D. Gottesman, An introduction to quantum error correction and fault-tolerant quantum computation, arXiv:0904.2557 (2009).

[31] D. S. Abrams and S. Lloyd, Nonlinear Quantum Mechanics Implies Polynomial-Time Solution for *NP*-Complete and #*P* Problems, Phys. Rev. Lett. **81**, 3992 (1998).

[32] G. 't Hooft, Deterministic quantum mechanics: The mathematical equations, arXiv:2005.06374 (2020).

[33] S. L. Adler, Quantum theory as an emergent phenomenon: Foundations and phenomenology, J. Phys.: Conf. Ser. **361**, 012002 (2012).

[34] V. Vanchurin, Entropic mechanics: Towards a stochastic description of quantum mechanics, Found. Phys. **50**, 40 (2020).

[35] V. Vanchurin, The world as a neural network, Entropy **22**, 1210 (2020).

[36] M. I. Katsnelson and V. Vanchurin, Emergent quantumness in neural networks, Found. Phys. **51**, 94 (2021).

[37] E. Nelson, Review of stochastic mechanics, J. Phys.: Conf. Ser. **361**, 012011 (2012).

[38] S. Wolfram, A class of models with the potential to represent fundamental physics, Complex Systems **29**, 107 (2020).

[39] M. J. W. Hall, D.-A. Deckert, and H. M. Wiseman, Quantum Phenomena Modeled by Interactions between Many Classical Worlds, Phys. Rev. X **4**, 041013 (2014).

[40] T. N. Palmer, Discretisation of the Bloch sphere, fractal invariant sets and Bell's theorem, Proc. R. Soc. A **476**, 20190350 (2020).

[41] R. Gallego Torromé, Foundations for a theory of emergent quantum mechanics and emergent classical gravity, arXiv:1402.5070 (2014).

[42] J. Walleczek and G. Grössing, Is the world local or nonlocal? Towards an emergent quantum mechanics in the 21st century, J. Phys.: Conf. Ser. **701**, 012001 (2016).

[43] R. Orús, Tensor networks for complex quantum systems, Nat. Rev. Phys. **1**, 538 (2019).

[44] S.-H. Lin, M. P. Zaletel, and F. Pollmann, Efficient simulation of dynamics in two-dimensional quantum spin systems with isometric tensor networks, Phys. Rev. B **106**, 245102 (2022).

[45] E. M. Stoudenmire and S. R. White, Real-space parallel density matrix renormalization group, Phys. Rev. B **87**, 155137 (2013).

[46] A. Jacot, F. Gabriel, and C. Hongler, Neural tangent kernel: Convergence and generalization in neural networks, arXiv:1806.07572 (2018).

[47] I. López Gutiérrez and C. B. Mendl, Real time evolution with neural-network quantum states, Quantum **6**, 627 (2022).

[48] M. Schmitt and M. Heyl, Quantum Many-Body Dynamics in Two Dimensions with Artificial Neural Networks, Phys. Rev. Lett. **125**, 100503 (2020).

[49] S.-H. Lin and F. Pollmann, Scaling of neural-network quantum states for time evolution, Phys. Status Solidi B **259**, 2100172 (2022).

[50] B. Jónsson, B. Bauer, and G. Carleo, Neural-network states for the classical simulation of quantum computing, arXiv:1808.05232 (2018).

[51] M. Levin and X.-G. Wen, Colloquium: Photons and electrons as emergent phenomena, Rev. Mod. Phys. **77**, 871 (2005).

[52] X.-G. Wen, Classifying gauge anomalies through symmetry-protected trivial orders and classifying gravitational anomalies through topological orders, Phys. Rev. D **88**, 045013 (2013).

[53] J. S. Bell, On the einstein podolsky rosen paradox, Phys. Phys. Fiz. **1**, 195 (1964).

[54] B. S. Cirel'son, Quantum generalizations of Bell's inequality, Lett. Math. Phys. **4**, 93 (1980).

[55] S. Liberati, Tests of Lorentz invariance: A 2013 update, Class. Quantum Gravity **30**, 133001 (2013).

[56] J. Lee, L. Xiao, S. S. Schoenholz, Y. Bahri, R. Novak, J. Sohl-Dickstein, and J. Pennington, Wide neural networks of any depth evolve as linear models under gradient descent, J. Stat. Mech.: Theory Exp. (2020) 124002.

[57] S. Weinberg, Testing quantum mechanics, Ann. Phys. **194**, 336 (1989).

[58] H. Everett, "Relative state" formulation of quantum mechanics, Rev. Mod. Phys. **29**, 454 (1957).

[59] L. Hardy, Quantum ontological excess baggage, Stud. Hist. Philos. Sci. B **35**, 267 (2004).

[60] M. S. Leifer, Is the quantum state real? An extended review of ψ-ontology theorems, Quanta **3**, 67 (2014).

[61] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, Found. Phys. **24**, 379 (1994).

[62] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Exact and approximate unitary 2-designs and their application to fidelity estimation, Phys. Rev. A **80**, 012304 (2009).

[63] S. Aaronson, D. Grier, and L. Schaeffer, The classification of reversible bit operations, arXiv:1504.05155 (2015).

[64] E. L. Post, *The Two-Valued Iterative Systems of Mathematical Logic* (Cambridge University Press, Cambridge, 1941).

[65] S. Hoory and A. Brodsky, Simple permutations mix even better, arXiv:math/0411098 (2004).

[66] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, Commun. Math. Phys. **346**, 397 (2016).

[67] E. Kaplan, M. Naor, and O. Reingold, Derandomized constructions of *k*-wise (almost) independent permutations, Algorithmica **55**, 113 (2009).

[68] S. Aaronson, Is quantum mechanics an island in theoryspace? arXiv:quant-ph/0401062 (2004).

[69] L. Hardy, Quantum theory from five reasonable axioms, arXiv:quant-ph/0101012 (2001).

[70] A. Kapustin, Is there life beyond quantum mechanics? arXiv:1303.6917 (2013).

[71] G. Chiribella, G. Mauro D'Ariano, and P. Perinotti, Quantum from principles, in *Quantum Theory: Informational Foundations and Foil*, edited by G. Chiribella and R. W. Spekkens (Springer, Dordrecht, 2015), pp. 171–221, arXiv:1506.00398.

[72] G. Vidal, Class of Quantum Many-Body States That Can Be Efficiently Simulated, Phys. Rev. Lett. **101**, 110501 (2008).

[73] G. Evenbly and G. Vidal, Tensor Network Renormalization Yields the Multiscale Entanglement Renormalization Ansatz, Phys. Rev. Lett. **115**, 200401 (2015).

[74] P. Mehta and D. J. Schwab, An exact mapping between the Variational Renormalization Group and deep learning, arXiv:1410.3831 (2014).

[75] D. Jimenez Rezende and S. Mohamed, Variational inference with normalizing flows, arXiv:1505.05770 (2015).

[76] W. Grathwohl, R. T. Q. Chen, J. Bettencourt, I. Sutskever, and D. Duvenaud, FFJORD: Free-form continuous dynamics for scalable reversible generative models, arXiv:1810.01367 (2018).

[77] S. Aaronson, Quantum computing, postselection, and probabilistic polynomial-time, arXiv:quant-ph/0412187 (2004).

[78] S. Carlip, Is quantum gravity necessary? Class. Quantum Gravity **25**, 154010 (2008).

[79] V. E. Hubeny, The AdS/CFT correspondence, Class. Quantum Gravity **32**, 124010 (2015).

[80] D. Deutsch, Quantum theory of probability and decisions, Proc. R. Soc. London A **455**, 3129 (1999).

[81] D. Wallace, A formal proof of the Born rule from decision-theoretic assumptions, arXiv:0906.2718 (2009).

[82] S. M. Carroll and C. T. Sebens, Many worlds, the Born rule, and self-locating uncertainty, in *Quantum Theory: A Two-Time Success Story* (Springer-Verlag Italia, Milan, Italy, 2014), pp. 157–169.

[83] W. H. Zurek, Quantum theory of the classical: Quantum jumps, Born's rule and objective classical reality via quantum Darwinism, Philos. Trans. R. Soc. London A **376**, 20180107 (2018).

[84] L. Masanes, T. D. Galley, and M. P. Müller, The measurement postulates of quantum mechanics are operationally redundant, Nat. Commun. **10**, 1361 (2019).

[85] S. Hossenfelder, A derivation of Born's rule from symmetry, Ann. Phys. **425**, 168394 (2021).

[86] S. Saunders, J. Barrett, A. Kent, and D. Wallace, *Many Worlds? Everett, Quantum Theory, & Reality* (Oxford University Press, Oxford, 2012).

[87] A. Kent, One world versus many: The inadequacy of Everettian accounts of evolution, probability, and scientific confirmation, arXiv:0905.0624 (2009).

[88] M. McKague, M. Mosca, and N. Gisin, Simulating Quantum Systems Using Real Hilbert Spaces, Phys. Rev. Lett. **102**, 020505 (2009).

[89] A. Aleksandrova, V. Borish, and W. K. Wootters, Real-vector-space quantum theory with a universal quantum bit, Phys. Rev. A **87**, 052106 (2013).

[90] E. C. G. Stueckelberg, Quantum theory in real Hilbert-space, Helv. Phys. Acta **33**, 727 (1960).

[91] J. Myrheim, Quantum mechanics on a real Hilbert space, arXiv:quant-ph/9905037 (1999).

[92] M.-O. Renou, D. Trillo, M. Weilenmann, T. P. Le, A. Tavakoli, N. Gisin, A. Acín, and M. Navascués, Quantum theory based on real numbers can be experimentally falsified, Nature (London) **600**, 625 (2021).

[93] Z.-D. Li, Y.-L. Mao, M. Weilenmann, A. Tavakoli, H. Chen, L. Feng, S.-J. Yang, M.-O. Renou, D. Trillo, T. P. Le *et al.*, Testing Real Quantum Theory in an Optical Quantum Network, Phys. Rev. Lett. **128**, 040402 (2022).

[94] M.-C. Chen, C. Wang, F.-M. Liu, J.-W. Wang, C. Ying, Z.-X. Shang, Y. Wu, M. Gong, H. Deng, F. T. Liang *et al.*, Ruling Out Real-Valued Standard Formalism of Quantum Theory, Phys. Rev. Lett. **128**, 040403 (2022).

[95] C. M. Caves, C. A. Fuchs, and P. Rungta, Entanglement of formation of an arbitrary state of two rebits, arXiv:quant-ph/0009063 (2000).

[96] C. M. Caves, C. A. Fuchs, and R. Schack, Unknown quantum states: The quantum de Finetti representation, J. Math. Phys. **43**, 4537 (2002).

[97] W. Son, L. Amico, and V. Vedral, Topological order in 1D Cluster state protected by symmetry, Quant. Info. Proc. **11**, 1961 (2012).

[98] D. Radicevic, Spin structures and exact dualities in low dimensions, arXiv:1809.07757 (2018).

[99] B. Roy, V. Juričić, and I. F. Herbut, Emergent Lorentz symmetry near fermionic quantum critical points in two and three dimensions, J. High Energy Phys. 04 (2016) 018.

[100] G. Bednik, O. Pujolàs, and S. Sibiryakov, Emergent Lorentz invariance from strong dynamics: Holographic examples, J. High Energy Phys. 11 (2013) 064.

[101] A. Belenchia, A. Gambassi, and S. Liberati, Lorentz violation naturalness revisited, J. High Energy Phys. 06 (2016) 049.

[102] J. Collins, A. Perez, D. Sudarsky, L. Urrutia, and H. Vucetich, Lorentz Invariance and Quantum Gravity: An Additional Fine-Tuning Problem? Phys. Rev. Lett. **93**, 191301 (2004).

[103] J. Polchinski, Comment on "Small Lorentz violations in quantum gravity: Do they lead to unacceptably large effects?", Class. Quantum Gravity **29**, 088001 (2012).

[104] R. B. Laughlin, Emergent relativity, Int. J. Mod. Phys. A **18**, 831 (2003).

[105] S. Carlip, Challenges for emergent gravity, arXiv:1207.2504 (2012).

[106] M. Hermele, M. P. Fisher, and L. Balents, Pyrochlore photons: The $U(1)$ spin liquid in a $S = \frac{1}{2}$ three-dimensional frustrated magnet, Phys. Rev. B **69**, 064404 (2004).

[107] Z.-C. Gu and X.-G. Wen, A lattice bosonic model as a quantum theory of gravity, arXiv:gr-qc/0606100 (2006).

[108] A. Yu. Kitaev, Fault-tolerant quantum computation by anyons, Ann. Phys. **303**, 2 (2003).

[109] S. Bravyi, D. P. DiVincenzo, and D. Loss, Schrieffer-Wolff transformation for quantum many-body systems, Ann. Phys. **326**, 2793 (2011).