

Neural quantum states for emitter dynamics in waveguide QED

Tatiana Vovk^{1,2,*}, Anka Van de Walle^{3,4,5}, Hannes Pichler^{1,2} and Annabelle Bohrdt^{3,4}

¹*Institute for Theoretical Physics, University of Innsbruck, 6020 Innsbruck, Austria*

²*Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, 6020 Innsbruck, Austria*

³*Ludwig Maximilian University Munich, 80333 Munich, Germany*

⁴*Munich Center for Quantum Science and Technology, 80799 Munich, Germany*

⁵*Department of Physics and Astronomy, Ghent University, 9000 Ghent, Belgium*



(Received 13 August 2025; revised 11 March 2026; accepted 20 April 2026; published 6 May 2026)

Quantum emitters coupled to one-dimensional waveguides constitute a paradigmatic quantum-optical platform for exploring collective phenomena in open quantum many-body systems. For appropriately spaced emitters, they realize the Dicke model, whose characteristic permutation symmetry allows for efficient exact solutions featuring superradiance. When the emitters are arbitrarily spaced, however, this symmetry is lost and general analytical solutions are no longer available. In this work, we introduce a numerical method to study the dynamics of such systems by extending the time-dependent neural quantum state (t-NQS) framework to open quantum systems. We benchmark our approach across a range of waveguide QED settings and compare its performance with tensor-network calculations. Our results demonstrate that the t-NQS approach is competitive with other numerical methods and highlight the potential of t-NQSs for studying open quantum many-body systems out of equilibrium.

DOI: [10.1103/qccj-6vyt](https://doi.org/10.1103/qccj-6vyt)

I. INTRODUCTION

Waveguide quantum electrodynamics (WQED) platforms are a class of open quantum many-body systems, where multiple quantum emitters (such as qubits) interact through a shared photonic environment [1–5]. In these systems, emitters are coupled identically to a one-dimensional bath, giving rise to photon-mediated long-range interactions [2,6,7]. Such setups enable the exploration of rich many-body dynamics [8–12] and serve as platforms for practical applications, including dissipative state preparation [13–15], quantum memory [16–18], and quantum information processing [13,19,20]. Experimental realizations of WQED systems span a broad range of physical systems, including cold atoms near nanophotonic waveguides [21,22], superconducting qubits coupled to microwave transmission lines [23,24], and solid-state emitters such as color centers [25] and quantum dots [26] coupled to nanofibers.

Although an extensive range of WQED and quantum spin network models have been studied [2,14,27–34], much of the foundational theoretical and numerical work has been built around the paradigmatic Dicke model [35–37]. In this model, an ensemble of n excited qubits emits light collectively, which is referred to as Dicke superradiance [38]. The Dicke model

implies full permutational symmetry among emitters, which allows for a dramatic reduction of the effective Hilbert space. Specifically, an ensemble of n qubits can be described as a single collective spin- $n/2$, with a Hilbert space dimension of $(n+1)$ [39]. This symmetry facilitates analytical treatment and enables efficient numerical simulations [37,40–42].

It is of interest to understand how the collective effects including superradiance are modified once this idealized symmetry is broken, which is inevitable in realistic setups. In WQED, for instance, this occurs when the emitters are displaced from positions commensurate with the wavelength corresponding to the transition frequency. It is important to emphasize that superradiance is not eliminated by breaking the emitter position symmetry; rather, in this case the dynamics is no longer confined to the symmetric Dicke subspace and can populate additional collective states of lower symmetry. In such cases, the short-time dynamics and local observables can still be computed efficiently [39,41,43]. However, since the system's state is no longer permutationally symmetric, determining the exact transient and long-time dynamics becomes substantially challenging, as one typically needs to work in the full Hilbert space, which scales exponentially with n . This requires development of novel numerical methods capable of capturing complex open quantum many-body dynamics [41].

Among the emerging numerical methods for complex quantum many-body systems, neural quantum states (NQSs) have recently gained significant attention [44–46]. In this approach, neural networks (NNs) are employed to represent quantum many-body states, offering potentially greater expressibility than conventional tensor network (TN) techniques [47]. In the domain of open quantum systems, NQSs have been successfully applied to problems involving local noise,

*Contact author: tatiana.vovk@uibk.ac.at

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

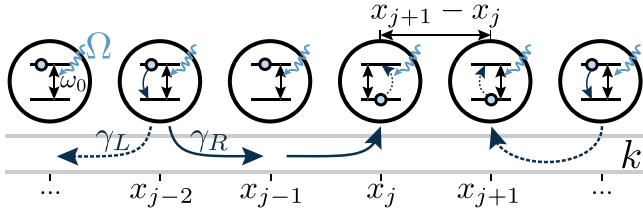


FIG. 1. Chain of spin-1/2 emitters with a resonance frequency ω_0 located at positions x_j near a one-dimensional Markovian waveguide with wavenumber k and decay rates into the left- and right-propagating photonic modes, denoted by γ_L and γ_R , respectively. The emitter chain can be additionally globally driven with a Rabi frequency Ω .

particularly for steady-state computations [48–52] and the simulation of dissipative dynamics [51–54]. In contrast, collective phenomena in WQED setups typically involve global quantum noise. Within the NQS framework, such global noise channels have so far been addressed only in the context of steady-state analysis [55]. Here, we extend the NQS framework to capture the full real-time dynamics as well.

In this work, we present a NQS-based numerical method for simulating the transient dynamics of quantum spin networks with arbitrary arrangement of qubits coupled to a Markovian one-dimensional (1D) waveguide. We employ the recently introduced time-dependent NQS (t-NQS) approach [56], extending it to the case of open-system dynamics. We demonstrate that our method captures the dynamics in the generic, symmetry-broken regime beyond the standard Dicke model: in the case of incommensurate emitter spacings, it exhibits improved convergence compared to the simulations accessible by the conventional TN method. We also demonstrate the versatility of physical regimes in WQED that can be simulated using our method. Finally, we apply the developed approach to investigate how superradiance is affected by a specific symmetry breaking that shifts the system away from the Dicke model, confirming a quadratic decrease in the peak emission intensity as the distance between emitters increases.

This paper is organized as follows. In Sec. II, we introduce the notation and give the description of the WQED setup studied in this work. In Sec. III, we analyze the performance of standard TN methods when applied to systems without permutational symmetry. Section IV introduces our t-NQS-based method for solving the dynamics of open quantum many-body systems. In Sec. V, we benchmark both approaches across several physical regimes, and then use the t-NQS method to calculate emitter dynamics in generic WQED settings.

II. PHYSICAL DESCRIPTION

In this section, we briefly describe the WQED model under investigation. Consider a system that consists of a chain of n emitters, each with a local Hilbert space dimension $d = 2$ (also referred to as qubits or spin-1/2 particles), trapped near a one-dimensional, Markovian, open-ended waveguide that acts as a shared environment, as illustrated in Fig. 1. Emitters can undergo a radiative transition from the excited state to the ground state by emitting a photon into the left- and right-propagating modes of the waveguide with rates γ_L and γ_R ,

respectively. Throughout the text, we denote the two states of the j th spin as $|0\rangle_j$ and $|1\rangle_j$, corresponding to spin-down and spin-up, respectively. We are interested in the dynamics of a spin chain that is initialized in the all-spin-up state:

$$\rho_0 = \bigotimes_{j=1}^n |1\rangle_j \langle 1|. \quad (1)$$

The spin system obeys the following Lindblad master equation (ME) [2,57,58]:

$$\partial_t \rho = -i[H + H_{\text{diss}}, \rho] + \mathcal{D}[\rho], \quad (2)$$

where $H = \Omega \sum_j \sigma_j^x$ is the system Hamiltonian corresponding to a local driving with Rabi frequency Ω , H_{diss} is the coherent term induced by the dissipation and $\mathcal{D}[\rho]$ is the dissipation superoperator [2]. We call the ME (2) balanced, when the emission rates into the left- and right-moving bath modes shown in Fig. 1 are the same, $\gamma_L = \gamma_R \equiv \gamma$. In this case, the terms H_{diss} and $\mathcal{D}[\rho]$ take the form:¹

$$H_{\text{diss}} = \gamma \sum_{j,\ell} \sin v_{j\ell} \sigma_j^+ \sigma_\ell^-, \quad (3a)$$

$$\mathcal{D}[\rho] = 2\gamma \sum_{j,\ell} \cos v_{j\ell} \left(\sigma_\ell^- \rho \sigma_j^+ - \frac{1}{2} \{ \sigma_j^+ \sigma_\ell^-, \rho \} \right),$$

$$\text{with } v_{j\ell} = k\delta x_{j\ell} \equiv k|x_j - x_\ell|, \quad (3b)$$

where k is the wave number of the guided mode, x_j are the spin positions along the waveguide, $\sigma_j^+ = |1\rangle_j \langle 0|$ and $\sigma_j^- = |0\rangle_j \langle 1|$ are the j th spin operators. We note that the decoherence rates are rescaled with the system size, $\gamma = \Gamma/n$, and we use the constant Γ as a base time scale throughout this work [59,60]. Equation (2) generally involves all-to-all two-body interactions and leads to a complicated intermediate dynamics and, in case of zero driving, $\Omega = 0$, a trivial steady state, $\rho_{\text{ss}} = \bigotimes_{j=1}^n |0\rangle_j \langle 0|$ [61].

When the distances between the emitters (or emitter spacings) are commensurate with the wavelength corresponding to the transition frequency, such that $k\delta x_{j\ell} = 2\pi m$, $\forall j, \ell$ with $m \in \mathbb{Z}^+$, the coherent part (3a) of the balanced ME vanishes and the system reduces to a permutationally symmetric Dicke model² [35,38]:

$$\partial_t \rho = -i[H, \rho] + 2\gamma \sum_{j,\ell} \left(\sigma_\ell^- \rho \sigma_j^+ - \frac{1}{2} \{ \sigma_j^+ \sigma_\ell^-, \rho \} \right). \quad (4)$$

When the system is initialized as (1) and evolves under (4), its dynamics can be reduced to that of a collective spin- $n/2$ and can be solved exactly in an efficient way [37,40–42]. In this setting, when $H = 0$, the system exhibits the phenomenon of Dicke superradiance, when the spins emit photons collectively with an intensity that is larger by a factor of n^2 than that of an

¹We refer the reader to Appendix A for the general form of the ME in the case when the waveguide is imbalanced, $\gamma_L \neq \gamma_R$.

²Throughout this work, we use the term “Dicke model” to refer to the effective spin-only formulation obtained after eliminating the photonic modes under the Born-Markov approximation; more precisely, this corresponds to a reduced open Dicke model with collective decay [2,7,38].

independent emitter [7]. This corresponds to a purely collectively dissipative regime, whereas finite driving $\Omega > 0$ brings the system to a different, driven-dissipative regime [59].

When, however, the emitter spacings are incommensurate with k , the permutational symmetry is broken and the dynamics can no longer be reduced to that of a single large spin. In this case, the many-body ME (2) has to be solved explicitly in the unconstrained Hilbert space. The exact diagonalization (ED) methods in this case are limited to small system sizes due to the exponential increase of the Hilbert space. In the following section we demonstrate an attempt to simulate the incommensurate chain dynamics with the standard tensor network (TN) technique based on the vectorization of the ME (2) [62] and show that with this method it is difficult to converge the simulations for larger system sizes.

III. CHALLENGES OF TENSOR NETWORK SOLUTIONS

In this section we present a TN-based approach to solve Eq. (2) and limitations of this type of solutions. We first provide the reader with the basic considerations concerning the implementation of the matrix product state (MPS) solution and then show how, once deviating from the standard Dicke model (4), it becomes hard for the TN approach to represent the exact state dynamics.

The standard TN technique for open quantum many-body systems involves vectorizing the density matrix, $\rho = \sum_{\alpha,\beta} \rho_{\alpha\beta} |\alpha\rangle\langle\beta| \rightarrow |\rho\rangle\rangle = \sum_{\alpha,\beta} \rho_{\alpha\beta} |\alpha\rangle \otimes |\beta\rangle$, and representing it as an MPS in the Liouville (double Hilbert) space [62,63]. In this case, the operator of the state propagation under the ME (2) can be efficiently encoded into a matrix product operator (MPO) using a finite automata approach [64,65]. We refer the reader to Appendix B for an explicit form of the propagator. Importantly, the propagator MPO has a fixed bond dimension $\chi_{\text{prop}} = 10$ for any system size.

For a given MPS, the computational cost of the TN processing $\mathcal{C}_{\text{TN}}$ is determined by its bond dimension χ as well as the system size n (see also Appendix F):

$$\mathcal{C}_{\text{TN}}(\chi, n) = n\chi^3. \quad (5)$$

An associated operator entanglement (OE) entropy acts as a proxy of this cost [66,67]:

$$E = - \sum_{i=1}^{\chi} \mu_i^2 \log_2 \mu_i^2, \text{ with } \mu_i = \frac{\tilde{\mu}_i}{\sqrt{\sum_i \tilde{\mu}_i^2}}, \quad (6)$$

where $\tilde{\mu}_i$ (μ_i) are the (normalized) singular values of the MPS's middle cut. Depending on the amount of correlations present in the state $|\rho\rangle\rangle$, its full TN representation may require different bond dimensions χ , and thus different levels of computational effort. This difference can be quantified by analyzing how the bond dimension χ and the corresponding OE (6) scale with the system size n [68]. For instance, if the OE scales linearly with n , the corresponding state exhibits a volume-law scaling of entanglement, and the exact representation requires the bond dimension to grow exponentially with n , matching the scaling of a naive full-state representation. On the contrary, when the OE saturates at a constant value (an area-law scaling), the bond dimension grows linearly, allow-

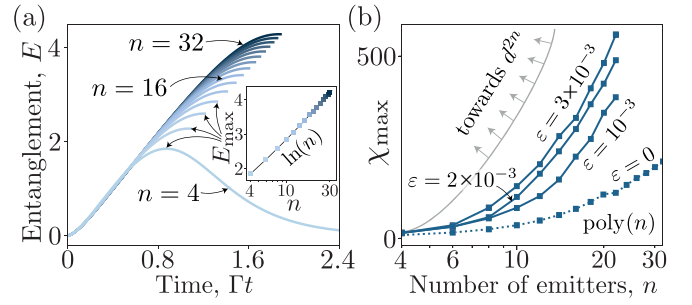


FIG. 2. Characteristics of tensor network (TN) solutions for the long-range WQED problem. (a) Operator entanglement (OE) dynamics in the Dicke regime for various system sizes ranging from $n = 4$ to $n = 32$ (indicated by different colors). Inset: scaling of the maximum OE with system size n . (b) Scaling of the maximum bond dimension χ_{max} with system size, for emitter spacings parametrized as in Eq. (7).

ing an efficient state representation. An intermediate situation implies bond dimension growth that is polynomial in n .

To assess the computational cost of solving the dynamics of ρ under the ME, we consider the scaling of the maximal OE encountered during the dynamics E_{max} with the system size n . In Fig. 2(a) we solve the Dicke ME (4) (commensurate emitter spacings, $k\delta x_{j,j+1} = 2\pi$). We start from the state (1) and show evolutions of the OE for different system sizes. One can observe a rapid entanglement buildup up to a maximal value E_{max} with a subsequent relaxation towards zero entanglement in the steady state, $E(\rho_{\text{ss}}) = 0$. This phenomenon is referred to as the entanglement barrier and is typical for open quantum systems [66,67]. In the inset of Fig. 2(a) we confirm the logarithmic scaling of E_{max} with the system size n in the Dicke regime. This corresponds to the polynomial scaling of the maximal bond dimension χ_{max} required to represent the state with the OE E_{max} , which is shown as a dashed line in Fig. 2(b).

It turns out that the entanglement dynamics significantly depends on the emitter positions x_j , even though the MPO propagator's size does not depend on this (see Appendix B). Once the emitter spacings are slightly incommensurate:

$$k\delta x_{j,j+1} = 2\pi(1 + \varepsilon), \quad (7)$$

with the spacing parameter $\varepsilon \ll 1$, the required bond dimension χ_{max} tends to grow faster with the system size, as shown in Fig. 2(b). This rapid change in computational cost indicates a significant challenge for the exact TN simulations of arbitrary spin chains with environment-induced long-range interactions. In the following section we demonstrate another way to solve this problem by using an alternative quantum many-body state representation based on NNs, i.e., so-called time-dependent neural quantum states (t-NQSSs).

IV. NEURAL QUANTUM STATES

In this section we present an alternative approach to solve the ME (2) that is based on NNs representing the system state, referred to as NQSSs [47]. We base our approach on Ref. [56], where explicitly time-dependent NQSSs (t-NQSSs) were introduced, and extend this concept to open-system dynamics.

Below we provide the reader with the basic ideas behind NQSs as well as important building blocks and computational cost estimates of the t-NQS approach used in this work.

The density matrix ρ can be represented within a NQS framework through various approaches, such as mixed-state purification [69] or vectorization of the Liouville space [55]. In this work, we choose the positive operator-valued measure (POVM) formalism to represent the density matrix [70]. In the case of informationally complete POVM (IC POVM), ρ is unambiguously defined via the probability distribution $p(\mathbf{a})$:

$$p(\mathbf{a}) = \text{tr}(\rho M_{(\mathbf{a})}), \quad (8)$$

where $\mathbf{a}$ spans over d^{2n} configurations and $\{M_{(\mathbf{a})}\}$ is a set of positive semidefinite operators called the IC POVM frame. Since the map (8) is unambiguous, there exists an inverse transformation that recovers the state, $\rho = \sum_{\mathbf{b}} p(\mathbf{b}) N^{(\mathbf{b})}$, where $\{N^{(\mathbf{b})}\}$ is the dual frame. We refer the reader to Refs. [44,54] for more information about IC POVM representations in the context of NQSs and to Appendix D for the explicit formulation of IC POVM used in this work as well as our motivation for adopting it.

NNs can be viewed as universal multivariate function approximators [71], and in our case such a function is the distribution $p(\mathbf{a})$. The state probabilities from this distribution can be approximated by the NNs in NQSs: $p(\mathbf{a}) \approx p_{\theta}(\mathbf{a})$, where θ is a vector of variational NN parameters. When the quantum state $\rho(t)$ evolves in time under the ME (2), so do the corresponding probabilities:

$$\partial_t p(\mathbf{a}, t) = \sum_{\mathbf{b}} A_{\mathbf{ab}} p(\mathbf{b}, t), \quad (9)$$

where the definition of $A_{\mathbf{ab}}$ can be found in Appendix C. There are several ways to encode the time dependence into the NQSs. Pioneering approaches include time-dependent variational Monte Carlo (t-VMC) [72,73] and projected t-VMC [74], which describe the dynamics through time-dependent NN parameters, $p(\mathbf{a}, t) \approx p_{\theta(t)}(\mathbf{a})$. On a practical level this approach implies a separate NN trained for each time step of the evolution. A more recent alternative consists in promoting t to an input parameter of the probability distribution, such that $p(\mathbf{a}, t) \approx p_{\theta}(\mathbf{a}, t)$. In this case, a single NN is trained globally to represent the quantum state at all times within a given time interval $t \in [0, T]$. This idea was independently introduced by Refs. [56,75], with the following key difference. Ref. [75] includes time dependence via a special wave function representation called Galerkin ansatz. In turn, Ref. [56] incorporates time as an additional input parameter that is being encoded in the NN.

In our work we use the latter method based on autoregressive transformer NNs. We refer the reader to Appendix D for details on autoregressive architectures. Transformer NNs are universal sequence-to-sequence function approximators initially introduced for language models [76]. Their use as NQSs spans a plethora of ground- and steady-state search methods for various physical models, including impurity models [77], dissipative models [55], and frustrated systems such as the transverse-field Ising model [78,79], J_1 - J_2 model [80], Fermi-Hubbard model [81], as well as neutral atom arrays [82,83]. Transformer's (self-)attention mechanism is particu-

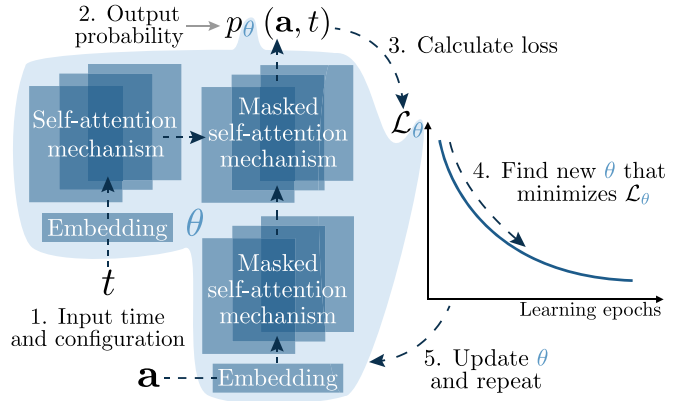


FIG. 3. Training routine of the time-dependent NQS (t-NQS) based on an autoregressive transformer NN. The NN with the parameter vector θ receives the time t and state configuration $\mathbf{a}$ as inputs and produces the corresponding probability distribution approximation $p_{\theta}(\mathbf{a}, t)$ as an output. Based on this distribution, the loss function (10) is computed, allowing us to perform a single NN optimization step (learning epoch) using standard NN backpropagation techniques. The procedure is repeated until the loss is sufficiently minimized (see Appendix E for further details).

larly useful to learn complex long-range correlation patterns in quantum many-body systems, in some cases potentially outperforming existing TN methods [45,56,79,80]. The model (2) considered in our work is thus an interesting test case for transformer NQSs, as it involves long-range dissipative interactions that rapidly generate quantum correlations in the system, rendering TN simulations hard, as we have shown in the previous section.

In Fig. 3 we show a sketch of the transformer model used to approximate the probability distribution $p(\mathbf{a}, t)$. The model has a so-called encoder-decoder structure [76], which allows us to naturally separate the configuration input $\mathbf{a}$ and time input t of the approximate function $p_{\theta}(\mathbf{a}, t)$. The encoder embeds the time input t into a higher-dimensional vector of size d_h called hidden dimension. This vector is then passed to an attention layer followed by a feed-forward NN. The attention layer dynamically assigns weights in the NQS, making the model effectively attend to correlations between different emitters. This makes capturing long-range interactions particularly efficient. A feed-forward NN enhances this representation by applying dense nonlinear transformations. Within this structure, the model learns correlations in the input and constructs a context vector, which is subsequently passed to the decoder. The decoder takes the input state configuration in the form of n tokens, $\mathbf{a} = (a_1, \dots, a_n)$ (see Appendix D), and embeds each of them into vectors of size d_h . Passing these vectors through attention layers, the decoder integrates them with the time context vector coming from the encoder. This information is then passed through another feed-forward NN with a softmax activation function and then is finally decoded into an approximation of the probability distribution, $p_{\theta}(\mathbf{a}, t)$.

The goal of the NN training is to minimize the loss function, which measures the deviation of the NQS output $p_{\theta}(\mathbf{a}, t)$ from the actual distribution $p(\mathbf{a}, t)$. In the case of t-NQSs, the loss function is simply the error of Eq. (9). Such a direct loss estimation is possible due to the fact that t is the input of the

NN, allowing for an exact and efficient calculation of time derivatives of $p_\theta(\mathbf{a}, t)$ using automated differentiation [84]. Once such a loss function is sufficiently globally minimized during the NN training, the resulting t-NQS represents the correct solution of Eq. (9). In the variational Monte Carlo (VMC) framework, this loss function can be written as

$$\mathcal{L}_\theta = \frac{dt}{|\tau|} \sum_{t \in \tau} \sum_{\mathbf{a} \sim p_\theta(\mathbf{a}, t)} \left| \frac{\partial_t p_\theta(\mathbf{a}, t) - \sum_{\mathbf{b}} A_{\mathbf{ab}} p_\theta(\mathbf{b}, t)}{p_\theta(\mathbf{a}, t)} \right|, \quad (10)$$

where the outer sum runs over time steps t from a discrete subset of the whole time interval, $\tau \subset [0, T]$ with dt being the discretization step, and the inner sum involves Born-rule sampling of state configurations. In order to take into account the initial state of the system as well as to simulate the dynamics for several time intervals of duration T , we utilize the concatenation technique introduced in Ref. [56]. We refer the reader to Appendix E for the details on the time interval concatenation, the choice of the interval subset τ as well as other technical aspects of the t-NQS model.

For the t-NQS model described above, the computational cost is determined by its hidden dimension d_h and the system size n [76]:

$$C_{\text{NQS}}(d_h, n) = n^4 d_h + n^3 d_h^2. \quad (11)$$

The high-order polynomial scaling with n is related to the quadratic scaling of transformer's self-attention and the separate application of two-body terms from the ME (2) during the estimation of the loss (10). We refer the reader to Appendix F for the detailed definitions of the computational cost. In practice, such a scaling of the cost allows us to perform computations with system sizes up to $n = 20$, as done in Sec. V.

V. RESULTS

In this section we compare the performances of the TN method presented in Sec. III and the t-NQS method presented in Sec. IV. Specifically, we apply both methods to solve the ME (2) that generates the dynamics of a spin-1/2 chain near a 1D waveguide. In Figs. 4 and 5 we compare the convergence properties of both methods for two instances of the balanced ME: the commensurate (Dicke) regime with $k\delta x_{j,j+1} = 2\pi$ and the incommensurate regime with $k\delta x_{j,j+1} = \pi/\sqrt{3}$. In addition, in Fig. 6 we demonstrate the capability of the t-NQS method to solve various instances of the general ME (as defined in Appendix A). Finally, using the t-NQS simulations, in Fig. 7 we consider how the WQED physics changes once moving away from the permutationally symmetric Dicke regime.

We consider how several system observables evolve under the ME (2). Specifically, we consider expectation values of the magnetization and fluctuations (cumulative two-body correlations) along the z direction of the spin chain:

$$\langle Z \rangle \equiv \frac{1}{n} \sum_{j=1}^n \langle \sigma_j^z \rangle, \quad (12)$$

$$\langle F_z^2 \rangle \equiv \frac{1}{n} \sum_{j,\ell=1}^n (\langle \sigma_j^z \sigma_\ell^z \rangle - \langle \sigma_j^z \rangle \langle \sigma_\ell^z \rangle) \quad (13)$$

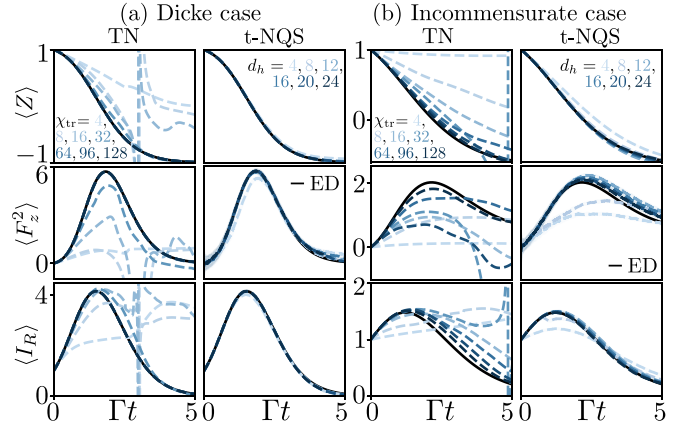


FIG. 4. Dynamics of physical observables of a spin chain near a waveguide for (a) commensurate and (b) incommensurate emitter spacings. The first and second rows feature the evolution of magnetization and fluctuations along the z axis, respectively. The third row shows the output intensity in the right-propagating waveguide mode (see main text). Left and right subcolumns in (a) and (b) correspond to TN and t-NQS simulations. Different line colors correspond to different values of truncation bond dimension χ_{tr} or hidden dimension d_h for TN and t-NQS, respectively. The system size is $n = 20$ in (a) and $n = 12$ (maximum accessible by the ED in the full Hilbert space) in (b). The shaded regions around t-NQS curves indicate the sampling error.

as well as the intensity $\langle I_R \rangle \equiv \langle c_R^\dagger c_R \rangle / n$ of the output emission in the right direction of the waveguide governed by the jump operator $c_R = \sum_j e^{ikx_j} \sigma_j^-$ (see Appendix A and Ref. [2] for details). This choice allows us to probe the performances of our numerical methods for both linear (magnetization) and quadratic (fluctuations, output intensity) observables.

In Fig. 4 we consider the dynamics of the mentioned observables in a system with a fixed size n , with different colors indicating varying numerical precision of the respective methods. In the Dicke case shown in Fig. 4(a) we observe a stable convergence by both the TN and t-NQS methods for all of the chosen observables. Specifically, the TN method coincides with the ED at all times already for the truncation³ bond dimension $\chi_{\text{tr}} = 64$, which confirms the polynomial scaling from Fig. 2(b). Incidentally, the t-NQS method gives correct predictions at a rather low hidden dimension $d_h = 8$. Moreover, we observe a good approximation with even lower values of χ_{tr} or d_h for the magnetization $\langle Z \rangle$.

The incommensurate case is different from the Dicke regime, as can be seen from comparing Fig. 4(b) with Fig. 4(a). Specifically, for both methods the convergence to the ED is slower for all considered observables. In the context of TNs, this is consistent with the finding shown in Fig. 2(b) of Sec. III. In this case, one can also observe a slightly better convergence for the linear observable in comparison to the quadratic observables for both methods. On the other hand, the t-NQS lines approach the ED lines closer, even if not converging as good as in the Dicke case. It is important to note

³Truncation bond dimension χ_{tr} is the maximal bond dimension allowed during simulations.

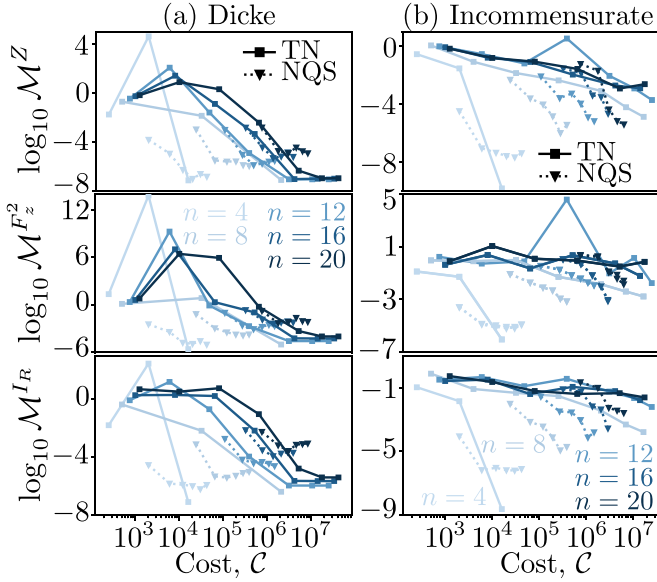


FIG. 5. Dependence of the average square deviation on the computational cost values for Dicke (a) and incommensurate (b) scenarios. As in Fig. 4, the rows display the magnetization and fluctuations along the z axis (first and second rows), and the right-propagating waveguide mode output (third row). Solid lines with squares and dotted lines with triangles represent average square errors for TN and NQS methods, respectively. The lines' colors indicate different system sizes.

that increasing the truncation bond dimension χ_{tr} of TNs and the hidden dimension d_h of NQSs leads to different scaling of the associated computational costs, as indicated in Eqs. (5) and (11). For instance, increasing χ_{tr} from 4 to 8 results in an eightfold increase in computational cost $\mathcal{C}_{TN}$, whereas increasing d_h from 4 to 8 leads to only a fourfold increase of $\mathcal{C}_{NQS}$. In light of this, together with the results of Fig. 4, we find that for a comparable increase in computational cost, the NQSs exhibit faster convergence than the TNs in the specific scenarios considered.

The results of Fig. 4 motivate comparing the convergence properties of the studied numerical methods. As a measure of how well the simulation converges for a given observable (O), we take the time average of the squared deviation of the numerical data from the ED solution for a fixed evolution time T :

$$\mathcal{M}_{\text{num}}^O(\mathcal{C}) = \frac{1}{T} \int_0^T dt (\langle O \rangle_{\text{num}}(t) - \langle O \rangle_{\text{ED}}(t))^2, \quad (14)$$

where the subscript $\text{num} \in \{\text{TN}, \text{NQS}\}$ indicates the numerical method, O is the operator corresponding to a physical observable and $\mathcal{C}$ is the corresponding computational cost as defined in Eqs. (5) and (11). We note that the error (14) stems from two different sources. The first is the error of the state approximation that is defined in TNs by the truncation bond dimension χ_{tr} and in NQSs by the hidden dimension d_h . The second source is related to all residual errors, e.g., an error from a finite time step in the first-order time propagation in TNs or finite sampling size, limited optimization capabilities and time interval discretization in t-NQSs. In the following analysis we focus on the former error of the state

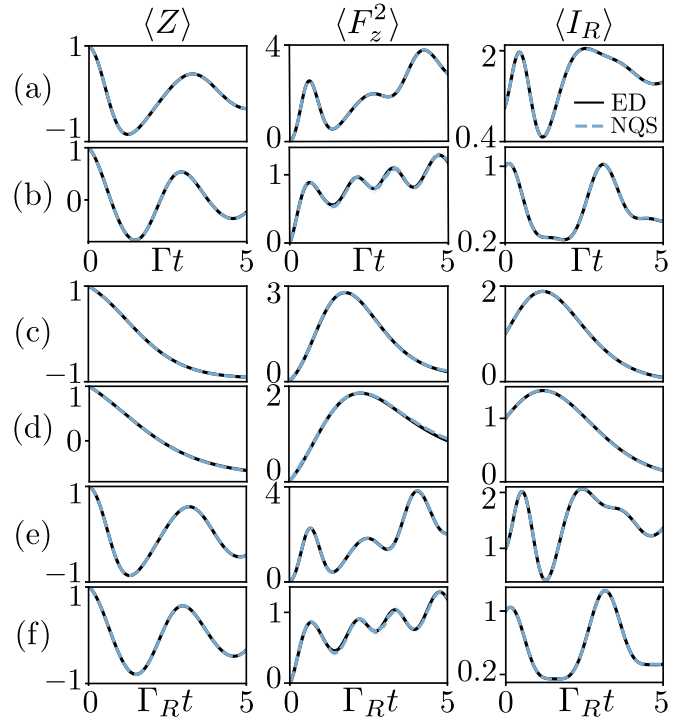


FIG. 6. Dynamics of physical observables of the spin chain near a waveguide for several cases of the generalized master equation (ME) as defined in Appendix A: (a) driven Dicke model ($\Gamma_L = \Gamma_R \equiv \Gamma = \Omega$, $k\delta x_{j,j+1} = 2\pi$), (b) incommensurate case with driving ($\Gamma_L = \Gamma_R \equiv \Gamma = \Omega$, $k\delta x_{j,j+1} = \pi/\sqrt{3}$), (c) commensurate chiral case ($\Gamma_L = 0.5\Gamma_R$, $\Omega = 0$, $k\delta x_{j,j+1} = 2\pi$), (d) incommensurate chiral case ($\Gamma_L = 0.5\Gamma_R$, $\Omega = 0$, $k\delta x_{j,j+1} = \pi/\sqrt{3}$), (e) commensurate driven chiral case ($\Gamma_L = 0.5\Gamma_R$, $\Omega = \Gamma_R$, $k\delta x_{j,j+1} = 2\pi$), (f) incommensurate driven chiral case ($\Gamma_L = 0.5\Gamma_R$, $\Omega = \Gamma_R$, $k\delta x_{j,j+1} = \pi/\sqrt{3}$). The columns correspond to the same observables as in Figs. 4 and 5. The system size is $n = 8$, the hidden dimension of the t-NQS is $d_h = 16$ for (a), (c), (e) and $d_h = 20$ for (b), (d), and (f). The semitransparent shadings around the t-NQS plot lines indicate the sampling error.

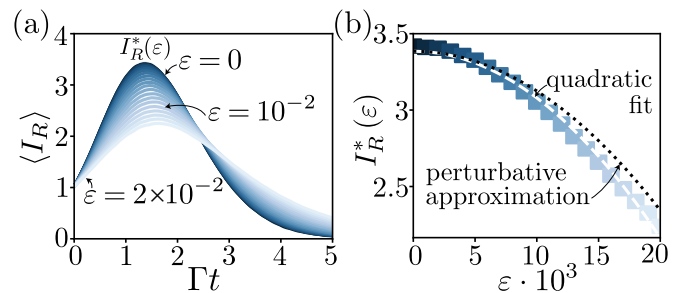


FIG. 7. Dependence of the emission intensity dynamics on the emitter spacing. (a) Time evolution of the emission intensity into the right-propagating waveguide mode, $\langle I_R \rangle$, for various emitter spacings (see text). The emission intensity grows during the transient dynamics, reaching the peak value $I_R^*(\epsilon)$, and then goes down to zero as the system approaches the steady state. The semitransparent shadings around the t-NQS plot lines indicate the sampling error. (b) The peak intensity I_R^* as a function of the spacing parameter ϵ . The perturbative approximation and quadratic fit are shown as dotted black and dashed white lines, respectively. The system size is $n = 16$, the hidden dimension of the t-NQS is $d_h = 20$.

approximation, while keeping the residual errors small. Residual errors thus set the lower limit for the value of the average square deviation (14). This means that, upon reducing the state approximation error by increasing the associated computational cost $\mathcal{C}$, the value of $\mathcal{M}_{\text{num}}^O$ reduces and eventually saturates at the value of the residual error.

We therefore aim to find out how much computational cost does a given state ansatz require to reach the threshold set by the residual errors. We thus state that the simulation is converged in terms of the state approximation for a given observable O if the average square error $\mathcal{M}_{\text{num}}^O$ reaches a plateau at some sufficient cost $\mathcal{C}^*$,

$$\mathcal{M}_{\text{num}}^O(\mathcal{C} \geq \mathcal{C}^*) = \tilde{\mathcal{M}}_{\text{num}}^O, \quad (15)$$

where the value of the plateau $\tilde{\mathcal{M}}_{\text{num}}^O$ is set by the residual errors.

In Fig. 5 we analyze the reduction the overall average square deviations of both methods, $\mathcal{M}_{\text{TN}}^O$ and $\mathcal{M}_{\text{NQS}}^O$, in terms of increasing computational cost.⁴ The TN and NQS average square deviations are indicated by solid and dashed lines, respectively. We consider the Dicke and incommensurate regimes in Fig. 5(a) and 5(b) and indicate various system sizes n by different colors. We note that in the Dicke case, the average square error is evaluated against the ED for all system sizes. In the incommensurate case, the ED is used for the system sizes up to $n = 12$, while for larger n , $\mathcal{M}_{\text{TN}}^O$ and $\mathcal{M}_{\text{NQS}}^O$ are calculated relative to the solutions with the largest bond dimension χ_{tr} or hidden dimension d_h , respectively. From the Dicke case shown in Fig. 5(a), one can conclude that both methods satisfy the convergence conditions. One can observe that the TN average square error is smaller compared to the NQS case, $\tilde{\mathcal{M}}_{\text{TN}}^O < \tilde{\mathcal{M}}_{\text{NQS}}^O$. This is expected because in this case the TNs represent the state exactly up to the second order in the time step [85], while the residual errors in NQSs are larger and depend on the system size. This dependence can be explained by the quadratically growing number of Lindbladian terms in the loss (10) and henceforth increased learning effort with the system size, and also by the discretization of time intervals in t-NQSs. A completely different scenario can be seen for the incommensurate regime shown in Fig. 5(b). Here, only for the smallest system size, $n = 4$, the TNs provide a solution that is close to the ED,⁵ while for larger system sizes the TN solutions fail to satisfy the convergence condition (15) for all of the observables. In turn, the t-NQS solutions show a clearer trend toward satisfying condition (15) for the cases of the magnetization $\langle Z \rangle$ and output intensity $\langle I_R \rangle$, and in the case of fluctuations $\langle F_z^2 \rangle$ t-NQSs confidently exhibit lower deviation values with increasing cost. Even though the deviation here is larger compared to the one

in the Dicke case, it can be confirmed that for all observables, the deviations accessible by NQSs are lower than the deviations accessible by TNs, $\mathcal{M}_{\text{NQS}}^O(\mathcal{C}_{\text{NQS}}^*) < \mathcal{M}_{\text{TN}}^O(\mathcal{C}_{\text{TN}}^*)$, with $\mathcal{C}_{\text{num}}^*$ being the maximal cost accessible by the respective numerical method. This confirms that, compared to the simulations accessible by the TN method, the t-NQS predictions are more accurate for incommensurate spin dynamics, while demonstrating no substantial gain in the case of permutationally symmetric Dicke model.

With the t-NQS method presented here, one is able to study the dynamics of various WQED settings beyond the examples considered so far. In Fig. 6 we take several instances of the general ME (presented in Appendix A) and plot the dynamics of different physical observables generated by the t-NQS method, comparing them to the ED results. One can see that the t-NQS is capable of capturing complex dynamics of driven systems with collective decay as well as emitters coupled to a chiral waveguide. We note that the incommensurate cases shown in Fig. 6(b), 6(d), and 6(f), while requiring a higher hidden dimension ($d_h = 20$) compared to the commensurate case ($d_h = 16$), converge smoothly to the ED curves, demonstrating the ability of the t-NQS to accurately represent the dynamics even in more complex settings.

Using the t-NQS method, one can explore the physics of the WQED beyond the standard Dicke regime [43]. Specifically, in Fig. 7 we consider the dynamics of the emission intensity $\langle I_R \rangle$ and how the peak intensity value I_R^* scales with the emitter spacing once deviating from the commensurate regime. This deviation is parametrized by ε in Eq. (7) and is also studied with TNs in Fig. 2(b). Figure 7(a) shows that the parameter ε has a pronounced impact on the dynamics of $\langle I_R \rangle$, with growing ε leading to a gradual suppression of the intensity peak I_R^* . Within the perturbation-theoretic approximation, one can expect I_R^* to decrease quadratically with sufficiently small ε , which is indicated by a black dotted line in Fig. 7(b) (we refer the reader to Appendix G for the derivation of the perturbative approximation of the quadratic scaling). We confirm this behavior with our t-NQS simulations, the results of which appear as square data points in Fig. 7(b). We additionally provide a quadratic fit of the data, shown as a white dashed line in Fig. 7(b) (we refer the reader to Appendix G for the parameters of the numerical quadratic fit). The results presented in Fig. 7 indicate that a slight deviation from commensurate emitter spacings, a common occurrence in experimental settings, leads to a quadratic suppression of the superradiant emission rate.

VI. CONCLUSIONS AND OUTLOOK

In this work, we introduce a neural quantum state (NQS) approach for simulating the dynamics of spin chains with dissipative all-to-all interactions mediated by a one-dimensional Markovian waveguide. Our method builds on the recently proposed time-dependent NQS (t-NQS) framework, which incorporates time as an explicit input to the neural network [56], extending the approach to the density matrix representation using informationally complete positive operator-valued measures (IC POVM) [44]. We benchmark the performance of t-NQS against a conventional tensor network (TN) method based on the matrix product state (MPS) representation of the

⁴We note that the costs for TNs and NQSs from Eqs. (5) and (11) are defined as scalings in different relative units and should not be compared quantitatively. Moreover, the considered ranges of computational costs differ for TNs and NQSs. This stems from our choice of the bond dimensions χ_{tr} and hidden dimensions d_h .

⁵For TNs with low bond dimensions one can observe a larger deviation for the Dicke model compared to the incommensurate case. This can potentially be related to the nonpositivity of the TN solution upon truncation of the bond dimension [62,86].

vectorized density matrix [63,66,67]. Our results demonstrate that the t-NQS method is able to capture spin dynamics in generic, symmetry-breaking configurations, albeit currently limited to system sizes of up to $n = 20$. In particular, our method predicts the dynamics of the incommensurate emitter spacings more accurately than the TN approach across accessible bond dimensions, while performing on par with TNs in the permutationally symmetric Dicke regime. Additionally, we demonstrate that the developed numerical method is able to simulate various physical regimes of the WQED, including emitter systems with driving, incommensurate spacings and chiral dissipative couplings. Finally, with the developed tool we study how the phenomenon of superradiance changes once moving away from the permutationally symmetric regime, and observe that the peak emission intensity is quadratically suppressed with increasing perturbation in the emitter spacings, in agreement with perturbative predictions.

Several directions remain open for future research. From the physics perspective, an important next step is to further investigate the superradiance in symmetry-broken regimes. In particular, the perturbative deviation from the Dicke model presented in Fig. 7 could be extended using finite-size scaling argument, to characterize the fate of superradiance in the thermodynamic limit. On the methodological side, the t-NQS framework developed in this work can naturally accommodate a broader class of physical models. For instance, our implementation readily extends to systems with time-dependent parameters, enabling the study of complex processes such as optimal control and dissipative state preparation [87–89]. Further improvements may come from exploring alternative NQS representations. Specifically, employing neural network ansätze operating on the vectorized density matrix [55] could offer advantages over the POVM-based sampling used in this work. Moreover, the vectorized form of the NQS could open doors for developing mappings between the TN and NN representations of quantum states [90], facilitating switching between different methods in order to overcome entanglement barriers such as ones shown in Fig. 2(a). A key technical challenge in this direction remains the efficient treatment of all-to-all interactions within the NQS framework. Addressing this will be essential for scaling to larger system sizes and more complex geometries. In addition, while we compared our approach with the standard TN representation of the vectorized density matrix [63,66,67], we did not consider quantum trajectory based TN approaches [41,60,91–94]. A systematic comparison with such methods remains an important future research direction.

We also note that extending the present t-NQS framework to WQED systems with explicit consideration of mechanical degrees of freedom [95] is conceptually straightforward within the POVM-based formulation, although computationally more demanding, and constitutes a promising topic of future research. Another interesting generalization of the considered case is independently distributed position deviations $\{\varepsilon_i\}$ drawn from, e.g., a Gaussian distribution $\mathcal{N}(\bar{\varepsilon}, \sigma^2)$ with mean $\bar{\varepsilon}$ and variance σ . This should introduce phase disorder to the collective jump operators and is expected to suppress the peak emission rate $\langle I_R^* \rangle$ in a manner that depends on both $\bar{\varepsilon}$ and σ . A systematic analysis of this scenario is left for future work. Finally, while driven WQED models, considered in

parts of this work, may, in addition, exhibit chaotic dynamics in certain parameter regimes [36,96], in this work we focus on convergence properties arising from symmetry breaking and the resulting increase in state complexity. A systematic investigation of how chaos may affect numerical simulability remains an interesting direction for future research. Taken together, these directions suggest that NQS-based methods may offer a promising route for simulating open quantum many-body dynamics, especially as their flexibility and scalability are further explored.

ACKNOWLEDGMENTS

We thank Ana Maria Rey, Darrick Chang, Ana Asenjo-Garcia, Alexander Poddubny, Ephraim Shahmoon, Markus Schmitt, Dominik Wild, Luca Arceci, Fan Yang, Giuliano Giudici, Katherine Van Kirk, Bennet Windt, and Nikita Leppen for fruitful discussions. The computational results presented here have been achieved using the Linux Cluster of the Institute for Theoretical Physics of the University of Innsbruck, LEO HPC infrastructure of the University of Innsbruck and the HPC infrastructure of the Arnold Sommerfeld Center for Theoretical Physics at the Ludwig Maximilian University in Munich. This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC-2111–390814868), the European Union's Horizon Europe research and innovation program under Grant Agreement No. 101113690 (PASQuanS2.1), the ERC Starting grant QARA (Grant No. 101041435), the EU-QUANTERA project TNiSQ (N-6001), the Austrian Science Fund (FWF) (Grant No. DOI 10.55776/COE1). For open access purposes, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

DATA AVAILABILITY

The data that support the findings of this article are openly available [97].

APPENDIX A: GENERAL MASTER EQUATION FOR THE IMBALANCED MARKOVIAN WAVEGUIDE

In the case when the waveguide has different dissipation rates for the left- and right-propagating modes, $\Gamma_L \neq \Gamma_R$, the dissipation-induced terms of the ME (2) can be written as sums over the left and right contributions:

$$H_{\text{diss}} = H_L + H_R, \quad (\text{A1})$$

$$\mathcal{D}[\rho] = D_L[\rho] + D_R[\rho], \quad (\text{A2})$$

where the coherent contributions (A1) have the following form:

$$H_L = -\frac{i\gamma_L}{2} \sum_{j<\ell} (e^{ik|x_j-x_\ell|} \sigma_j^+ \sigma_\ell^- - \text{H.c.}), \quad (\text{A3a})$$

$$H_R = -\frac{i\gamma_R}{2} \sum_{j>\ell} (e^{ik|x_j-x_\ell|} \sigma_j^+ \sigma_\ell^- - \text{H.c.}), \quad (\text{A3b})$$

with $\gamma_{L,R} = \Gamma_{L,R}/n$ are the decoherence rates rescaled with n [59,60]. The dissipators (A2) read:

$$D_\alpha[\rho] = \gamma_\alpha (c_\alpha \rho c_\alpha^\dagger - \frac{1}{2} \{c_\alpha^\dagger c_\alpha, \rho\}), \quad (\text{A4})$$

where $\alpha \in \{L, R\}$ and $c_{L,R} = \sum_j e^{\mp i k x_j} \sigma_j^\pm$. We also refer the reader to Refs. [2,98] for further information on chiral spin networks.

APPENDIX B: TENSOR-NETWORK MASTER EQUATION PROPAGATOR

A site- j tensor from an MPO that encodes the propagation under the balanced ME (2)–(3) can be written as [64,65]

$$A_j = \begin{pmatrix} \mathbb{I} & c_-^\rightarrow & c_+^\rightarrow & (c_-^\rightarrow)^\dagger & (c_+^\rightarrow)^\dagger & c_-^\leftarrow & c_+^\leftarrow & (c_-^\leftarrow)^\dagger & (c_+^\leftarrow)^\dagger & \mathcal{L}_{\text{loc}} \\ 0 & \mathbb{I} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -(c_-^\rightarrow)^\dagger + (c_-^\leftarrow)^\dagger \\ 0 & 0 & \mathbb{I} & 0 & 0 & 0 & 0 & 0 & 0 & (c_+^\rightarrow)^\dagger \\ 0 & 0 & 0 & \mathbb{I} & 0 & 0 & 0 & 0 & 0 & c_-^\rightarrow/2 - c_-^\leftarrow/2 \\ 0 & 0 & 0 & 0 & \mathbb{I} & 0 & 0 & 0 & 0 & -c_+^\rightarrow \\ 0 & 0 & 0 & 0 & 0 & \mathbb{I} & 0 & 0 & 0 & (c_-^\leftarrow)^\dagger/2 - (c_-^\rightarrow)^\dagger/2 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I} & 0 & 0 & -(c_+^\leftarrow)^\dagger \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I} & 0 & (-c_-^\leftarrow + c_-^\rightarrow) \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I} & c_+^\rightarrow \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I} \end{pmatrix}_j, \quad (\text{B1})$$

where $\mathbb{I} = \mathbb{1} \otimes \mathbb{1}$ is the identity operator in the Liouville space, $c_\pm^\rightarrow = \sqrt{\gamma dt}(\sigma^- \otimes \mathbb{1})e^{\pm i k x}$ and $c_\pm^\leftarrow = \sqrt{\gamma dt}(\mathbb{1} \otimes \sigma^-)e^{\pm i k x}$ are the short notations for jump operators in the Liouville space, and the local Lindbladian term being

$$\begin{aligned} \mathcal{L}_{\text{loc}} = & -i\Omega dt(\sigma^x \otimes \mathbb{1} - \mathbb{1} \otimes \sigma^x) \\ & + 2\gamma dt(\sigma^- \otimes \sigma^+ - \sigma^+ \sigma^- \otimes \mathbb{1} - \mathbb{1} \otimes \sigma^+ \sigma^-). \end{aligned}$$

Note that the site index j is omitted but implied for every operator appearing in (B1). The resulting MPO has a fixed bond dimension $\chi_{\text{prop}} = 10$ for any system size and is able to encode both permutationally symmetric (Dicke) and incommensurate propagators. In the former case, the exponents in jump operators $c_\pm^\rightarrow$ and $c_\pm^\leftarrow$ correspond to a trivial gauge and can be safely dropped, thus leading to identical MPO tensors that do not depend on the position j . The generalization to the imbalanced ME presented in Appendix A is straightforward.

APPENDIX C: MASTER EQUATION AS AN EQUATION OF MOTION FOR PROBABILITY DISTRIBUTIONS

The balanced ME (2)–(3) for the density matrix $\rho(t)$ can be rewritten in IC POVM representation as Eq. (9), which is an equation of motion for the probability distribution $p(\mathbf{a}, t)$ defined in Eq. (8). In this ME, the generator of time evolution consists of coherent and incoherent contributions, $A_{\text{ab}} = C_{\text{ab}} + D_{\text{ab}}$, which can be written as follows:

$$\begin{aligned} C_{\text{ab}} &= -i\text{tr}((H + H_{\text{diss}})[N^{(\mathbf{b})}, M_{(\mathbf{a})}]), \\ D_{\text{ab}} &= \gamma \sum_{j,\ell} \cos v_{j\ell} \mathbb{D}_{j\ell}^{\text{ab}}, \\ \mathbb{D}_{j\ell}^{\text{ab}} &= \text{tr}(2\sigma_\ell^- N^{(\mathbf{b})} \sigma_j^+ M_{(\mathbf{a})} - \sigma_j^+ \sigma_\ell^- \{N^{(\mathbf{b})}, M_{(\mathbf{a})}\}), \end{aligned} \quad (\text{C1})$$

where $\mathbf{a}$ and $\mathbf{b}$ are the state configuration vectors, $M_{(\mathbf{a})}$ and $N^{(\mathbf{b})}$ are the IC POVM frame and dual frame, respectively, and definitions of the remaining terms can be found in the

discussion to Eqs. (2) and (3) of the main text. The generalization to the imbalanced ME presented in Appendix A is straightforward.

APPENDIX D: LOCAL PROBABILITY REPRESENTATIONS

As mentioned in the main text, the t-NQSs used in this work are based on autoregressive NN sampling. This approach consists in representing the NN output, i.e., the probability distribution $p(\mathbf{a})$, in a local basis:

$$p(\mathbf{a}) = p(a_1, \dots, a_n), \quad (\text{D1})$$

where each a_i spans $d^2 = 4$. Such a representation is generated with local IC POVM frames, $\{M_{(\mathbf{a})}\} = \{M_{(a_1)} \otimes \dots \otimes M_{(a_n)}\}$. Each of the local frames $M_{(a_j)}$ can be chosen in multiple ways as long as it corresponds to an IC POVM representation. In this work we choose each of the local frames to form a so-called Pauli-4 basis [54], which can be written as follows:

$$\begin{aligned} M_{(0)_j} &= \frac{1}{3}|0\rangle_j\langle 0|, \\ M_{(1)_j} &= \frac{1}{3}|+\rangle_j\langle +|, \\ M_{(2)_j} &= \frac{1}{3}|r\rangle_j\langle r|, \\ M_{(3)_j} &= \mathbb{1}_j - M_{(0)_j} - M_{(1)_j} - M_{(2)_j}. \end{aligned} \quad (\text{D2})$$

Here $|0\rangle_j$, $|+\rangle_j$ and $|r\rangle_j$ are the eigenvectors with eigenvalue +1 of spin operators σ_j^x , σ_j^y , and σ_j^z , respectively (j th spin Pauli operators in the x , y , and z direction).

This allows us to subsequently sample local configurations from conditional probabilities, $p(\mathbf{a}) = \prod_{j=1}^n p(a_j|a_1, \dots, a_{j-1})$, and avoid global sampling with computationally costly Markov chain Monte Carlo techniques [48,54]. Such an economic sampling is possible in our case, since the probability distributions remain normalized under

the trace-preserving ME (2). The NN architectures suitable for autoregressive sampling include recurrent NNs (RNNs) [99–101], masked convolutional neural networks (CNNs) [44,90,102], and transformers [54,78,79,82,103], the latter being used in the present work.

APPENDIX E: TECHNICAL DETAILS ON THE NEURAL NETWORK MODEL

The t-NQS described in Sec. IV is trained to represent the mixed state of the system at all times t in a given time interval $t \in [0, \mathcal{T}]$. Since the intermediate dynamics depends on the initial state ρ_0 , it is crucial to set up the model at the correct state at time $t = 0$. We follow Ref. [56] by incorporating an initialization (concatenation) layer at the end of the NN architecture, which blends fixed initial IC POVM probabilities $p_0(\mathbf{a}) = \text{tr}(\rho_0 M_{(\mathbf{a})})$ with the output values of the trained NN with parameters θ_1 . This layer forces the total t-NQS output to have the form of a simple convex combination in time domain:

$$\text{concat}(p_0, p_{\theta_1}, t) = f(t)p_0(\mathbf{a}) + (1 - f(t))p_{\theta_1}(\mathbf{a}, t) \quad (\text{E1})$$

with $t \in [0, \mathcal{T}]$,

where $f(t)$ is an arbitrary differentiable function with boundary conditions $f(0) = 1$ and $f(\mathcal{T}) = 0$, and $p_{\theta_1}(\mathbf{a}, t)$ corresponds to the output of the NN parametrized by θ_1 . For all the results presented in this work, we choose the simplest concatenation function, $f(t) = 1 - t/\mathcal{T}$, as we found it to facilitate the NN training.

We note that this approach can be extended beyond a single time interval, e.g., it can be used to concatenate two subsequent intervals. In this case, the concatenation layer blends the previous NN output, $p_{\theta_1}(\mathbf{a}, \mathcal{T})$, taken at the last time step of its interval, and the next network, $p_{\theta_2}(\mathbf{a}, t)$, at the subsequent interval $t \in [\mathcal{T}, 2\mathcal{T}]$:

$$\text{concat}(p_{\theta_1}, p_{\theta_2}, t) = f(t)p_{\theta_1}(\mathbf{a}, \mathcal{T}) + (1 - f(t))p_{\theta_2}(\mathbf{a}, t) \quad (\text{E2})$$

with $t \in [\mathcal{T}, 2\mathcal{T}]$.

The first NN is treated as an initial state and its parameters θ_1 are fixed, i.e., there is no training occurring for this NN, and the second NN with parameters θ_2 is being trained for the corresponding time interval. In this way one can facilitate the training for a larger time range, $[0, m\mathcal{T}]$ with $m \in \mathbb{Z}^+$, by dividing the range into m smaller intervals of duration $\mathcal{T}$, training a separate NN for each of them and concatenating them one after another in a sequence.

The t-NQS output at a given interval is used to estimate the loss function $\mathcal{L}_\theta$ provided in Eq. (10) of the main text. The training of the corresponding transformer NN consists in minimization of $\mathcal{L}_\theta$ with respect to the NN parameters θ . This minimization is performed using the gradient-based backpropagation in the PYTORCH framework [84]. We find that, for the purpose of optimization, it is sufficient to include only the gradient of the first term in the loss function (10) that corresponds to the time derivative, while neglecting the gradient contribution from the second, Lindbladian term. The reason for this is evident from Fig. 8, where we plot cumulative absolute values of gradients summed across all the NN parameters θ of the corresponding loss terms. These values are simple indicators of the importance of the correspond-

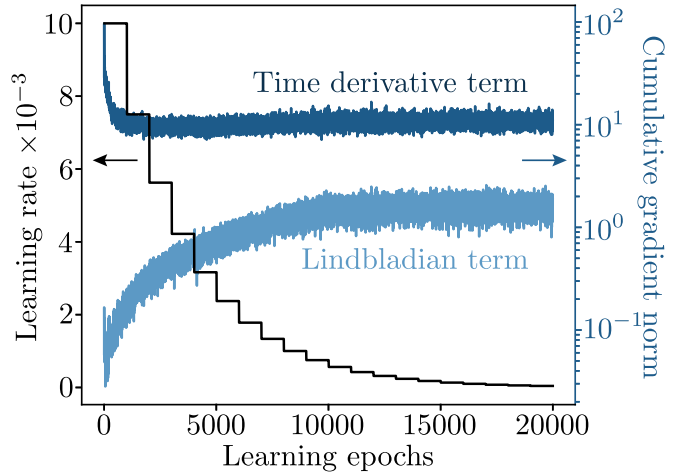


FIG. 8. A typical example of the gradient absolute value accumulated for all the NN parameters θ for the time derivative (dark blue solid line) and Lindbladian (light blue solid line) contributions to the loss (10) during the training of a single t-NQS. The black solid line indicates the learning rate changes during the training. Here the physical parameters are: the system size is $n = 4$, the distance between the emitters is $k|x_j - x_{j+1}| = \pi/\sqrt{3}$, the driving is $\Omega = 0.2\Gamma_R$, and decoherence rates are $\Gamma_L = 0.5\Gamma_R$. The NN model has hidden dimension $d_h = 20$, with the rest of the model parameters being the same as in the text.

ing gradient contributions for the overall learning process. Additionally we plot the learning rate schedule (see further technical details in the next paragraph). We observe a clear dominance of the time derivative term in the beginning of training when the learning rate is the highest. Upon the end of training, when the model is no longer learning efficiently, the gradients of the time derivative and Lindbladian contributions are still separated by an order of magnitude. This justifies the described gradient approximation, and we find that our method still successfully minimizes the loss (10) and converges to the correct dynamics. This heuristic significantly reduces memory consumption, thereby enabling the simulation of larger system sizes.

The fact that retaining only the gradient contribution from the time-derivative term in the loss (10) is sufficient can be understood both from a computational and physical points of view. Since time is treated as an explicit network input, the gradient of the time-derivative term involves mixed second-order derivatives $\partial_{\theta_i}\partial_t p_\theta(\mathbf{a}, t)$, whereas the Lindbladian contribution involves only first-order derivatives $\partial_{\theta_i} p_\theta(\mathbf{a}, t)$. For randomly initialized NNs, the output function is typically irregular in its inputs, leading to comparatively large curvature in time and thus larger second-order derivatives at early stages of training. As a result, the gradient from the time-derivative term dominates the optimization landscape, as confirmed by Fig. 8. From a physical point of view, the ME enforces temporal smoothness, which leads to reduction of the second-order derivative contribution; moreover, at the exact solution both gradient contributions should vanish. However, since the NQS provides only an approximate representation, the gradients remain, with the time-derivative gradient being dominant throughout training. Neglecting

the Lindbladian gradient contribution therefore constitutes a controlled approximation: the loss is evaluated according to the full expression (10), and only the dominant gradient channel is used in the back propagation.

For minimization of the loss (10) we use the standard adam optimizer for the NN optimization with multistep learning rate scheduler [84]. The scheduler parameters are as follows: total number of epochs is $N_{\text{epochs}} = 2 \times 10^4$, the learning milestones happen every 10^3 epochs, the initial learning rate is 10^{-2} and the reduction (gamma) factor is 0.75. We train our model in batches of size 200 per epoch, and each batch includes sampled probability configurations from 25 different time steps that form the subset of the target time interval, $\tau \subset [0, T]$. In our work we set $\mathcal{T} = 0.25/\Gamma$ and the time step $dt = 10^{-3}/\Gamma$ ($\mathcal{T} = 0.25/\Gamma_R$ and the time step $dt = 10^{-3}/\Gamma_R$ in the case of the chiral models, where $\Gamma_L < \Gamma_R$), such that the number of steps in $[0, T]$ is equal to 250. In Ref. [56] the τ subset was composed of time steps randomly picked from the interval $[0, T]$ for every learning epoch, such that it covered all the steps homogeneously on average. In the present work, τ range is increased gradually from 0 to $\mathcal{T}$ during an initial stage of training to facilitate a smoother loss reduction. Specifically, we start with the range of just first five steps of the total interval, increasing it gradually each 200 epochs by 5, reaching the full range after 10^4 epochs. The condition for exiting the training is either reaching the maximum number of epochs N_{epochs} or reaching the loss value that is smaller than some threshold, $\mathcal{L}_\theta < \epsilon_{\mathcal{L}}$, upon reaching the full range of τ . Throughout our work the loss threshold is set to $\epsilon_{\mathcal{L}} = 10^{-2}$.

Upon exiting the training stage, the t-NQS output $p_\theta(\mathbf{a}, t)$ is used to calculate the dynamics of physical observables appearing in Sec. V. For a given physical observable and corresponding operator O , the expectation value can be calculated as [54]:

$$\langle O \rangle \approx \sum_{\mathbf{a} \sim p_\theta(\mathbf{a}, t)} \text{tr}(OM_{(\mathbf{a})}), \quad (\text{E3})$$

where $M_{(\mathbf{a})}$ is the IC POVM frame used to generate the probability distribution $p_\theta(\mathbf{a}, t)$. At each time step of each t-NQS interval, observables are estimated using 10^3 configurations sampled from the probability distribution, $\mathbf{a} \sim p_\theta(\mathbf{a}, t)$. For all of the t-NQS results appearing in this work we additionally apply Savitzky-Golay filter to decrease the noise in the data [104]. The Savitzky-Golay filter parameters we used in the present work: the window length is $201dt$ and the polynomial order is 3.

APPENDIX F: COMPUTATIONAL COST ESTIMATION

In this Appendix we estimate the computational costs for both methods to numerically solve the ME (2), namely, the TN approach based on the vectorized density matrix presented in Sec. III and the t-NQS method based on IC-POVM representations of the density matrix presented in Sec. IV. It is important to note that the cost estimates for these methods should be compared with caution. Even though both of the approaches essentially rely on tensor multiplications and contractions, the basic operations behind the methods are structurally different. The TN approaches rely on a so-called singular value

decomposition (SVD, see Ref. [63]) and subsequent tensor contractions. In turn, the NQS simulations imply two parts: the first one is the NN learning, where the forward and backward pass as well as the loss estimations are made, and the second one is the sampling, where the physical observables are being calculated based on the data output of the trained NN [45]. Apart from the principal differences in construction, the computer hardware on which the mentioned operations are performed is also typically different: while the NN processing uses graphics processing units (GPUs), the TN procedures still mostly rely on central processing units (CPUs), even though a significant effort is being put to transfer TN processing to GPUs [105]. Due to these reasons it is very challenging to draw any conclusions from comparing the absolute performance times of the TN- and NN-based methods.

We take an alternative approach to compare computational costs by defining a set of tunable parameters that influence performance times and comparing their scalings in terms of the big O notation [106]. For instance, in the TN method the most computationally costly part is the SVD sweep, during which the TN that represents the state is simultaneously propagated and brought to the canonical form [63]. The cost behind such a sweep defines the general cost of the TN method and can be estimated as $\mathcal{O}(\mathcal{C}_{\text{TN}})$ with $\mathcal{C}_{\text{TN}}$ given in Eq. (5) of the main text. Importantly, $\mathcal{C}_{\text{TN}}(\chi, n)$ is a function of the system size n and the TN bond dimension χ . The local Hilbert space dimension d^2 can also enter this estimate, however we ignore it as it is a constant (nontunable) parameter.

The computational cost of the NQS is defined by two tunable parameters: the number of qubits n and the hidden dimension of the network d_h . We do not include the other NN parameters such as the number of layers or the batch size in our analysis due to the fact that these parameters are set constant throughout this work. Different stages of the NQS methods are influenced by the parameters n and d_h in a different way. For instance, the forward pass during the learning stage scales as $\mathcal{O}(n^2 d_h + d_h^2 n)$ [76], while the autoregressive sampling stage requires one order of n more, thus scaling as $\mathcal{O}(n^3 d_h + n^2 d_h^2)$. The most significant computational cost comes from the estimation of the loss (10), where the all-to-all Lindbladian has to be explicitly applied and which scales as $\mathcal{O}(\mathcal{C}_{\text{NQS}})$ with $\mathcal{C}_{\text{NQS}}$ given in Eq. (11) of the main text. We note that $\mathcal{C}_{\text{NQS}}(d_h, n)$ is a function of the system size n and the hidden dimension d_h .

APPENDIX G: PERTURBATIVE DERIVATION OF QUADRATIC SCALING IN FIG. 7

In this Appendix we give a brief derivation of the quadratic scaling of the peak emission intensity I_R^* with the emitter spacing parameter ε (see Fig. 7 and the corresponding discussion for more details). Consider the balanced ME given in Eqs. (2) and (3) of the main text and assume equidistant emitter spacings, $k\delta x_{j,j+1} = 2\pi(1 + \varepsilon)$, with $\varepsilon \ll 1$ being a perturbation parameter. The output emission operator in this case can be written as:

$$I_R \equiv c_R^\dagger c_R / n = \frac{1}{n} \sum_{j,\ell} e^{i2\pi(j-\ell)\varepsilon} \sigma_\ell^+ \sigma_j^-. \quad (\text{G1})$$

One can use Taylor expansion:

$$e^{i2\pi(j-\ell)\varepsilon} = 1 + i2\pi(j-\ell)\varepsilon - 2\pi^2(j-\ell)^2\varepsilon^2 + \dots, \quad (\text{G2})$$

and, assuming that the system's state at the emission peak time, ρ^* , is mostly defined by the permutationally symmetric sector, one can write the output emission intensity as a function of the spacing perturbation ε :

$$\begin{aligned} \langle I_R \rangle_{\rho^*} &= \frac{1}{n} \sum_{j,\ell} \langle \sigma_\ell^+ \sigma_j^- \rangle_{\rho^*} + \varepsilon \frac{i2\pi}{n} \sum_{j \neq \ell} \langle \sigma_\ell^+ \sigma_j^- \rangle_{\rho^*} (j-\ell) \\ &\quad - \varepsilon^2 \frac{2\pi^2}{n} \sum_{j \neq \ell} \langle \sigma_\ell^+ \sigma_j^- \rangle_{\rho^*} (j-\ell)^2 + \mathcal{O}(\varepsilon^3), \end{aligned} \quad (\text{G3})$$

where we used the notation $\langle O \rangle_{\rho^*} = \text{tr}(O\rho^*)$. Due to the approximate permutational symmetry of ρ^* , the contributions odd in power of ε are equal by modulus but opposite by sign and therefore are vanishing. If we assume that the emission intensity peaks exactly at half-filling, $\sum_j \langle \sigma_j^+ \sigma_j^- \rangle_{\rho^*} / n = 1/2$, then we can write:

$$I_R^*(\varepsilon = 0) \equiv \frac{1}{2} + \frac{1}{n} \sum_{j \neq \ell} \langle \sigma_\ell^+ \sigma_j^- \rangle_{\rho^*}. \quad (\text{G4})$$

Using the numerical value of the peak output intensity in the Dicke regime with $\varepsilon = 0$, $I_R^*(\varepsilon = 0) \approx 3.377$ (see Fig. 7), we can deduce:

$$\frac{1}{n} \sum_{j \neq \ell} \langle \sigma_\ell^+ \sigma_j^- \rangle_{\rho^*} = I_R^*(0) - \frac{1}{2} \approx 2.877, \quad (\text{G5})$$

such that with the approximate permutational symmetry of ρ^* we can write:

$$\frac{1}{n} \langle \sigma_\ell^+ \sigma_j^- \rangle_{\rho^*} \approx \frac{2.877}{n(n-1)} \forall j \neq \ell. \quad (\text{G6})$$

In this case, the quadratic term of Eq. (G3) can be factorized as:

$$-\varepsilon^2 2\pi^2 \frac{2.877}{n(n-1)} \sum_{j \neq \ell} (j-\ell)^2. \quad (\text{G7})$$

The sum appearing in Eq. (G7) can be simplified:

$$\sum_{j \neq \ell} (j-\ell)^2 = \frac{1}{6} n^2 (n^2 - 1), \quad (\text{G8})$$

such that the final expression for the quadratic term of Eq. (G3) is

$$-\varepsilon^2 \pi^2 \frac{2.877}{3} n(n+1). \quad (\text{G9})$$

This result appears as a black dotted line in Fig. 7. The quadratic fit function sketched as a white dotted line in Fig. 7 is $I_R^*(\varepsilon) = f(\varepsilon \times 10^2)$ with $f(x) \approx -0.302x^2 + 3.377$.

-
- [1] H. J. Kimble, The quantum internet, *Nature (London)* **453**, 1023 (2008).
 - [2] H. Pichler, T. Ramos, A. J. Daley, and P. Zoller, Quantum optics of chiral spin networks, *Phys. Rev. A* **91**, 042116 (2015).
 - [3] D. E. Chang, J. S. Douglas, A. González-Tudela, C.-L. Hung, and H. J. Kimble, *Colloquium*: Quantum matter built from nanoscopic lattices of atoms and photons, *Rev. Mod. Phys.* **90**, 031002 (2018).
 - [4] P. Türschmann, H. Le Jeannic, S. F. Simonsen, H. R. Haakh, S. Götzinger, V. Sandoghdar, P. Lodahl, and N. Rotenberg, Coherent nonlinear optics of quantum emitters in nanophotonic waveguides, *Nanophoton.* **8**, 1641 (2018).
 - [5] A. S. Sheremet, M. I. Petrov, I. V. Iorsh, A. V. Poshakinskiy, and A. N. Poddubny, Waveguide quantum electrodynamics: Collective radiance and photon-photon correlations, *Rev. Mod. Phys.* **95**, 015002 (2023).
 - [6] E. Shahmoon and G. Kurizki, Nonradiative interaction and entanglement between distant atoms, *Phys. Rev. A* **87**, 033831 (2013).
 - [7] W.-K. Mok, A. Asenjo-Garcia, T. C. Sum, and L.-C. Kwek, Dicke superradiance requires interactions beyond nearest neighbors, *Phys. Rev. Lett.* **130**, 213605 (2023).
 - [8] S. J. Masson, I. Ferrier-Barbut, L. A. Orozco, A. Browaeys, and A. Asenjo-Garcia, Many-body signatures of collective decay in atomic chains, *Phys. Rev. Lett.* **125**, 263601 (2020).
 - [9] N. Fayard, L. Henriët, A. Asenjo-Garcia, and D. E. Chang, Many-body localization in waveguide quantum electrodynamics, *Phys. Rev. Res.* **3**, 033233 (2021).
 - [10] S. Cardenas-Lopez, S. J. Masson, Z. Zager, and A. Asenjo-Garcia, Many-body superradiance and dynamical mirror symmetry breaking in waveguide QED, *Phys. Rev. Lett.* **131**, 033605 (2023).
 - [11] S. Mahmoodian, G. Calajó, D. E. Chang, K. Hammerer, and A. S. Sørensen, Dynamics of many-body photon bound states in chiral waveguide QED, *Phys. Rev. X* **10**, 031011 (2020).
 - [12] A. V. Poshakinskiy, J. Zhong, and A. N. Poddubny, Quantum chaos driven by long-range waveguide-mediated interactions, *Phys. Rev. Lett.* **126**, 203602 (2021).
 - [13] J. Agustí, X. H. H. Zhang, Y. Minoguchi, and P. Rabl, Autonomous distribution of programmable multiqubit entanglement in a dual-rail quantum network, *Phys. Rev. Lett.* **131**, 250801 (2023).
 - [14] S. Agarwal, E. Chaparro, D. Barberena, A. Piñeiro Orioli, G. Ferioli, S. Pancaldi, I. Ferrier-Barbut, A. Browaeys, and A. M. Rey, Directional superradiance in a driven ultracold atomic gas in free space, *PRX Quantum* **5**, 040335 (2024).
 - [15] O. Rubies-Bigorda, S. J. Masson, S. F. Yelin, and A. Asenjo-Garcia, Deterministic generation of photonic entangled states using decoherence-free subspaces, *Phys. Rev. Lett.* **134**, 213603 (2025).
 - [16] A. V. Gorshkov, A. André, M. Fleischhauer, A. S. Sørensen, and M. D. Lukin, Universal approach to optimal photon storage in atomic media, *Phys. Rev. Lett.* **98**, 123601 (2007).
 - [17] B. Gouraud, D. Maxein, A. Nicolas, O. Morin, and J. Laurat, Demonstration of a memory for tightly guided light in an optical nanofiber, *Phys. Rev. Lett.* **114**, 180503 (2015).

- [18] C. Sayrin, C. Clausen, B. Albrecht, P. Schneeweiss, and A. Rauschenbeutel, Storage of fiber-guided light in a nanofiber-trapped ensemble of cold atoms, *Optica* **2**, 353 (2015).
- [19] B. Royer, A. L. Grimsmo, A. Choquette-Poitevin, and A. Blais, Itinerant microwave photon detector, *Phys. Rev. Lett.* **120**, 203602 (2018).
- [20] A. L. Grimsmo, B. Royer, J. M. Kreikebaum, Y. Ye, K. O'Brien, I. Siddiqi, and A. Blais, Quantum metamaterial for broadband detection of single microwave photons, *Phys. Rev. Appl.* **15**, 034074 (2021).
- [21] A. Goban, C.-L. Hung, S.-P. Yu, J. Hood, J. Muniz, J. Lee, M. Martin, A. McClung, K. Choi, D. E. Chang, *et al.*, Atom-light interactions in photonic crystals, *Nat. Commun.* **5**, 3808 (2014).
- [22] N. V. Corzo, J. Raskop, A. Chandra, A. S. Sheremet, B. Gouraud, and J. Laurat, Waveguide-coupled single collective excitation of atomic arrays, *Nature (London)* **566**, 359 (2019).
- [23] M. Mirhosseini, E. Kim, X. Zhang, A. Sipahigil, P. B. Dieterle, A. J. Keller, A. Asenjo-Garcia, D. E. Chang, and O. Painter, Cavity quantum electrodynamics with atom-like mirrors, *Nature (London)* **569**, 692 (2019).
- [24] B. Kannan, M. J. Ruckriegel, D. L. Campbell, A. Frisk Kockum, J. Braumüller, D. K. Kim, M. Kjaergaard, P. Krantz, A. Melville, B. M. Niedzielski, *et al.*, Waveguide quantum electrodynamics with superconducting artificial giant atoms, *Nature (London)* **583**, 775 (2020).
- [25] A. Sipahigil, R. E. Evans, D. D. Sukachev, M. J. Burek, J. Borregaard, M. K. Bhaskar, C. T. Nguyen, J. L. Pacheco, H. A. Atikian, C. Meuwly, *et al.*, An integrated diamond nanophotonics platform for quantum-optical networks, *Science* **354**, 847 (2016).
- [26] A. P. Foster, D. Hallett, I. V. Iorsh, S. J. Sheldon, M. R. Godsand, B. Royall, E. Clarke, I. A. Shelykh, A. M. Fox, M. S. Skolnick, *et al.*, Tunable photon statistics exploiting the Fano effect in a waveguide, *Phys. Rev. Lett.* **122**, 173603 (2019).
- [27] J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi, Quantum state transfer and entanglement distribution among distant nodes in a quantum network, *Phys. Rev. Lett.* **78**, 3221 (1997).
- [28] D. E. Chang, A. S. Sørensen, E. A. Demler, and M. D. Lukin, A single-photon transistor using nanoscale surface plasmons, *Nat. Phys.* **3**, 807 (2007).
- [29] A. González-Tudela and D. Porras, Mesoscopic entanglement induced by spontaneous emission in solid-state quantum optics, *Phys. Rev. Lett.* **110**, 080502 (2013).
- [30] J. Kumlin, S. Hofferberth, and H. Peter Büchler, Emergent universal dynamics for an atomic cloud coupled to an optical waveguide, *Phys. Rev. Lett.* **121**, 013601 (2018).
- [31] J. Huber, P. Kirton, and P. Rabl, Phase-space methods for simulating the dissipative many-body dynamics of collective spin systems, *SciPost Phys.* **10**, 045 (2021).
- [32] C. D. Mink and M. Fleischhauer, Collective radiative interactions in the discrete truncated Wigner approximation, *SciPost Phys.* **15**, 233 (2023).
- [33] C.-R. Mann, F. Andreoli, V. Protsenko, Z. Lenarčič, and D. Chang, Selective radiance in super-wavelength atomic arrays, *arXiv:2402.06439*.
- [34] D. Ilin, A. V. Poshakinskiy, and A. S. Solntsev, Correlated relaxation and emerging entanglement in arrays of Λ -type atoms, *Phys. Rev. A* **111**, 043702 (2025).
- [35] R. H. Dicke, Coherence in spontaneous radiation processes, *Phys. Rev.* **93**, 99 (1954).
- [36] N. Leppen and E. Shahmoon, Quantum correlated steady states under competing collective and individual decay, *arXiv:2404.02134*.
- [37] R. Holzinger, N. S. Bassler, J. Lyne, F. G. Jimenez, J. T. Gohsrich, and C. Genes, Solving Dicke superradiance analytically: A compendium of methods, *Phys. Rev. A* **113**, 043718 (2026).
- [38] M. Gross and S. Haroche, Superradiance: An essay on the theory of collective spontaneous emission, *Phys. Rep.* **93**, 301 (1982).
- [39] J. T. Lee, S. Cardenas-Lopez, S. J. Masson, R. Trivedi, and A. Asenjo-Garcia, Exact many-body quantum dynamics in one-dimensional baths via “superspins”, *arXiv:2505.00588*.
- [40] N. S. Bassler, Absence of entanglement growth in Dicke superradiance, *Phys. Rev. A* **112**, 053713 (2025).
- [41] X. H. H. Zhang, D. Malz, and P. Rabl, Unraveling superradiance: Entanglement and mutual information in collective decay, *Phys. Rev. Lett.* **135**, 033602 (2025).
- [42] P. Rosario, L. O. R. Solak, A. Cidrim, R. Bachelard, and J. Schachenmayer, Unraveling Dicke superradiant decay with separable coherent spin states, *Phys. Rev. Lett.* **135**, 133602 (2025).
- [43] R. Holzinger and S. F. Yelin, Scaling of superradiant peak emission in spatially extended emitter arrays, *arXiv:2506.12649*.
- [44] J. Carrasquilla and G. Torlai, How to use neural networks to investigate quantum many-body physics, *PRX Quantum* **2**, 040201 (2021).
- [45] H. Lange, A. Van de Walle, A. Abedinnia, and A. Bohrdt, From architectures to applications: A review of neural quantum states, *Quantum Sci. Technol.* **9**, 040501 (2024).
- [46] M. Schmitt and M. Heyl, Simulating dynamics of correlated matter with neural quantum states, *arXiv:2506.03124*.
- [47] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, *Science* **355**, 602 (2017).
- [48] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, Variational neural-network ansatz for steady states in open quantum systems, *Phys. Rev. Lett.* **122**, 250503 (2019).
- [49] N. Yoshioka and R. Hamazaki, Constructing neural stationary states for open quantum many-body systems, *Phys. Rev. B* **99**, 214306 (2019).
- [50] A. Nagy and V. Savona, Variational quantum Monte Carlo method with a neural-network ansatz for open quantum systems, *Phys. Rev. Lett.* **122**, 250501 (2019).
- [51] M. J. Hartmann and G. Carleo, Neural-network approach to dissipative quantum many-body dynamics, *Phys. Rev. Lett.* **122**, 250502 (2019).
- [52] F. Vicentini, R. Rossi, and G. Carleo, Positive-definite parametrization of mixed quantum states with deep neural networks, *arXiv:2206.13488*.
- [53] M. Reh, M. Schmitt, and M. Gärtner, Time-dependent variational principle for open quantum systems with artificial neural networks, *Phys. Rev. Lett.* **127**, 230501 (2021).
- [54] D. Luo, Z. Chen, J. Carrasquilla, and B. K. Clark, Autoregressive neural network for simulating open quantum systems via a probabilistic formulation, *Phys. Rev. Lett.* **128**, 090501 (2022).

- [55] L. Wei, Z. Jia, Y. Wang, D. Kaszlikowski, and H. Ling, Variational transformer ansatz for the density operator of steady states in dissipative quantum many-body systems, *Phys. Rev. B* **112**, 064303 (2025).
- [56] A. Van de Walle, M. Schmitt, and A. Bohrdt, Many-body dynamics with explicitly time-dependent neural quantum states, *Mach. Learn.: Sci. Technol.* **6**, 045011 (2025).
- [57] G. Lindblad, Completely positive maps and entropy inequalities, *Commun. Math. Phys.* **40**, 147 (1975).
- [58] W.-K. Mok, A. Poddar, E. Sierra, C. C. Rusconi, J. Preskill, and A. Asenjo-Garcia, Universal scaling laws for correlated decay of many-body quantum systems, *arXiv:2406.00722*.
- [59] F. Carollo and I. Lesanovsky, Exact solution of a boundary time-crystal phase transition: Time-translation symmetry breaking and non-Markovian dynamics of correlations, *Phys. Rev. A* **105**, L040202 (2022).
- [60] A. Cabot, L. S. Muhle, F. Carollo, and I. Lesanovsky, Quantum trajectories of dissipative time crystals, *Phys. Rev. A* **108**, L041303 (2023).
- [61] I. M. Mirza and J. C. Schotland, Multiqubit entanglement in bidirectional-chiral-waveguide QED, *Phys. Rev. A* **94**, 012302 (2016).
- [62] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, Matrix product density operators: Simulation of finite-temperature and dissipative systems, *Phys. Rev. Lett.* **93**, 207204 (2004).
- [63] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, *Ann. Phys. (NY)* **326**, 96 (2011).
- [64] G. M. Crosswhite and D. Bacon, Finite automata for caching in matrix product algorithms, *Phys. Rev. A* **78**, 012356 (2008).
- [65] C. Hubig, I. P. McCulloch, and U. Schollwöck, Generic construction of efficient matrix product operators, *Phys. Rev. B* **95**, 035129 (2017).
- [66] D. Wellnitz, G. Preisser, V. Alba, J. Dubail, and J. Schachenmayer, Rise and fall, and slow rise again, of operator entanglement under dephasing, *Phys. Rev. Lett.* **129**, 170401 (2022).
- [67] G. Preisser, D. Wellnitz, T. Botzung, and J. Schachenmayer, Comparing bipartite entropy growth in open-system matrix-product simulation methods, *Phys. Rev. A* **108**, 012616 (2023).
- [68] J. Eisert, M. Cramer, and M. B. Plenio, *Colloquium: Area laws for the entanglement entropy*, *Rev. Mod. Phys.* **82**, 277 (2010).
- [69] G. Torlai and R. G. Melko, Latent space purification via neural density operators, *Phys. Rev. Lett.* **120**, 240503 (2018).
- [70] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).
- [71] G. Cybenko, Approximation by superpositions of a sigmoidal function, *Math. Control Signals Syst.* **2**, 303 (1989).
- [72] G. Carleo, F. Becca, M. Schiró, and M. Fabrizio, Localization and glassy dynamics of many-body quantum systems, *Sci. Rep.* **2**, 243 (2012).
- [73] G. Carleo, L. Cevolani, L. Sanchez-Palencia, and M. Holzmann, Unitary dynamics of strongly interacting Bose gases with the time-dependent variational Monte Carlo method in continuous space, *Phys. Rev. X* **7**, 031026 (2017).
- [74] A. Sinibaldi, C. Giuliani, G. Carleo, and F. Vicentini, Unbiasing time-dependent variational Monte Carlo by projected quantum evolution, *Quantum* **7**, 1131 (2023).
- [75] A. Sinibaldi, D. Hendry, F. Vicentini, and G. Carleo, Time-dependent neural Galerkin method for quantum dynamics, *Phys. Rev. Lett.* **136**, 120402 (2026).
- [76] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, Attention is all you need, *Adv. Neural Inf. Process. Syst.* **30** (2017).
- [77] X. Cao, Z. Zhong, and Y. Lu, Vision transformer neural quantum states for impurity models, *Phys. Rev. B* **112**, 235155 (2025).
- [78] Y.-H. Zhang and M. Di Ventura, Transformer quantum state: A multi-purpose model for quantum many-body problems, *Phys. Rev. B* **107**, 075147 (2023).
- [79] H. Lange, G. Bornet, G. Emperauger, C. Chen, T. Lahaye, S. Kienle, A. Browaeys, and A. Bohrdt, Transformer neural networks and quantum simulators: A hybrid approach for simulating strongly correlated systems, *Quantum* **9**, 1675 (2025).
- [80] L. L. Viteritti, R. Rende, and F. Becca, Transformer variational wave functions for frustrated quantum spin systems, *Phys. Rev. Lett.* **130**, 236401 (2023).
- [81] E. Ibarra-García-Padilla, H. Lange, R. G. Melko, R. T. Scalettar, J. Carrasquilla, A. Bohrdt, and E. Khatami, Autoregressive neural quantum states of Fermi Hubbard models, *Phys. Rev. Res.* **7**, 013122 (2025).
- [82] K. Sprague and S. Czischek, Variational Monte Carlo with large patched transformers, *Commun. Phys.* **7**, 90 (2024).
- [83] D. Fitzek, Y. Hong Teoh, H. Cyrus Fung, G. A. Dagnev, E. Merali, M. Schuyler Moss, B. MacLellan, and R. G. Melko, RydbergGPT, *Mach. Learn.: Sci. Technol.* **6**, 045057 (2025).
- [84] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, *et al.*, PyTorch: An imperative style, high-performance deep learning library, *Adv. Neural Inf. Process. Syst.* (2019).
- [85] M. P. Zaletel, R. S. K. Mong, C. Karrasch, J. E. Moore, and F. Pollmann, Time-evolving a matrix product state with long-ranged interactions, *Phys. Rev. B* **91**, 165112 (2015).
- [86] M. Kliesch, D. Gross, and J. Eisert, Matrix-product operators and states: NP-hardness and undecidability, *Phys. Rev. Lett.* **113**, 160503 (2014).
- [87] K. Stannigel, P. Rabl, and P. Zoller, Driven-dissipative preparation of entangled states in cascaded quantum-optical networks, *New J. Phys.* **14**, 063014 (2012).
- [88] A. Pocklington, Y.-X. Wang, Y. Yanay, and A. A. Clerk, Stabilizing volume-law entangled states of fermions and qubits using local dissipation, *Phys. Rev. B* **105**, L140301 (2022).
- [89] A. Pocklington and A. A. Clerk, Accelerating dissipative state preparation with adaptive open quantum dynamics, *Phys. Rev. Lett.* **134**, 050603 (2025).
- [90] D. Wu, R. Rossi, F. Vicentini, and G. Carleo, From tensor-network quantum states to tensorial recurrent neural networks, *Phys. Rev. Res.* **5**, L032001 (2023).
- [91] A. J. Daley, Quantum trajectories and open many-body quantum systems, *Adv. Phys.* **63**, 77 (2014).
- [92] M. T. Manzoni, D. E. Chang, and J. S. Douglas, Simulating quantum light propagation through atomic ensembles using matrix product states, *Nat. Commun.* **8**, 1743 (2017).
- [93] T. Vovk and H. Pichler, Entanglement-optimal trajectories of many-body quantum Markov processes, *Phys. Rev. Lett.* **128**, 243601 (2022).

- [94] T. Vovk and H. Pichler, Quantum trajectory entanglement in various unravelings of Markovian dynamics, *Phys. Rev. A* **110**, 012207 (2024).
- [95] D. Ilin, A. S. Solntsev, and I. Iorsh, Polaron effect in waveguide quantum optomechanics, *Phys. Rev. B* **111**, 155424 (2025).
- [96] F. Ferrari, L. Gravina, D. Eeltink, P. Scarlino, V. Savona, and F. Minganti, Steady-state quantum chaos in open quantum systems, *Phys. Rev. Res.* **7**, 013276 (2025).
- [97] T. Vovk, A. Van de Walle, H. Pichler, and A. Bohrdt, Neural quantum states for emitter dynamics in waveguide QED [Data set], Zenodo (2025), doi:[10.5281/zenodo.16850005](https://doi.org/10.5281/zenodo.16850005).
- [98] D. G. Suárez-Forero, M. Jalali Mehrabad, C. Vega, A. González-Tudela, and M. Hafezi, Chiral quantum optics: Recent developments and future directions, *PRX Quantum* **6**, 020101 (2025).
- [99] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, Reconstructing quantum states with generative models, *Nat. Mach. Intell.* **1**, 155 (2019).
- [100] K. Donatella, Z. Denis, A. Le Boité, and C. Ciuti, Dynamics with autoregressive neural quantum states: Application to critical quench dynamics, *Phys. Rev. A* **108**, 022210 (2023).
- [101] J. Nys, Z. Denis, and G. Carleo, Real-time quantum dynamics of thermal states with neural thermofields, *Phys. Rev. B* **109**, 235120 (2024).
- [102] A. van den Oord, N. Kalchbrenner, L. Espeholt, O. Vinyals, A. Graves, *et al.*, Conditional image generation with PixelCNN decoders, *Adv. Neural Inf. Process. Syst.* **29** (2016).
- [103] J. Carrasquilla, D. Luo, F. Pérez, A. Milsted, B. K. Clark, M. Volkovs, and L. Aolita, Probabilistic simulation of quantum circuits using a deep-learning architecture, *Phys. Rev. A* **104**, 032610 (2021).
- [104] A. Savitzky and M. J. Golay, Smoothing and differentiation of data by simplified least squares procedures, *Anal. Chem.* **36**, 1627 (1964).
- [105] M. Fishman, S. R. White, and E. M. Stoudenmire, The ITensor software library for tensor network calculations, *SciPost Phys. Codebases* **4** (2022).
- [106] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms* (MIT Press, Cambridge, 2022).