nature portfolio

Peer Review File

Variational Approach to Open Quantum Systems with Long-Range Competing Interactions

Corresponding Author: Mr Dawid Hryniuk

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
In this manuscript, the authors proposed a time-dependent varational Monte Carlo method based on matrix product operators to simulate open systems. A key advantage of this approach is its compatibility with long-range Ising interactions due to its sampling-based nature. Therefore, it is well-suited for the study of open system dynamics under long-ranged interactions.

This work is a natural and timely extension of the author's previous variational matrix product operator Monte Carlo method to the time-dependent settings. Its efficiency is benchmarked against tMPS method in the case of anisotropic Heisenberg model. Furthermore, the authors demonstrate the strength of their approach in handling systems with long-range and competing interactions, a regime where traditional tensor-network-based methods often fail.

In summary, this paper makes an original and valuable contribution by proposing a method to simulate open systems under certain long-ranged interactions. It opens a promising avenue to study the interaction-induced effects and non-equilibrium phase transitions in such settings. Therefore, I am pleased to recommend the publication of this paper on Communications Physics.

Below are some minor comments/questions the authors, that they might consider consider changing in the next version.

(1) Since this paper is mainly about a new algorithm, although a detailed flowchart in SM, a condensed version or schematic summarizing the algorithm's main steps in the main text would significantly improve accessibility and readability

(2) The table at the end of the SM lists the parameters used in Figures 3–5, but lacks sufficient explanation. A brief discussion on it would be helpful.

(3) Within the table, the bond dimension D and Monte Carlo sampling number N_MC are some empirical numbers in the algorithm. It would be beneficial if the authors could provide some intuition/understanding on how these numbers scale with the system properties (e.g. interaction strength, ...) or affect the convergence and accuracy of this method.

Reviewer #2

(Remarks to the Author)
The manuscript by Hryniuk and Szymanska discusses the application of a generic MPO ansatz in conjunction with time dependent Variational Monte Carlo (t-VMC) algorithms in order to simulate the dissipative dynamics of a system governed by a Lindblad Master Equation.

The manuscript is of interest to the community working on numerical methods for Open Systems and to the communities developing Strongly-correlated methods in general.
Indeed, several thematically-related manuscripts have appeared in recent times in the literature discussing similar VMC-style optimization of tensor networks, such as PRB 95, 195154 (2017), PRB 103, 235155 (2021), PRB 111, 075102 (2025)

and ArXiV:2502.13454, which discuss the optimization of PEPS using VMC+SR-style algorithms for both ground-states and finite temperature calculations.
To the extent of my knowledge the present manuscript would be the first work extending those approaches to the liouvillian dynamics, even though it eschews citing even a single of those relevant previous works.

While the work is in principle interesting, it lacks a thorough discussion and benchmarking of the previously unexplored combination of tVMC+MPO, leaving several questions unanswered.
Results are presented without supporting scaling arguments explaining and motivating the choice of hyperparameters for some well understood toy problems, which is crucial for a manuscript discussing numerical methods.
For the main hyperparameters of the algorithm, namely the bond dimension, the number of Monte-Carlo samples, and timestep size no convergence scaling is reported.

A major issue with the manuscript in the present form is that figures like Fig.4 (c-d) benchmark against a smaller system size, and it's impossible to know whether the reported calculations are accurate or not, and how changing the control hyperparameters will improve the accuracy (or not).
Similarly, in Fig.3 panels (a-c) simply tell 'the method works' but provide no insight into the method, with only Fig.3d reporting some interesting data (and again, there is no study on the effect of dt on the method itself).
A comprehensive scaling of the hyperparameters, also in numerically challenging regimes of the toy model chosen to highlight the limits of the method, should be included in a revised version of the manuscript.

Another major issue that bothers me, more about the presentation than the actual content of the manuscript, is the way the 'methodological approach' is discussed. Overall, I feel that the 't-VMPOMC approach' is 'simply' the intelligent combination t-VMC with an MPO ansatz.
I find that this is not obvious and is not surfaced clearly, distancing this manuscript from a whole body of relevant literature both in the realm of tensor networks AND variational algorithms that already studied the problem.
By presenting the algorithms as 'one element' it is hard to distinguish what is estabilished from what is not.

I believe the manuscript could be considered for publication if the various shortcoming of the present manuscript will be addressed.

Major open questions:

- MPS/MPO are known to be limited in the amount of entanglement they can represent, as the amount of entanglement along the contraction direction they can represent is bounded by the bond dimension. Nethertheless, in this manuscript you claim that your ansatz can encode long range entanglement.

If i compare your work purely against the literature of TN methods, such as PRL 114, 220601 (2015), you are using the same variational ansatz, only optimizing it with a different protocol that allows for a more favorable scaling of the computation when in presence of long range terms. However, the limitations of the ansatz remain the same.

Nevertheless, the manuscript does not analyze at all the validity of the bond dimension chosen. Is it saturated during the calculation or not? how does it depend on time (and as it is not varied over time, you can still plot the schmitt rank of the individual tensors).

In general, i believe the manuscript misses several 'numerical cross checks', starting with a systematic study of the effects of the truncation. I would expect that if interaction decayed sufficiently slowly, the minimum required bond dimension would start to explode. When does this happen? What is the minimum bond dimension for the calculations reported?

Moreover, does changing the number of samples affect the minimum required bond dimension to observe a convergence?

- The density matrix represented here is not necessarily positive, and i would expect that the non-positivity would increase as the simulation goes to later times? Moreover, i'd expect the number of samples used to have an effect of some sort to this quantity. This quantity, or some other witnesses of the non physicality of the state, should be reported.

- There is not a single convergence study in terms of number of samples in the manuscript, which is worrying for a VMC algorithm.

Minor/major points

- "In this work, we propose a new time-dependent variational matrix product operator Monte Carlo approach (t-VMPOMC for short)" I wonder how is t-VMPOMC any different from t-VMC?

- "Eq.2/3" in the literature of variational algorithms it is customary to use a 'normalised' version of the TDVP where the distance would be computed between the updates normalised according to the respective states. This is important as states are in general unnormalized and leads to an Eq.3 that would have the 'centered' jacobian appearing.
In the TN literature this is uncommon, as it is trivial to enforce the normalization of the state.

Also note that the well known 'centered expressions' for S and F that emerge from assuming 'unnormalized' states are in general considered to have considerable lower variance when estimated with MC integration because of the Cramer-Rao theorem.

I would suggest mentioning explicitly your choice to clarify your approach, and would also suggest briefly investigating whether chosing a centered or not centered expression has effects on your accuracy.

- Mention that the Z normalisation constant you chose is fundamentally unphysical?

- The S matrix is regularised or not?

- "where partially contracted matrix products are saved to memory and reused during subsequent lattice sweeps." can you elaborate somewhere about the asymptotical gain of this technique? If MCMC sampling from rho has complexity O(chi^alpha L^beta+1) assuming a subsampling of a factor of L as is common, what is the asymptotic complexity with the caching? Also can you better explain what you are doing exactly? It's hard to follow, and I would assume that every time you update a local degree of freedom you'd have to update several terms?

- In general, can you add sections/letters to the sections of the supplementary and mention it when referring to it in the main text?

- "Unlike in comparable variational Monte Carlo (VMC) algorithm" you are discussing a variational monte carlo algorithm at all effects. The only difference is that your ansatz is more structured, allowing for direct tensor contractions between the ansatz and operators (more precisely, your ansatz is multilinear, while a normal NN is nonlinear). This is similar to how autoregressive NQS are more structured than general neural networks (but less than a tensor network) thus allowing the direct calculation of marginal probabilities which can be leveraged to perform ancestral sampling.

I think presenting your algorithm as 'different from VMC' is misleading and strongly oppose this view, instead insist that you stress that this only emerges because you chose a (more structured/constrained) ansatz.

For reference, another another attempt at reducing the cost of local energy calculation by introducing more structure/constraints is arxiv:2212.11296.

- "This may be particularly valuable for chemical calculations, where the large number of non-zero matrix elements in the computation of the local estimator scales poorly with the size of the system" for normal nonlinear ansatze the number of connected entries in a second-quantized chemistry hamiltonian scales as M^4 where M is the number of orbitals. If you want to claim that the tensor contraction is advantageous, please prove it. I don't see any a-priori argument for why the contraction would not also scale as M^4.

Also explicitly mention 'second quantised'. First quantised approaches completely side-step this problem, albeit in exchane of the complexity of computing the laplacian.

- "or with arbitrary non-local terms with purely-diagonal" unless i'm mistaken, this is not a decomposition, but rather you are assuming that the long range terms of your Linbladian act diagonally? You should more clearly state here and in the introduction of the manuscript that your approach relies on this foundational assumption.

Also note that diagonal terms are 'free' for standard NQS, as they do not decrease the sparsity, which is at odds with some of your previous statements.

- "we leverage the latter observation to simulate spin lattices with otherwise prohibitively-difficult competing long-ranged Ising interactions in a subsequent part of this manuscript." I struggle to understand this sentence. As i mentioned above, this is not an observation but rather you are explicitly looking at a specific class of lindlbadians.

- "In contrast to many common tensor network time-evolution algorithms [56, 57], the variational approach does not rely on a Suzuki-Trotter decomposition of the propagator e L t , and as such is free of a Trotter error [45]." This sentence is misleading, as the TDVP is only leads to first-order updates in dt, as you had to linearize the unitary evolution?

- "In contrast again to modern VMC approaches" as discussed before, this is in contrast to nonlinear neural-network ansatze.

- METHODS section: "Tensor network results" refers to the supplementary materials, giving no insight. Can we at least get an overview of the main conclusions (possibly without proofs) in the main text?

- PRL 128, 090501 (2022) using autoregressive networks, PRL 114, 220601 (2015) and PRA 92, 022116 (2015) on using MPO for finding steady states are possibly relevant citations.

Reviewer #3

(Remarks to the Author)

Authors propose a variational tensor network method for calculation of dynamics and steady states in long-range interacting, open systems. The novelty of their approach is in combining Monte Carlo sampling with tensor network anzatz for the density matrix, using time-dependent variational principle to update the tensor network parameters. While traditional tensor network approaches are considered limited when applied to long-range interacting Hamiltonians, limiting the expenses by sampling different configuration seems to be a good alternative. They perform impressive L=200 sites calculations in one-dimension, and somewhat less impressive 4x4 calculations in two dimensions. Comparison with exact diagonalization and other tensor network methods (when applicable) convincingly shows that their method is stable and has additional advantages, such as being free of Trotter error. In the supplementary material they give a well written detailed explanation / derivation on how to optimize the procedure and what are the numerical costs of different steps. While their results do not discuss any particularly exotic physical results, this is not that problematic as their paper is methodologically oriented.

I think the paper is well written, it gives a new perspective on using tensor network in combination with Monte Carlo sampling for simulation of dissipative dynamics which could find use in many studies. I appreciate that their code is publicly available on GitHub, which even further increases the chances of it being broadly used. Therefore I would recommend the publication in the next iteration, after the below points are addressed:

Suggestions and question:

- Authors mentioned “Our implemention offers uniques technical and methodological advantages over comparable state-of-the-art tensor and neural network approaches whilst demonstrating competitive performance”. Since in their manuscript they do not benchmark different (tensor and neural network) approaches side-by-side, this statement does not seem to be based on factual comparison and should be modified accordingly.

- Add a section with more concrete explanation of advantage of this approach compared to standard tensor network contraction methods.

- While one dimensional calculations are performed on impressive L=200 sites, two-dimensional calculations results are shown on less impressive 4x4=16 sites.
-- Could the authors comment whether this method is indeed less applicable for 2d calculations?
-- Is the problem in their ‘one-dimensional’ covering of the two-dimensional sistem?

- Limitation on the time propagation were not discussed much. Most plots show t<=4 results, with the exception of Fig3(d). Authors should comment on whether there are some limitations on what propagation times can be considered.

- Can it happen that the intermediate dynamics is not capture precisely, while the steady state calculation can be rather accurate nonetheless, i.e., can intermediate times require larger bond dimension than the steady state? Do the error pile up, or can they get reduced at later times due to the dissipative nature of the setup?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I think questions raised in previous round have been properly addressed by authors, and thus I recommend its publication.

Reviewer #2

(Remarks to the Author)
I would like to commend the authors for the through work they have done, and recommend the manuscript for publication.

RESPONSE TO REFEREES' COMMENTS:
"Variational Approach to Open Quantum Systems with Long-Range Competing Interactions"

We thank all Referees for their detailed review of our work. Below we respond to all points raised by the Referees and describe the modifications we have made to the manuscript to address them. We indicated changes made to the manuscript with a red font (with the exception of the entirely new section II in the Supplementary Material, which we kept in a black font for ease of reading).

We believe that these additions further improve the clarity and comprehensiveness of our work and that the new version of the manuscript will be suitable for publication in Communications Physics.

Sincerely yours,

D. A. Hryniuk and M. H. Szymanska

RESPONSE TO REFEREE A

We thank the Referee for their highly-positive review and enthusiasm in recommending the publication of our manuscript. The Referee suggests a number of potential improvements to the readability and useability of our approach, which we respond to below.

In this manuscript, the authors proposed a time-dependent varational Monte Carlo method based on matrix product operators to simulate open quantum systems. A key advantage of this approach is its compatibility with long-range Ising interactions due to its sampling-based nature. Therefore, it is well-suited for the study of open system dynamics under long-ranged interactions.

This work is a natural and timely extension of the author's previous variational matrix product operator Monte Carlo method to the time-dependent settings. Its efficiency is benchmarked against tMPS method in the case of anisotropic Heisenberg model. Furthermore, the authors demonstrate the strength of their approach in handling systems with long-range and competing interactions, a regime where traditional tensor-network-based methods often fail.

In summary, this paper makes an original and valuable contribution by proposing a method to simulate open quantum systems under certain long-ranged interactions. It opens a promising avenue to study the interaction-induced effects and non-equilibrium phase transitions in such settings. Therefore, I am pleased to recommend the publication of this paper on Communications Physics.

Below are some minor comments/questions the authors, that they might consider consider changing in the next version.

Comment 1: *Since this paper is mainly about a new algorithm, although a detailed flowchart in SM, a condensed version or schematic summarizing the algorithm's main steps in the main text would significantly improve accessibility and readability.*

Reply: We found that expanding the relevant part of the "Methods" section by providing a summary of the main steps of the algorithm would be most beneficial to the reader. The revised manuscript provides these changes.

Comment 2: *The table at the end of the SM lists the parameters used in Figures 3–5, but lacks sufficient explanation. A brief discussion on it would be helpful.*

Reply: We believe that the extensive discussion on the parameters and hyperparameters in our algorithm included in the newly added section II on convergence analysis in the SM equally meets the Referee's request for clarification on the same parameters listed in the table.

Comment 3: *Within the table, the bond dimension D and Monte Carlo sampling number N_{MC} are some empirical numbers in the algorithm. It would be beneficial if the authors could provide some intuition/understanding on how these numbers scale with the system properties (e.g. interaction strength, ...) or*

affect the convergence and accuracy of this method.

Reply: The optimal hyperparameter values required for convergence are naturally strongly model- and model-parameter-dependent. In general, the required bond dimension grows with the amount of entanglement present during the transient dynamics, while the number of Monte Carlo samples must be sufficient enough to accommodate the increase in the number of variational parameters at larger bond dimensions and of the increase of the size of the Hilbert space when increasing the size of the system. Both hyperparameters are chosen empirically, as the Referee rightly pointed out. Our revised SM now includes comprehensive convergence checks across six distinct models and with respect to five key hyperparameters used during our simulations. In particular, we compare the convergence for two variants of a dissipative Ising chain with weak and strong long-ranged competing interactions (Figs. 12-31), where we expectedly observe more stringent demands on the bond dimension and number of samples when the interactions are stronger. We summarise some common observations on the convergence and form recommendations on suitable hyperparameter values.

RESPONSE TO REFEREE B

We thank the Referee for their detailed review of our manuscript and for recommending its publication after their remarks are addressed. The Referee's main concerns relate to the lack of convergence studies in our manuscript, the presentation of our method and its relation to existing literature, the correct use of acronyms, as well as clarification of several other ambiguities. However, we believe that several of the Referee's criticisms of our work stem from a misunderstanding of: the intricacies of open quantum systems and non-unitary evolution, the associated literature, the specifics of our method and its implementation, and the scope and purposes of our results. We address all points raised by the Referee below.

The manuscript by Hryniuk and Szymanska discusses the application of a generic MPO ansatz in conjunction with time dependent Variational Monte Carlo (t-VMC) algorithms in order to simulate the dissipative dynamics of a system governed by a Lindblad Master Equation.

Comment 1: *The manuscript is of interest to the community working on numerical methods for Open Quantum Systems and to the communities developing Strongly-correlated methods in general. Indeed, several thematically-related manuscripts have appeared in recent times in the literature discussing similar VMC-style optimization of tensor networks, such as PRB 95, 195154 (2017), PRB 103, 235155 (2021), PRB 111, 075102 (2025) and ArXiV:2502.13454, which discuss the optimization of PEPS using VMC+SR-style algorithms for both ground-states and finite temperature calculations.*
To the extent of my knowledge the present manuscript would be the first work extending those approaches to the liouvillian dynamics, even though it eschews citing even a single of those relevant previous works.

Reply 1: We respectfully disagree with the classification of our work as an extension of the above referenced approaches. PRB 95, 195154, PRB 103, 235155, and ArXiV:2502.13454 describe methods for finding the ground states of correlated spins or electrons using VMC with a PEPS ansatz, which is substantially different in both theme and technique from our approach to solving for the *non-unitary dynamics* of *open quantum systems* using *time-dependent* VMC (t-VMC) with an *MPO* ansatz. PRB 111, 075102, arguably the closest paper to ours from the above list in that it utilizes MPDO (among other ansätze) to treat mixed (thermal) states, does so by performing imaginary-time evolution to minimize a free energy functional. Of the available literature on hybrid tensor network and VMC methods, the original paper by Sandvik and Vidal [1], detailing VMC optimization of translationally-invariant MPS, was most relevant to us in the completion of our prior work on finding non-equilibrium steady states with VMC and MPO [2] (and which we cite in that paper), from which the present manuscript is a continuation of. As such, we do not consider the works listed by the Referee as meriting citation in our manuscript on this basis, as opposed to the over 20 relevant works on (t-)VMC and tensor network simulation methods for many-body (open) quantum lattices that we built our approach on and which are already cited in our manuscript. However, we are happy to reference some of the above related works to better motivate and contextualize our approach among other hybrid proposals for problems in other types of strongly-correlated systems, which we now do in our revised introduction.

Comment 2: *While the work is in principle interesting, it lacks a thorough discussion and benchmarking of the previously unexplored combination of tVMC+MPO, leaving several questions unanswered. Results are presented without supporting scaling arguments explaining and motivating the choice of hyperparameters for some well understood toy problems, which is crucial for a manuscript discussing numerical methods. For the*

main hyperparameters of the algorithm, namely the bond dimension, the number of Monte-Carlo samples, and timestep size no convergence scaling is reported.

A major issue with the manuscript in the present form is that figures like Fig.4 (c-d) benchmark against a smaller system size, and it's impossible to know whether the reported calculations are accurate or not, and how changing the control hyperparameters will improve the accuracy (or not). Similarly, in Fig.3 panels (a-c) simply tell 'the method works' but provide no insight into the method, with only Fig.3d reporting some interesting data (and again, there is no study on the effect of dt on the method itself). A comprehensive scaling of the hyperparameters, also in numerically challenging regimes of the toy model chosen to highlight the limits of the method, should be included in a revised version of the manuscript.

Reply 2: The Referee will be happy to know that our newly-revised SM now includes extensive convergence tests with respect to five crucial hyperparameters across six illustrative models of dissipative spin chains with short- and long-range competing interactions, comprising a total of 60 distinct figures. These additions constitue a convergence analysis more comprehensive than of any other works known to us, whether for closed or open quantum systems.

We note that the convergence analysis for the step size was done in terms of the convergence with respect to the adaptive error tolerance ϵ_{tol} instead of the step size dt as explicitly requested by the Referee in their comment above. A fixed-step forward Euler integration was employed only in Fig. of the main text for the sole purpose of comparison with a fixed-step-sized Trotterized MPO time-evolution algorithm; this integration scheme is very inefficient and inaccurate when compared to the second-order adaptive Heun integration method used in all other of our simulations.

Comment 3: *Another major issue that bothers me, more about the presentation than the actual content of the manuscript, is the way the 'methodological approach' is discussed. Overall, I feel that the 't-VMPOMC approach' is 'simply' the intelligent combination t-VMC with an MPO ansatz. I find that this is not obvious and is not surfaced clearly, distancing this manuscript from a whole body of relevant literature both in the realm of tensor networks AND variational algorithms that already studied the problem. By presenting the algorithms as 'one element' it is hard to distinguish what is estabilished from what is not.*

Reply 3: It is unfortunate that the Referee's impression of our use of the quoted acronym amounts to a malign misplacement of our new proposal among existing algorithms. We did not intend to deny categorising our method as a t-VMC algorithm, nor as a tensor network algorithm, and we are happy for our method to be categorised as either by the readers of our manuscript.

However, we strongly oppose the Referee's claim that we "[distanced] this manuscript from a whole body of relevant literature both in the realm of tensor networks AND variational algorithms that already studied the problem". References to relevant papers on (t-)VMC and tensor network methods number more than 20 in the original version of our manuscript. Common methodological aspects, such as computation of local estimators and observables, regularization of the variational metric, time-integration, and tensor contractions, are discussed in relation to existing works in multiple places in our manuscript and the SM. Finally, as is emphasized several times across our manuscript, the principal goal of this work was to enable the simulation of dissipative systems with competing long-ranged interactions, which have **not** been studied by existing tensor network and variational algorithms, as stated or implied by the Referee in their comment.

I believe the manuscript could be considered for publication if the various shortcoming of the present manuscript will be addressed.

Comment 4: *MPS/MPO are known to be limited in the amount of entanglement they can represent, as the amount of entanglement along the contraction direction they can represent is bounded by the bond dimension. Nethertheless, in this manuscript you claim that your ansatz can encode long range entanglement.*

Reply 4: This is false. Our manuscript makes no claims whatsoever about the ability of the ansatz to encode long range entanglement. Instead, our results demonstrate that our novel algorithm can accurately capture the properties of our targeted models. It is well known to practitioners of the field that, generically, dissipation limits long-range correlations due to entangling dynamics as a result of decoherence, which contributes the success of our approach.

Comment 5: *If i compare your work purely against the literature of TN methods, such as PRL 114, 220601 (2015), you are using the same variational ansatz, only optimizing it with a different protocol that allows for a more favorable scaling of the computation when in presence of long range terms. However, the*

limitations of the ansatz remain the same.

Reply 5: This is again false. The ansatz used in the quoted reference is a generic MPO with open boundary conditions, for which canonical forms can be found. Our prescription makes use primarily of translationally-invariant MPOs with fixed unit cell and periodic boundary conditions, for which canonical forms may not be found in general (and due to which, many established techniques are inapplicable or inefficient for said class of MPOs [3, 4]). To our best knowledge, our approach is the first to make use of such looped MPOs for the purpose of simulating the dynamics of dissipative many-body quantum lattices, and to do so with remarkable success.

Furthermore, the referenced paper makes use of a deterministic time-independent variational optimisation algorithm to target the steady state only (completely bypassing the strongly-entangling transient dynamics), which is fundamentally different to our prescription which accurately captures the non-equilibrium dynamics (in addition to the steady state) through the use of time-dependent VMC in combination with efficient tensor contraction subroutines. Therefore, the comparison between the two methods cannot be reduced simply to the choice of the ansätze (which are not equivalent in the first place).

Comment 6: *Nevertheless, the manuscript does not analyze at all the validity of the bond dimension chosen. Is it saturated during the calculation or not? how does it depend on time (and as it is not varied over time, you can still plot the schmitt rank of the individual tensors).*

Reply 6: We kindly invite the Referee to consult the relevant figures in the newly-added convergence tests. For the Ising model with short- and long-range competing interactions, a bond dimension of $\chi = 20$ appears to be comfortably sufficient to accurately capture the dynamics of the system, as evidenced by the smallest Schmidt coefficients saturating at about 10^{-6} during the time-evolution. In contrast, for the XYZ model, the maximal bond dimension considered, $\chi = 24$, appears insufficient to accurately capture the dynamics of some of the highly-non-local correlation functions. We stress that these observations do not undermine the effectiveness of our algorithm; instead, they demonstrate that the chosen hyperparameter values and allocated computational resources (a hard limit of 480 combined CPU hours for each time series) should be considered insufficient in this example.

Comment 7: *In general, i believe the manuscript misses several 'numerical cross checks', starting with a systematic study of the effects of the truncation. I would expect that if interaction decayed sufficiently slowly, the minimum required bond dimension would start to explode. When does this happen? What is the minimum bond dimension for the calculations reported?*

Reply 7: The relevant figures from the convergence checks to compare are Figs. 12-21 and 32-41 in SM II. We are considering the convergence of competing long-ranged Ising chains with long-range competing interactions with decay exponents $(3, 6)$ and $(2, 4)$, without Kac-normalization. As one would expect, the more extended interactions put somewhat stringer requirements on the hyperparameters.

We chose decay constants with $\alpha > 1$ since interactions with $\alpha \leq 1$ are usually renormalized to satisfy size-extensivity; such renormalization would weaken the strength of the interaction at shorter distances, which may give the impression of an "easier" simulation at longer-ranged interactions. If we set $\alpha \ll 1$ without Kac-renormalization, the bond dimension would "explode" for larger systems, as put by the Referee. However, the usefulness of such size-dependent calculations seems dubious to us.

Comment 8: *Moreover, does changing the number of samples affect the minimum required bond dimension to observe a convergence?*

Reply 8: As can be seen in the new convergence tests, a larger number of Monte Carlo samples does not improve the convergence when the the bond dimension is insufficent (when compared to simulations at larger bond dimension). Therefore, an insufficient number of samples can preclude convergence, but an increased number of samples won't make up for an insufficient bond dimension.

In addition, a larger bond dimension leads to a larger total number of variational parameters, which effectively increases the number of required samples. Knowing also that computational complexity scales linearly with the number of samples but cubically with the bond dimension, it is threrefore not a practical consideration to reduce the number of samples at the expense of a larger bond dimension.

Comment 9: *The density matrix represented here is not necessarily positive, and i would expect that the non-positivity would increase as the simulation goes to later times? Moreover, i'd expect the number of samples used to have an effect of some sort to this quantity. This quantity, or some other witnesses of the*

non physicality of the state, should be reported.

Reply 9: It is correct to note that the variational manifold permits density matrices that are not positive semi-definite. However, we are not sure why the Referee seems to be under the impression that non-positivity should be increasing over time; the variational dynamics follow that of the Lindbladian, a completely-positive map.

In the added convergence checks in SM II, for simulations with $N = 10$ sites, we plot the value of the smallest eigenvalue of the Hermetized variational density matrix in subplots 7b. A negative minimum eigenvalue is equivalent to the density matrix being non-positive semi-definite. As can be seen, the minimum eigenvalue in all of our simulations stays positive (starting from 0). Therefore, positivity is maintained throughout and in all of our simulations analysed for $N = 10$ sites, regardless of any hyperparameter values (including number of samples).

Comment 10: *There is not a single convergence study in terms of number of samples in the manuscript, which is worrying for a VMC algorithm.*

Reply 10: SM II now features convergence analyses with the number of samples. However, we would like to point out that it is customary for cornerstone NQS VMC publications on open quantum system to not feature convergence checks with the number of samples [5, 6, 7, 8, 9], including in works recently published in Communications Physics [10]. Therefore, the Referee's concern above is at odds with common practices in the existing literature. Nevertheless, we agree that all-manner of comprehensive convergence checks is, in principle, bound to benefit the quality of methodological publications.

Comment 11: *"In this work, we propose a new time-dependent variational matrix product operator Monte Carlo approach (t-VMPOMC for short)" I wonder how is t-VMPOMC any different from t-VMC?*

Reply 11: The acronym "t-VMPOMC" was used by us to label the method introduced and showcased in our manuscript. It can be considered a time-dependent extension of our earlier published VMPOMC method [2]. The term "t-VMC" describes a particular family of variational stochastic time-evolution algorithms, of which t-VMPOMC is a member of.

Given the Referee's strong opposition to our use of the "t-VMPOMC" acronym, we have replaced it with "t-VMC+MPO" in the main text and SM.

Comment 12: *"Eq.2/3" in the literature of variational algorithms it is customary to use a 'normalised' version of the TDVP where the distance would be computed between the updates normalised according to the respective states. This is important as states are in general unnormalized and leads to an Eq.3 that would have the 'centered' jacobian appearing. In the TN literature this is uncommon, as it is trivial to enforce the normalization of the state.*

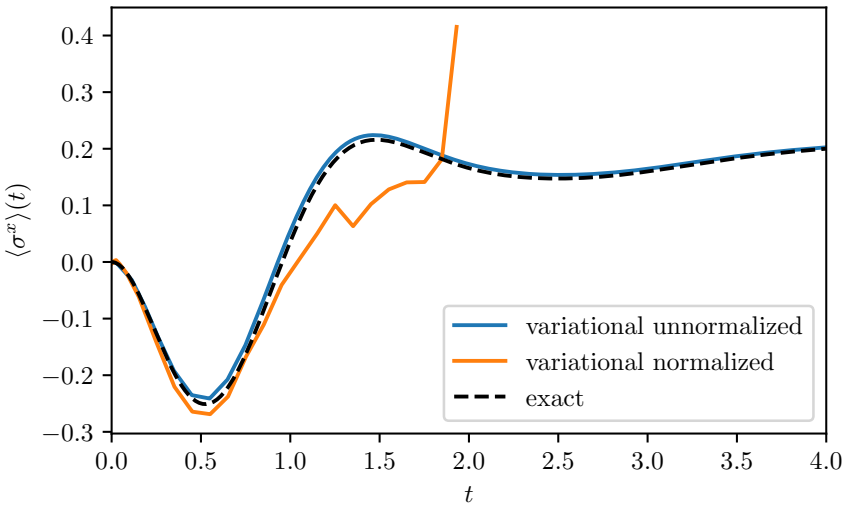
Also note that the well known 'centered expressions' for S and F that emerge from assuming 'unnormalized' states are in general considered to have considerable lower variance when estimated with MC integration because of the Cramer-Rao theorem.

I would suggest mentioning explicitly your choice to clarify your approach, and would also suggest briefly investigating whether chosing a centered or not centered expression has effects on your accuracy.

Reply 12: As is unambiguously stated by Eq. 3 in the main text, our results were obtained without further normalisation of the states entering the TDVP. While it is possible to enforce normalisation of the vectors according the the Hilbert-Schmidt norm, as imposed by our variational principle, which would lead to covariant expressions for the metric and variational forces, this would be erroneous as it would violate trace-preservation of the dynamics. Indeed, performing this normalisation, then renormalising the tensor elements after each iteration to restore unit trace, produces results that are inferior to our naive "unnormalised" approach, or which fail to converge at all. We explicitly demonstrate this in the figure below for the Ising chain.

In principle, it may be possible to enforce trace preservation by renormalising the vectors by the trace norm. However, this is not possible in practice for larger systems as it would require computing the square root of an uncontracted tensor network (the general impracticality of the trace norm in simulation algorithms for Liouvillian dynamics was touched upon in previous publications [11, 12, 5, 2]).

While we appreciate the Referee's concerns, we would like to stress that the non-unitary evolution of mixed states presents additional challenges not encountered with pure states; "customary" practices from



the TDVP literature may not be applicable in the more general Liouvillian case. As such, we don't believe we had much of a choice in the above matter in our implementation that we ought to explicitly discuss in our manuscript. Furthermore, since the main goal of this study was to enable simulation of otherwise inaccessible open quantum systems with complex long ranged interactions rather than performance chasing, we consider the application of potential variance reduction techniques as far beyond the scope of this work.

Comment 13: *Mention that the Z normalisation constant you chose is fundamentally unphysical?*

Reply 13: The normalisation constant Z is the purity of the density matrix, which we would argue is a physically meaningful property of the quantum system. However, we don't see how commenting on the (un)physicality of the normalisation constant in our algorithm has any bearing on its functionality and the obtained results.

Comment 14: *The S matrix is regularised or not?*

Reply 14: Yes, this is discussed in the SM (and now also in the revised METHODS section).

Comment 15: *"where partially contracted matrix products are saved to memory and reused during subsequent lattice sweeps." can you elaborate somewhere about the asymptotical gain of this technique? If MCMC sampling from ρ has complexity $O(\chi^\alpha L^\beta + 1)$ assuming a subsampling of a factor of L as is common, what is the asymptotic complexity with the caching? Also can you better explain what you are doing exactly? It's hard to follow, and I would assume that every time you update a local degree of freedom you'd have to update several terms?*

Reply 15: Sequentially updating N degrees of freedom one at a time via the Metropolis algorithm requires access to the overlaps $\langle \boldsymbol{x}|\rho\rangle$. With $\{L\}$ and $\{R\}$ matrix products cached, the computation of these overlaps per sample $\boldsymbol{x}$ requires 2 matrix multiplications per degree of freedom (joining the constant left partial matrix product, the matrix resulting from contracting the MPO unit cell along the updated physical index, and constant right partial matrix product). Asymptotically, this reduce the cost of sampling by a factor of N if the constant left and right matrix products were to be computed on the spot each time in a lazy implementation, though it comes at the minor expense of up-front memorisation of the set of partial matrix products $\{L\}$ and $\{R\}$. Crucially, the same $\{L\}$ and $\{R\}$ can be reused repeatedly in subsequent parts of the t-VMC+MPO algorithm for similar savings in cost. We hope the above helps clarify the Referee's questions.

Comment 16: *In general, can you add sections/letters to the sections of the supplementary and mention it when referring to it in the main text?*

Reply 16: Yes, we categorised our SM into ordered sections that we now refer to in the main text.

Comment 17: *"Unlike in comparable variational Monte Carlo (VMC) algorithm" you are discussing a variational monte carlo algorithm at all effects. The only difference is that your ansatz is more structured,*

allowing for direct tensor contractions between the ansatz and operators (more precisely, your ansatz is multilinear, while a normal NN is nonlinear). This is similar to how autoregressive NQS are more structured than general neural networks (but less than a tensor network) thus allowing the direct calculation of marginal probabilities which can be leveraged to perform ancestral sampling.

I think presenting your algorithm as 'different from VMC' is misleading and strongly oppose this view, instead insist that you stress that this only emerges because you chose a (more structured/constrained) ansatz.

For reference, another another attempt at reducing the cost of local energy calculation by introducing more structure/constraints is arxiv:2212.11296.

Reply 17: We are genuinely puzzled as to why the Referee seems to be under the impression that we deny categorizing our approach as a VMC algorithm. As we explained in an earlier reply, we are perfectly happy for our approach to be acknowledged as either a VMC or tensor network algorithm. The Referee's assessment that the additional structure possessed by a tensor network ansatz, and absent in other ansätze, is what enables—by means of contraction between the ansatz and operators—exact evaluation of quantities that must normally be sampled in VMC algorithms, is precisely what we explain in the sentence that is partially quoted by the Referee above (in the specific context of the local estimator, though the same arguments apply to related quantities). We did not intend for this sentence to imply that our algorithm does not belong to the class of VMC algorithms; we used the adjective "comparable" to indicate the precise opposite, i.e. that our method **should** be compared with other VMC approaches (that may differ in the use of a less structured ansatz and the consequences thereof). To make the comparison more explicit, we have modified the quoted sentence as follows: "Unlike in comparable variational Monte Carlo (VMC) algorithms reliant on non-linear neural-network architectures..."

Comment 18: *"This may be particularly valuable for quantum chemical calculations, where the large number of non-zero matrix elements in the computation of the local estimator scales poorly with the size of the system" for normal nonlinear ansatze the number of connected entries in a second-quantized chemistry hamiltonian scales as M^4 where M is the number of orbitals. If you want to claim that the tensor contraction is advantageous, please prove it. I don't see any a-priori argument for why the contraction would not also scale as M^4.*

Also explicitly mention 'second quantised'. First quantised approaches completely side-step this problem, albeit in exchane of the complexity of computing the laplacian.

Reply 18: This remark was targeted specifically at translationally invariant models as studied in quantum chemistry (e.g. Hubbard-type models), for which the improvement in computational complexity is discussed in detail in SM I.B. However, upon reading the referee's comments, we realised that this sentence would be misleading if understood more broadly and have made the decision to retract it.

Comment 19: *"or with arbitrary non-local terms with purely-diagonal" unless i'm mistaken, this is not a decomposition, but rather you are assuming that the long range terms of your Linbladian act diagonally? You should more clearly state here and in the introduction of the manuscript that your approach relies on this foundational assumption.*

Reply 19: We are not sure what the Referee is confused about in the first sentence. The above remark singles out Lindbladian terms that can be written diagonally in some basis, which leads to convenient cancellations that significantly reduce the overall computational cost. In general, the Lindbladian, just like the Hamiltonian, consists of both, diagonal and non-diagonal parts. In that sense, it's possible to decompose the Lindbladian into diagonal and non-diagonal parts, with the diagonal part being trivially accounted for.

The Referee seems to be also under the mistaken impression that our method is limited strictly to diagonal long-ranged interactions. The Ising model with (diagonal) competing long-ranged interactions was chosen primarily as a minimal model of an open quantum systems that exhibits novel modulated magnetic ordering in the steady state, secondarily due to its importance to modern experiments, and thirdly because diagonal long-ranged interactions are trivial to treat exactly (which this section of our manuscript highlights). However, our implementation is not at all limited to simulating diagonal long-ranged interactions, as we now demonstrate in our revised manuscript in which we simulate a dissipative spin chain with non-diagonal competing long-ranged anisotropic XYZ interactions.

Therefore, instead of proceeding to "clearly state here and in the introduction of the manuscript that your approach relies on this foundational assumption" as requested by the Referee, we further enlarged

our claim to explicitly include arbitrary (diagonal and non-diagonal, isotropic and anisotropic) long-ranged spatially-decaying interactions as among efficiently simulable via our approach.

Comment 20: *Also note that diagonal terms are 'free' for standard NQS, as they do not decrease the sparsity, which is at odds with some of your previous statements.*

"we leverage the latter observation to simulate spin lattices with otherwise prohibitively-difficult competing long-ranged Ising interactions in a subsequent part of this manuscript." I struggle to understand this sentence. As i mentioned above, this is not an observation but rather you are explicitly looking at a specific class of lindlbadians.

Reply 20: We are not aware any similar procedures of exploiting the diagonally of the Lindbladian to efficiently account for long-ranged interactions in open quantum systems in the available literature. Therefore, we consider the term "observation" as an appropriate descriptor of our recognition and subsequent exploitation of this feature of our proposal.

Further, we strongly, but respectfully, disagree with the Referee's characterization of our work as merely "explicitly looking at a specific class of lindlbadians." The area of Markovian and non-Markovian open quantum systems has seen growing demand over recent years for new methods capable of accurately solving the many-body problem. Models with long-ranged interactions make a particularly compelling case for the development of new methods owing to the emergence of new experimental platforms and quantum devices which rely on long-range couplings as a critical mechanism of action (as discussed in our introduction). In many such cases, including e.g. Rydberg arrays or some configurations of trapped ions, the long-ranged interactions take a diagonal form. Moreover, as evidenced by the new section on non-diagonal interactions added to the manuscript, our method is perfectly capable of treating non-diagonal long-ranged couplings; however, the diagonal interactions that we chose to concentrate on already underpin a significant segment of the experimentally-realizable platforms, all while exhibiting novel emergent behaviour in the competing case that we wished to highlight with our work. It is beyond any doubt to us that the highly efficient method we introduced in this work, which enables the treatment of dissipative lattices with diagonal and non-diagonal long-ranged interactions, will become a cornerstone approach to and of value to a growing number of researchers investigating such systems.

Comment 21: *"In contrast to many common tensor network time-evolution algorithms [56, 57], the variational approach does not rely on a Suzuki-Trotter decomposition of the propagator e L t , and as such is free of a Trotter error [45]." This sentence is misleading, as the TDVP is only leads to first-order updates in dt, as you had to linearize the unitary evolution?*

Reply 21: We admit that this sentence was phrased sloppily. We recognize the importance of time-discretization errors in both variational and Trotterized algorithms. Indeed, with the exception of Fig. 3d of the main text, we used a second-order adaptive integration scheme in all of our simulations precisely to minimize errors due to time-discretization. However, the comparison between the TDVP and Trotterized evolution in this plot and the quoted sentence was meant to be interpreted in the specific context of steady state calculations. Extraction of the steady state at long times from Trotterized time-evolution simulations is not an uncommon practice. In such approaches, time-discretization errors are expected to pile-up over time, leading to incorrect steady states. In contrast, one might expect the repeated projection in TDVP algorithms to suppress (to some limited extent) the acquired time-discretization errors when nearing the fixed point, i.e. the steady state, provided that the bond dimension is sufficently large for the variational manifold to (approximately) contain the steady state and any transient states leading towards it. Indeed, as Fig. 3d demonstrates, the steady-state magnetizations computed with second-order Trotterized t-MPS plateau at some values away from the true expected value, with the error decreasing as the time-step size decreases. In contrast, our crude variational approach with linear updates captures the steady states more accurately than any of our t-MPS calculations. We have added a follow-up sentence in the main text to expand upon this remark.

Comment 22: *"In contrast again to modern VMC approaches" as discussed before, this is in contrast to nonlinear neural-network ansatze.*

Reply 22: The Referee will be well aware that variational families of solutions to the quantum many-body problem extend beyond tensor network and neural quantum states, with e.g. Jastrow-type wavefunctions also devoid of the multilinear structure of tensor networks. However, it is true that, in the specific context of VMC methods for open quantum systems, all hitherto published approaches—to the best of our

knowledge—have been limited to NQS ansätze, so we are happy to qualify our comparison to NQS specifically.

Comment 23: *METHODS section: "Tensor network results" refers to the supplementary materials, giving no insight. Can we at least get an overview of the main conclusions (possibly without proofs) in the main text?*

Reply 23: We decided to include a step-by-step outline of a single iteration of our method in the revised METHODS section. Relevant sections from the SM that provide additional details on the involved computations are referenced where appropriate.

Comment 24: *PRL 128, 090501 (2022) using autoregressive networks, PRL 114, 220601 (2015) and PRA 92, 022116 (2015) on using MPO for finding steady states are possibly relevant citations.*

Reply 24: PRL 114, 220601 (2015) was already cited in our manuscript under reference 60 (now 66). We cited the remaining two works suggested by the referee in the above comment in our revised manuscript where appropriate.

RESPONSE TO REFEREE C

We thank the Referee for their positive reception of our work and for recommending its publication. The Referee's main concerns relate to the limitations of our approach (particularly for two-dimensional systems and with regards to time-propagation lengths), which we address below.

Authors propose a variational tensor network method for calculation of dynamics and steady states in long-range interacting, open quantum systems. The novelty of their approach is in combining Monte Carlo sampling with tensor network anzatz for the density matrix, using time-dependent variational principle to update the tensor network parameters. While traditional tensor network approaches are considered limited when applied to long-range interacting Hamiltonians, limiting the expenses by sampling different configuration seems to be a good alternative. They perform impressive L=200 sites calculations in one-dimension, and somewhat less impressive 4x4 calculations in two dimensions. Comparison with exact diagonalization and other tensor network methods (when applicable) convincingly shows that their method is stable and has additional advantages, such as being free of Trotter error. In the supplementary material they give a well written detailed explanation / derivation on how to optimize the procedure and what are the numerical costs of different steps. While their results do not discuss any particularly exotic physical results, this is not that problematic as their paper is methodologically oriented.

I think the paper is well written, it gives a new perspective on using tensor network in combination with Monte Carlo sampling for simulation of dissipative dynamics which could find use in many studies. I appreciate that their code is publicly available on GitHub, which even further increases the chances of it being broadly used. Therefore I would recommend the publication in the next iteration, after the below points are addressed:

Suggestions and question:

Comment 1: *Authors mentioned "Our implemention offers uniques technical and methodological advantages over comparable state-of-the-art tensor and neural network approaches whilst demonstrating competitive performance". Since in their manuscript they do not benchmark different (tensor and neural network) approaches side-by-side, this statement does not seem to be based on factual comparison and should be modified accordingly.*

Reply 1: We agree and have appropriately rephrased this sentence.

Comment 2: *Add a section with more concrete explanation of advantage of this approach compared to standard tensor network contraction methods.*

Reply 2: We clarified the differences of our approach to standard tensor network contraction methods in the relevant paragraph of the main text, particularly in relation to errors in the time-propagation and construction of the Lindbladian superoperator. We don't consider the differences to be extensive enough yet general (i.e. not specific to particular instances of tensor network algorithms) to merit the addition of a

separate section beyond what is now discussed in this paragraph, however.

Comment 3: *While one dimensional calculations are performed on impressive L=200 sites, two-dimensional calculations results are shown on less impressive 4x4=16 sites.*

Could the authors comment whether this method is indeed less applicable for 2d calculations?

Is the problem in their 'one-dimensional' covering of the two-dimensional sistem?

Reply 3: The principal limitation of our approach in the realm of two-dimensional lattices is due to the increase in the required number variational parameters. For translationally-invariant lattices, the one-dimensional covering inherits translational invariance along one dimension only. In the case of the 4×4 calculations reported, this necessitates the use of four times as many variational parameters as compared to the one-dimensional calculations considered in other parts of our work. Therefore, the difficulty of this calculation for a bond dimension of 10 is comparable to the difficulty of a calculation with bond dimension 20 in the latter case. In addition, the one-dimensional covering, by its construction, requires a much larger bond dimension to correctly account for the entanglement in the "transverse" direction, further pushing up the number of required variational parameters.

However, we would insist that the discrepancies seen in the plots are still primarily due to finite-size effects rather than errors. We revised the figure to include variational dynamics for a 3×3 lattice, where we note very good agreement with exact results up to around $t = 2$. Hence, we expect the strong differences between the dynamics for 3×3 and 4×4 lattices at early times (noting that the initial condition is an unentangled product state) to stem from the increase in the lattice size instead of errors, which are expected to become large at longer times once the systems becomes highly entangled.

We would like to take this opportunity to stress that, to our best knowledge, no other tensor network method capable of describing competing long-ranged interactions in one, let alone two, dimensions exists. While we recognise that our results for the two-dimensional case are far from perfect, we hope that the inclusion of this figure will prompt other researchers to formulate yet more powerful algorithms for this important class of dissipative models by encouraging them to improve upon our results.

Comment 4: *Limitation on the time propagation were not discussed much. Most plots show t_f=4 results, with the exception of Fig3(d). Authors should comment on whether there are some limitations on what propagation times can be considered.*

Reply 4: The time-regime of $0 < t < 4$ was chosen to highlight the more complex transient dynamics that our method is able to capture accurately. In the new convergence results added to the SM, the dynamics are plotted for twice as long, comfortably settling at the steady state (naturally, some simulations did not reach $t = 8$ given the limited allocated computational resources, or failed outright due to inappropriate hyperparameter values).

Generally, within the adaptive step-sizing method utilized in near all of our simulations, the step size would usually increase significantly if not capped owing to the drop in the error in the Heun integration scheme as the dynamics approaches the fixed point. Therefore, we expect that, generally, given sufficient resources and large-enough bond dimension to accurately overcome the highly-entangling transient dynamics, very long propagation times can be attained at modest additional cost (indeed, the steady-state results in Fig. 5 in the main text were extracted at large $t = 30$). However, this may be model-dependent.

Comment 5: *Can it happen that the intermediate dynamics is not capture precisely, while the steady state calculation can be rather accurate nonetheless, i.e., can intermediate times require larger bond dimension than the steady state? Do the error pile up, or can they get reduced at later times due to the dissipative nature of the setup?*

Reply 5: It is indeed the case that transient dynamics generally require a (sometimes significantly) larger bond dimension than the steady-state [13], where the entanglement is curtailed by dissipation (a simple example would be an interacting system with a product state as the steady state, initialized in a product state different to the steady state). Our new convergence results seem to indicate that the steady state can be captured accurately even if the variational transient dynamics are visibly distinct from the exact dynamics in some cases only; it appears that the errors in the transient dynamics must still be limited for the steady state to converge. We are unable to comment anymore precisely on this relationship without additional extensive testing that would veer off from the goals of this work, with any insights likely to be model-dependent.

We would like to also point out that time-independent variational minimization approaches can be applied to Lindblad master equation to obtain the steady state while completely bypassing costly transient dynamics [2, 5, 13], albeit at the cost of minimizing highly-non-linear cost functions. In such approaches, the degree of error in the estimate of the steady state is rigorously quantified by the value of the cost function that is strictly bounded by 0 from below, with the limit of 0 attained by an exactly-correct variational solution. While the same cost function can in principle be employed to judge the accuracy of the variational steady state in a TDVP algorithm (which we in fact do in Fig. 62 in the SM), it is not clear to us how this quantity would behave as function of the bond dimension.

References

[1] A. W. Sandvik and G. Vidal. Variational quantum monte carlo simulations with tensor-network states. *Phys. Rev. Lett.*, 99:220602, Nov 2007.

[2] Dawid A. Hryniuk and Marzena H. Szymańska. Tensor-network-based variational Monte Carlo approach to the non-equilibrium steady state of open quantum systems. *Quantum*, 8:1475, September 2024.

[3] Ulrich Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Annals of Physics*, 326(1):96–192, 2011.

[4] Pietro Silvi, Ferdinand Tschirsich, Matthias Gerster, Johannes Jünemann, Daniel Jaschke, Matteo Rizzi, and Simone Montangero. The Tensor Networks Anthology: Simulation techniques for many-body quantum lattice systems. *SciPost Phys. Lect. Notes*, page 8, 2019.

[5] Filippo Vicentini, Alberto Biella, Nicolas Regnault, and Cristiano Ciuti. Variational neural-network ansatz for steady states in open quantum systems. *Physical Review Letters*, 122, 2019.

[6] Alexandra Nagy and Vincenzo Savona. Variational quantum monte carlo method with a neural-network ansatz for open quantum systems. *Physical Review Letters*, 122, 2019.

[7] Michael J. Hartmann and Giuseppe Carleo. Neural-network approach to dissipative quantum many-body dynamics. *Phys. Rev. Lett.*, 122:250502, Jun 2019.

[8] Nobuyuki Yoshioka and Ryusuke Hamazaki. Constructing neural stationary states for open quantum many-body systems. *Phys. Rev. B*, 99:214306, Jun 2019.

[9] Johannes Mellak, Enrico Arrigoni, Thomas Pock, and Wolfgang von der Linden. Quantum transport in open spin chains using neural-network quantum states. *Phys. Rev. B*, 107:205102, May 2023.

[10] Johannes Mellak, Enrico Arrigoni, and Wolfgang von der Linden. Deep neural networks as variational solutions for correlated open quantum systems. *Communications Physics*, 7(1):268, Aug 2024.

[11] Hendrik Weimer, Augustine Kshetrimayum, and Román Orús. Simulation methods for open quantum many-body systems. *Rev. Mod. Phys.*, 93:015008, Mar 2021.

[12] Patrick J. Coles, M. Cerezo, and Lukasz Cincio. Strong bound between trace distance and hilbert-schmidt distance for low-rank states. *Phys. Rev. A*, 100:022103, Aug 2019.

[13] Jian Cui, J. Ignacio Cirac, and Mari Carmen Bañuls. Variational matrix product operators for the steady state of dissipative quantum systems. *Phys. Rev. Lett.*, 114:220601, Jun 2015.