Dear Editor,

We thank the referees for their thorough revision and relevant comments. We hope they find the revised version of the manuscript satisfactory. We have carefully considered their comments and suggestions, and have made corresponding revisions throughout. We believe that both the manuscript and this response now convey the motivation behind our work more clearly, and we hope it will be recognized as a valuable contribution to SciPost Physics. Thank you for your consideration.

Please find attached our point-by-point response to referees' comments and questions. A list of the changes made in the revised version of the manuscript is provided below.

Ljubljana & Garching November 4, 2025
Michele Coppola, Mari Carmen Bañuls, Zala Lenarčič

Referee: 2

Strengths

The paper is clearly written and introduces a highly relevant and original idea to the field.

We thank the referee for the positive comment.

Weaknesses

Certain final results, particularly the prediction of long-time observables, do not achieve very high accuracy.

We acknowledge this criticism and plan to address it in future studies on the topic.

Report

I believe the paper clearly deserves publication in SciPost, because of its fresh, novel ideas as well as its clear presentation. I only provide a few points below for the authors to consider and possibly address in the manuscript.

1. - *You write "a dynamical map is Markovian if and only if it is CP-divisible", while in https://arxiv.org/pdf/1901.05223 the authors claim "We show that completely positive (CP) divisible quantum processes can still involve non-Markovian temporal correlations." Can you comment on that?*

     The referee raises a valid point. However, the apparent contradiction stems from the fact that different definitions of Markovianity coexist in the literature. The paper cited by the referee (Phys. Rev. Lett. 123, 040401) adopts a more stringent and operational definition based on multi-time correlations within the process-tensor framework. In contrast, throughout this work we employ the older and weaker Rivas-Huelga-Plenio (RHP) criterion of Markovianity, which characterises dynamics that are completely positive and divisible. The RHP definition has also been shown to imply the Breuer-Laine-Piilo (BLP) criterion (see Rep. Prog. Phys. 77, 094001 (2014); Rev. Mod. Phys. 88, 021002 (2016)), where non-Markovianity is identified with information backflow. We have now clarified in the manuscript that multiple notions of Markovianity exist, and explicitly state that we adopt the RHP definition throughout.

   - *Moreover, while it may be obvious, why does the non-invertibility of the map imply that it is not divisible?*

     To answer, we need to go back to the definition of divisibility. By definition, the dynamical map $\epsilon_{t,0}$ is divisible iff we can always find a dynamical propagator ϵ_{t_2,t_1} such that

     $$\epsilon_{t_2,0} = \epsilon_{t_2,t_1}\epsilon_{t_1,0}\,, \qquad \forall t_1,\, t_2 \geq 0\,. \tag{1}$$

     Therefore,
     (a) if $\det(\epsilon_{t,0}) \neq 0\ \forall t \geq 0$ then $\epsilon_{t_2,t_1} = \epsilon_{t_2,0}\epsilon_{t_1,0}^{-1}$ is always well defined and the map is divisible;
     (b) if $\exists t_1 \ : \ \det(\epsilon_{t_1,0}) = 0$, divisibility depends on the behaviour of the kernel of

the dynamical map (see footnote in page 5). As shown in Journal of Modern Optics 54(12), 1695 (2007), if the dimension of the kernel of the dynamical map increases monotonically in time, it is possible to find a propagator ϵ_{t_2,t_1} that satisfies Eq. (1), and the map remains divisible; otherwise, it is not. In our work, we explicitly show that the behavior of the kernel is non-trivial (see Sec. 4), and the map alternates between times when it is singular and times when it is invertible. Therefore, we conclude that it is non-divisible.

- *And why negative gamma's imply no CP divisibility?*

  To answer this question, let us consider a divisible dynamical map $\epsilon_{t,0}$. According to the definition of divisibility, we can always find a dynamical propagator ϵ_{t_2,t_1} that satisfies (1).
  Let us assume, for simplicity, $\det(\epsilon_{t,0}) \neq 0 \; \forall t \geq 0$. It is easy to show that $\Lambda_t = \dot{\epsilon}_{t,0}\epsilon_{t,0}^{-1}$ is generator of the dynamics, i.e., $\dot{\hat{\rho}}_t = \Lambda_t[\hat{\rho}_t]$. It is known that Λ_t can be put in the pseudo-Lindblad form (see Eq. (8) of our manuscript), for a time-dependent effective Hamiltonian and time-dependent jump frequencies. There is an important theorem (see Communications in Mathematical Physics, 48, 119-130 or the book by Breuer and Petruccione "The theory of open quantum systems") which states that Λ_t is CP iff $c_{ij}(t)$ is positive semi-definite. Therefore, if $c_{ij}(t)$ has negative eigenvalues (negative gamma's) then Λ_t is not CP and $\epsilon_{t,0}$ is not CP divisible.

2. *"If the dynamical map is invertible, memory-kernel master equations can be rewritten in the time-local form..." To get this form, should the dynamical map be invertible at any time? If the map is not invertible at time t_0, can one still get a local form for time $t < t_0$ and $t > t_0$?*

   Yes, the referee is correct. However, in such a case, one cannot use the time-local form to reconstruct the full dynamical map $\epsilon_{t,0} \; \forall t > 0$; it can only be employed to obtain $\epsilon_{t,0} \; \forall t < t_0$ or the time propagator $\epsilon_{t,t_0} \; \forall t > t_0$.

3. *As far as I understood, the dynamical map is obtained explicitly, i.e. as a matrix, after the transverse TN contraction. If it is the case, why you constraint yourself to use the Frobenius norm, while I think other norms are better indicated for operators / channels? For instance norm-1, or the diamond norm. By the way, while obvious for TN experts, the acronym MPO at page 6 is never defined.*

   Again, the referee raised a valid point. In an ideal situation, one should indeed compare the action of maps on different initial states, e.g., by defining a loss function

$$\overline{\text{loss}}(t, \Lambda_t) \equiv \max_{\hat{\rho}_0} \|\epsilon_{t+\mathrm{d}t,0}[\hat{\rho}_0] - (1 + \Lambda_t \mathrm{d}t)\, \epsilon_{t,0}^{\text{OPT}}[\hat{\rho}_0]\|_1 \,, \tag{2}$$

   where $\|\hat{O}\|_1 = \text{tr}\sqrt{\hat{O}^\dagger \hat{O}}$ is the trace norm of $\hat{O}$ and $\hat{\rho}_0$ is the initial state. The corresponding error function is

$$\overline{\mathcal{E}}(t) \equiv \min_{\Lambda_t} \max_{\hat{\rho}_0} \|\epsilon_{t+\mathrm{d}t,0}[\hat{\rho}_0] - (1 + \Lambda_t \mathrm{d}t)\, \epsilon_{t,0}^{\text{OPT}}[\hat{\rho}_0]\|_1 \,. \tag{3}$$

   While the loss function (2) has the advantage of relying on the distinguishability of quantum states, it is evident that the error function (3), including optimization over (many-body) $\hat{\rho}_0$ initial states and the parameters of Λ_t, is numerically extremely costly.

The diamond norm, physically meaningful for channels, could be used instead. However, even though it can be computed efficiently via an SDP program, in practice it would also largely increase the computational cost of our optimization. Therefore we used an error function that is calculable in practice. The practical aspect behind our choice of Frobenius norm is motivated in the manuscript as well.
Notice that, since we only deal with small subsystems (one or two sites), and the different norms can be related to each other by polynomial factors in the dimension of the subsystem, in our case, all these choices are essentially equivalent. To illustrate it, we can define the loss function

$$\mathrm{loss}_3(\mathrm{t}, \Lambda_\mathrm{t}) \equiv \|\dot{\epsilon}_{\mathrm{t},0} - \Lambda_\mathrm{t}\, \epsilon_{\mathrm{t},0}\|_1 \,, \tag{4}$$

and the error function

$$\mathcal{E}_3(t) \equiv \mathrm{loss}_3(\mathrm{t}, \Lambda_\mathrm{t}^{\mathrm{OPT}}) \,, \tag{5}$$

which is the minimum of the loss function (4) at time t and a fixed number of optimization steps. As illustrated in Figs. 1 and 2 below, we do not find the definition (4) practically advantageous, since it does not yield any significant improvements in the reconstruction of local observables.
We thank the referee for pointing out that we forgot to define the acronym MPO in the manuscript.

4. *If I understand correctly, a separate optimization must be performed for each time t of interest. Is that right?*

   Yes, the referee is right.

5. *How the initial values of the parameters $c(t)$ and $h(t)$ are chosen? Is the learning procedure robust with respect to this initialization? More generally, is the learning procedure robust? For instance, does the optimization reliably converge?*

   As we aim to obtain a Liouvillian that is continuous in time t to better capture the physics of dissipative processes, we used $c(t-dt)$ and $h(t-dt)$ as the initial input values for $c(t)$ and $h(t)$ in the learning procedure. In doing so we reduce the number of required optimization steps. However, as we now more strongly emphasize in the manuscript, we acknowledge the presence of multiple local minima in the parameter space that yield comparably good results.
   For a discussion of the convergence of the optimization procedure, we refer the referee to Appendix A, where we illustrate how the loss function behaves as a function of the ADAM optimization steps. Naturally, when the dynamical map is invertible, we expect that, in the limit of an infinite number of optimization steps, the procedure should converge to a unique Liouvillian and the loss function should approach zero. However, this behavior is not expected at points where the dynamical map becomes singular.

6. *I didn't quite understand whether the optimization process discussed around Eq. (15) is the same as the one described around Eq. (10), and if not, how the two differ.*

   No, it is not. In the loss function (15), $\mathcal{L}_t$ represents a completely positive time-local form by construction, with $c(t) = d(t)^\dagger d(t)$, which is Hermitian and positive semi-definite. This imposes a constraint in our minimization. In contrast, in the loss function (10) we do not enforce such a constraint, and $c(t)$ is only required to be a Hermitian matrix.

7. *Do you have an explanation for why non-Markovianity in Fig.8 is higher in the integrable non-critical case than in chaotic one? Naively, one might expect the opposite.*

   One intuitive way of explaining the sources of non-Markovianity in integrable systems is to consider the dynamics of quasi-particles with different group velocities. For the critical parameters of the integrable model and our initial state, the occupation at momenta with large quasi-particle velocity is large, therefore particles pass through the sub-system quickly and the dynamics is rather Markovian and time-local. Notice that this behavior is the one that closest mimics that of dual unitaries, for which the subsystem dynamics is perfectly Markovian, for the boundary vectors turn out to be unentangled (see e.g. Ref. [58] Bertini et al, SciPost Phys. 8, 067 (2020)).

   For integrable non-critical parameters, there is a wider region of momenta with zero or small quasi-particle velocity. These quasi-particles stay close to the boarder between the sub-system and the rest for a longer time, contributing to the non-Markovian character of the dynamics.

   Chaotic dynamics cannot be addressed in terms of quasi-particles. Thus, the source of Markovianity is different. It has been argued that for generic dynamics and small enough subsystem in comparison to the whole system size, with high probability the evolution of the subsystem is almost Markovian [see Figueroa et al, Quantum 3, 136 (2019)]. Our results suggest that, indeed, our non-integrable system behaves in this way, and seemingly becomes more Markovian after a certain initial time. We added a similar discussion to the Sec. 6.

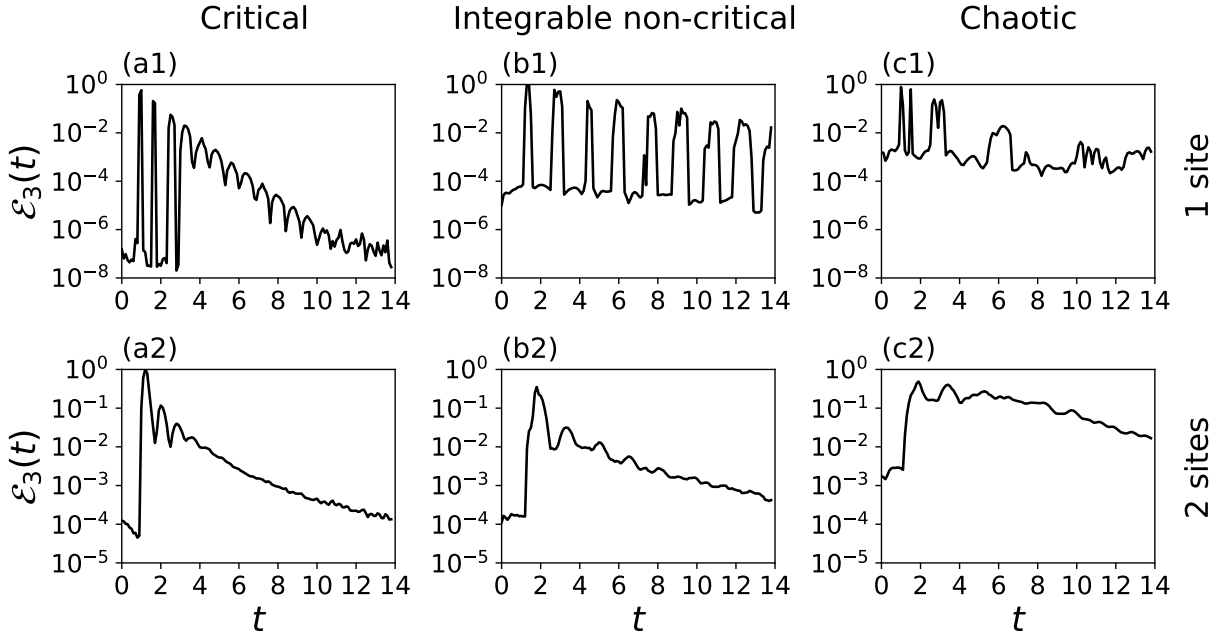

Figure 1: Error $\mathcal{E}_3(t)$ as a function of time t for single-site ($l = 1$) (first row) and two-site $l = 2$ (second row) subsystems, across the three case studies: (a1,a2) critical ($g^x = 1.0$, $g^z = 0.0$), (b1,b2) integrable non-critical ($g^x = 1.5$, $g^z = 0.0$) and (c1,c2) chaotic ($g^x = -1.05$, $g^z = 0.5$) quantum Ising chains. Qualitative behaviour for the error function $\mathcal{E}_3(t)$ using trace norm is comparable to the original error function (Eq. (11) in the manuscript) which makes use of the Frobenius norm (see Fig. 3 in the manuscript for comparison).

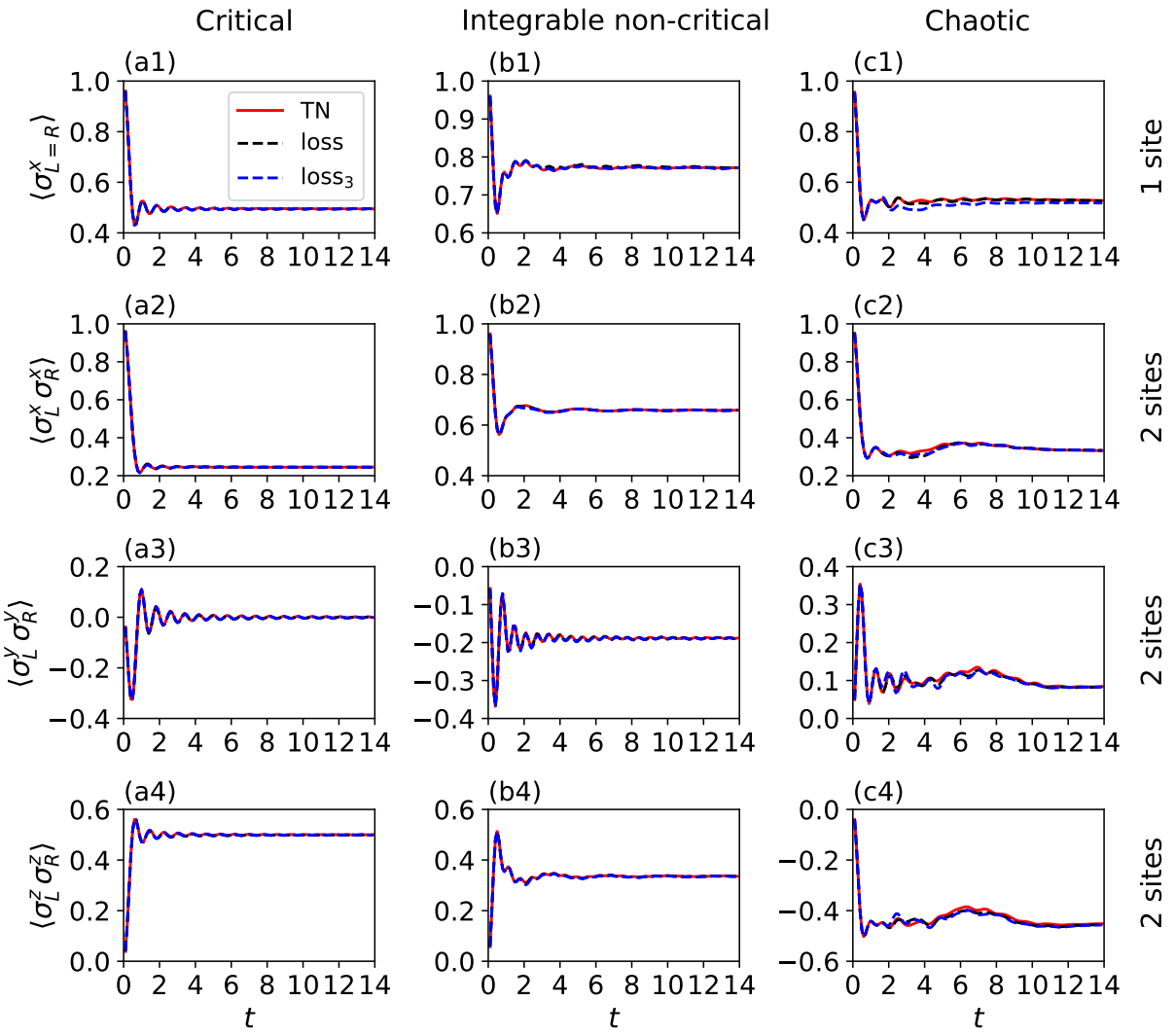

Figure 2: Comparison between the time evolution of observables generated by the dynamical map $\epsilon_{t,0}$ (red curves) and that reconstructed by minimizing the loss functions loss(t) (black dashed lines) and loss$_3(t)$ (blue dashed lines). We present results for single-site $l = 1$ (first row) and two-site $l = 2$ (last three rows) subsystems across the three case studies. First column: critical ($g^x = 1.0$, $g^z = 0.0$) Ising chain. Middle column: integrable non-critical ($g^x = 1.5$, $g^z = 0.0$) Ising chain. Third column: chaotic ($g^x = -1.05$, $g^z = 0.5$) Ising chain. Type of the observable considered is denoted on the left.

Recommendation

Publish (surpasses expectations and criteria for this Journal; among top 10%*).*

We thank the referee for the constructive comments on our manuscript. We trust that our responses and the corresponding revisions adequately address all concerns raised.

List of changes

- In Sec. 3 (page 6), we add "matrix product operator (MPO) ansatz"
- In Sec. 4 (page 8), we specify: "In addition, we would like to emphasize that the choice of the Frobenius norm in (10) is primarily motivated by practical reasons. Although alternative definitions of the loss function based on the trace norm and the distinguishability of quantum states could in principle be employed, they are typically of impractical use because of the high computational cost in minimization. For this reason, the Frobenius norm (10) was chosen for its efficiency."
- In Sec. 5 (page 9), we add: "Since we aim to obtain Liouvillians Λ_t that are as continuous in time as possible, we use the optimized $h_i(t-dt)$ and $c_{ij}(t-dt)$ as input values for the subsequent optimization at time t. However, we should note that a different initialization could result in another solution of a similar quality. We believe this is due to existence of several (local) minima of the loss function, possibly a consequence of singularity of the dynamical map and/or the norm used in the loss function."
- In Sec. 6 (page 14), we specify: "While there are different definitions of Markovianity in the literature, the RHP criterion states that a dynamical map is Markovian if and only if it is CP-divisible."
- In Sec. 6 (page 15), we specify: "On the other hand, for the non-critical integrable model, there are quasi-particles with zero or small velocities that transverse through the subsystem on longer timescales, and in this way contribute to the non-Markovian features of the dynamics. The chaotic model does not allow for a quasi-particle description thus the source of (non-)Markovianity is different there. Our observation with lower-level of non-Markovianity in the chaotic model compared to the gapped integrable case seems to be consistent with Ref. [60], which argues that, for generic dynamics and small enough subsystem, the evolution of the subsystem is almost Markovian with high probability."
- In the conclusions (page 18), we add: "An interesting research direction could be to investigate whether a more Markovian dynamics emerges from initial states at higher energies and, if so, to characterize the thermal nature of the environment by checking the detailed balance condition at the level of the learned Lindblad operators."