

Exponential concentration and symmetries in Quantum Reservoir Computing

Antonio Sannia,^{1,*} Gian Luca Giorgi,¹ and Roberta Zambrini¹

¹*Institute for Cross-Disciplinary Physics and Complex Systems (IFISC) UIB-CSIC, Campus Universitat Illes Balears, 07122, Palma de Mallorca, Spain.*

Quantum reservoir computing (QRC) is an emerging framework for near-term quantum machine learning that offers in-memory processing, platform versatility across analogue and digital systems, and avoids typical trainability challenges such as barren plateaus and local minima. The exponential number of independent features of quantum reservoirs opens the way to a potential performance improvement compared to classical settings. However, this exponential scaling can be hindered by exponential concentration, where finite-ensemble noise in quantum measurements requires exponentially many samples to extract meaningful outputs, a common issue in quantum machine learning. In this work, we go beyond static quantum machine learning tasks and address concentration in QRC for time-series processing using quantum-scrambling reservoirs. Beyond discussing how concentration effects can constrain QRC performance, we demonstrate that leveraging Hamiltonian symmetries significantly suppresses concentration, enabling robust and scalable QRC implementations. We illustrate our approach with concrete examples, including an established QRC design.

- Introduction Reservoir computing (RC) is a machine learning framework designed for time series analysis, originally introduced to address the trainability challenges of recurrent neural networks [1]. The core concept involves leveraging the dynamics of a system to process input signals, with the training phase focusing solely on optimizing a linear combination of measured observables. This approach enables a single reservoir computer to handle diverse tasks without encountering trainability issues. RC has been implemented across various platforms, ranging from digital systems [2, 3] to analog devices [4–6].

Recently, quantum reservoir computing (QRC) has attracted significant attention for its potential to enhance information processing by harnessing the unique properties of quantum systems [7, 8]. Platforms such as digital [9, 10] and analog [11] quantum circuits, bosonic [12–15] and fermionic [16, 17] models, photonic integrated circuits [18], and spin networks [7, 19–22] have been explored as quantum reservoirs. The quantum regime offers two key advantages: the ability to exploit exponentially larger degrees of freedom compared to classical systems and the capability to analyze quantum data without relying on full-state tomography.

While QRC is considered a promising candidate for overcoming the fundamental limitations of other algorithms based on Quantum Neural Networks (QNNs)—such as the exponential cost of training due to barren plateaus [23–27] and local minima [28]—it is not without challenges. Similar to quantum kernel methods [29] and quantum extreme learning machines [30], QRCs can suffer from exponential concentration of output values, inheriting this problem from static to temporal tasks. This means that as the system size increases, each output tends to become exponentially close to a fixed, input-independent value. In realistic scenarios with limited computational resources, this phenomenon can make the whole QRC algorithm inefficient, as the number of required measurements to obtain non-trivial outputs

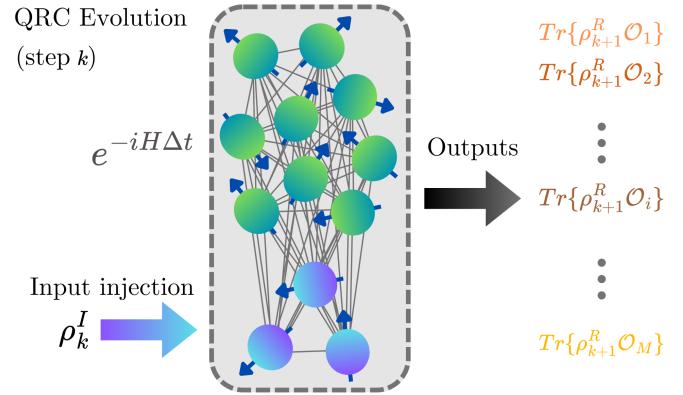


FIG. 1. Three-layer design of QRC governed by Eq. 1: Input injection in state ρ_k^I ; reservoir evolution; output measurements (colors indicate that they are distinguishable in the ideal case without statistical errors).

grows exponentially with system scaling.

In the ideal case, where an infinite number of measurements is available, the learning ability of QNNs has been linked to their capacity for scrambling information [31, 32]. However, training QNNs, which function as scramblers, is resource-inefficient [33, 34]. Similarly, QRC tends to achieve peak performance when reservoirs operate as information scramblers [10, 19, 35], even in the time-independent case of extreme learning machines [36]. Beyond ideal conditions, when finite-ensemble noise in quantum measurements is accounted for, we show that exponential concentration effects emerge, limiting achievable efficiency in terms of resources. On a positive note, we identify a strategy to sustain QRC scalability and show that concentration can be limited by incorporating symmetries [37, 38] into the reservoir model, the presence of which mitigates exponential concentration effects and preserves the distinguishability of output observables.

Quantum reservoir model – We consider spin-based

QRC models belonging to the class of erase-and-write maps, introduced in Ref. [7]. These systems have been extensively studied and have shown promising results in processing both classical and quantum data series [7, 20]. The QRC algorithm can be summarized in three main steps: (i) input injection; (ii) dynamical evolution of the reservoir; (iii) measurement of reservoir observables to construct the output layer [39].

Consider a reservoir of n qubits. At each step k of the QRC algorithm, an input of $m \leq n$ -qubit input states is injected (i), resetting a subset of the reservoir qubits. The entire reservoir then evolves according to its Hamiltonian, coupling the $n+m$ qubits. The reservoir updating rule (ii) is expressed as:

$$\rho_{k+1}^R = e^{-iH\Delta t} \rho_k^I \otimes \text{Tr}_m\{\rho_k^R\} e^{iH\Delta t}, \quad (1)$$

where Tr_m denotes the partial trace over the m qubits into which the inputs are injected, ρ_k^R is the reservoir state at time step k , H is the reservoir Hamiltonian, and Δt is the time between inputs (See Fig. 1 for a summary scheme). Notably, the input states ρ_k^I may either encode classical data or arise from prior quantum processes, making the procedure suitable for both quantum and classical tasks [8].

After each input injection, the reservoir outputs are derived from a set of M observables $\{\mathcal{O}_i\}_{i=1}^M$. The output at time step k is given by a linear combination:

$$y_k = \sum_{i=1}^M w_i \text{Tr}\{\rho_k \mathcal{O}_i\} = \sum_{i=1}^M w_i \langle \mathcal{O}_i \rangle_k, \quad (2)$$

where the w_i values are the free weights optimized during training, and $\langle \mathcal{O}_i \rangle_k = \text{Tr}\{\rho_k \mathcal{O}_i\}$. Minimal optimization resources are used to complete this final step (iii), typically through linear regression.

The state space and the number of distinct observables scale exponentially with the full system size in a quantum systems. In the ideal case, with unlimited resources, these expectation values can be computed exactly. However, in practice, only a finite number of measurements is available. This introduces unavoidable statistical errors that fundamentally limit the precision of reservoir output extraction [14, 21, 40]. Moreover, for the expectation values in Eq. (2) to be usable by the QRC algorithm, they must depend solely on the input series. This requirement necessitates that the reservoir dynamics satisfy the so-called *echo state property*, which means that the reservoir state forgets its initial condition over time, making the output layer a well-defined function of the input series [41].

Not all dynamics described by Eq. (1) satisfy the echo state property, and specific classes of reservoir Hamiltonians H must be chosen. In particular, it has been found that the unitary evolution of the reservoir must scramble input-state information across all the system qubits to capture the desired feature [19, 35].

Theory framework – Although the map in Eq. (1) involves both unitary evolution and nonunitary state reset, analyzing its unitary part alone suffices to address the concentration problem, as this already emerges in each of the time steps. Assuming a chaotic reservoir Hamiltonian H , its unitary evolution $U = e^{-iH\Delta t}$ generates information scrambling when the evolution time Δt exceeds a characteristic timescale. In this regime, U can be regarded as a random unitary sampled from a 2-design distribution, satisfying what is known as ergodicity in classical statistical mechanics [42]. This implies that, in the absence of Hamiltonian constraints, the state of the reservoir ρ_k^R will converge exponentially quickly (in the number of qubits n) toward the maximally mixed state, which renders it ineffective for information extraction. In this work, we propose to counteract this effect by identifying the proper class of scrambling reservoirs for QRC. We will consider block-diagonal Hamiltonians arising from the presence of symmetries in the system, which partition the Hamiltonian into invariant subspaces. This constrains the distribution from which U is sampled, and, consequently, the resulting reservoir state distribution. The absence of symmetry can be seen as a special instance where the only conserved quantity, during the unitary step, is the energy. Our first result, proven in the Supplementary Material [43], characterizes these constrained reservoir states:

Lemma 1: *Let H be the reservoir Hamiltonian, and let $\{S_i\}_{i=1}^K$ denote the set of its corresponding irreducible symmetries, such that $[S_i, H] = 0$ and $[S_i, S_j] = 0$ for all i, j . If the reservoir unitary evolution $U = e^{-iH\Delta t}$ generates scrambling dynamics, then the reservoir state, at each time step k , can be described as being sampled from a probability distribution whose mean $\bar{\rho}_k^R$ is given by the following direct sum:*

$$\bar{\rho}_k^R = \bigoplus_{l=1}^L \alpha_l^k \frac{\mathbb{P}_l}{D_l}. \quad (3)$$

Here, L is the number of irreducible symmetry sectors (joint eigenspaces of S_i), which we will call symmetric subspaces. If we denote by d_i the number of distinct eigenvalues of S_i , we have $L = \prod_{i=1}^K d_i$. The operators \mathbb{P}_l are projectors onto these subspaces, and D_l represents their corresponding dimension. Finally, the coefficients α_l^k depend on the particular input history in the QRC dynamical map (1), as shown in the supplementary material [43].

By incorporating a variance analysis of the reservoir state distribution, in alignment with results found in other quantum machine learning models [29, 30], we derive a condition that implies an exponential concentration for output observables. While previous analyses attribute this concentration to the exponential size of the entire Hilbert space, we anticipate that, in the presence of symmetries, it is a consequence of the possible expo-

nential size of the symmetric subspaces.

Theorem 1 (Exponential concentration): *Let \mathcal{O}_i be an output observable. If, for all the symmetric subspaces, labelled by l , the following conditions hold:*

$$\begin{aligned}\text{Tr}\{\mathbb{P}_l \mathcal{O}_i\} &= O(1); \\ D_l &= O(e^{cn})\end{aligned}$$

where c is a positive constant, then we have that

$$\text{Tr}\{\mathcal{O}_i \rho_k^R\} = O(e^{-cn}),$$

with probability exponentially close, with respect to n , to 1. In the absence of symmetries, the same result holds, with D_l dimension of the whole Hilbert space.

Before entering into the importance of this theorem, we would like to add a result that follows from its demonstration in the Supplementary Material [43] and is valid independently of the D_l scaling. The expectation value $\text{Tr}\{\mathcal{O}_i \rho_k^R\}$ is a sample from a probability distribution whose mean is the expectation value \mathcal{O}_i computed with respect to the state of Eq. (3) and the variance can still be evaluated asymptotically being at most a $O(1/\min_{D_l \neq 1} D_l)$, also when not all the D_l values scale exponentially. These considerations can be useful in addressing cases not affected by exponential concentration (see, for instance, the Ising model example below).

Under the hypothesis of Theorem 1, the presence of exponential concentration implies that a number of measurements that exponentially scale with the number of qubits is required to properly resolve the observable expectation values, rendering the reservoir computing algorithm inefficient. Nevertheless, Theorem 1 also provides clues for the strategies that allow one to circumvent concentration by violating its hypotheses. A first strategy could consist of considering reservoir Hamiltonians with a number of symmetries that scale with the size, ensuring that all the dimensions D_l are at most $O(\text{poly}(n))$. In this case, the number of resources required to construct the output layer scales at most polynomially with the reservoir size, ensuring the overall efficiency of the algorithm.

As a second strategy, exponential concentration can be avoided by selecting observables that are compatible with the state structure of Eq. (3). More formally, we state

Theorem 2 (Suitable observables): *Given a reservoir dynamics that fulfill Lemma 1 hypothesis, let \mathcal{O}_i be an output observable of the form:*

$$\mathcal{O}_i = \bigoplus_l \beta_l \mathbb{P}_l, \quad (4)$$

which aligns with Eq. (3). If the amplitudes α_l^k and β_l are independent of the system size, then $\text{Tr}\{\mathcal{O}_i \rho_k^R\}$ can be considered as sampled from a probability distribution whose mean is $\sum_l \alpha_l^k \beta_l = O(1)$, avoiding an exponential concentration.

Theorem 2, proven in the Supplementary Material [43], suggests that a natural choice of observables that avoids exponential concentration is the set of symmetry operators S_i themselves. Furthermore, we observe that the number of degrees of freedom that can be efficiently extracted (related to the output layer) is determined by the number of functions α_l^k , which is equivalent to the number of symmetric subspaces, L , in Eq. (3).

Another important implication of Theorem 2 is that it establishes a necessary condition for the effective operation of the reservoir computing model. Specifically, a dynamical reservoir response can only be achieved if the α_l^k coefficients vary at each step. This variation implies that the injection of input states, as described in Eq. (1), must change the expectation values of the symmetries with respect to the reservoir state.

As in the examples we show below, the echo state property is directly satisfied when input injections constrain the coefficients α_l^k to be fading memory functions, which means that they only depend on the recent input history [39]. In this case, with good approximation,

$$\alpha_l^k \simeq \alpha_l^k(\rho_{k-1}^I, \dots, \rho_{k-\tau_{FM}}^I),$$

where τ_{FM} is the fading memory time. It is now immediate to conclude that $\text{Tr}\{\mathcal{O}_i \rho_k\}$ is a sample from a distribution that only depends on the input series. This distribution is independent of the initial state, thereby satisfying the echo state property requirement.

Importantly, Theorems 1 and 2 imply that for a scrambling reservoir, as in the case of Eq. (1), symmetries are necessary to prevent exponential concentration. Formally, we state:

Corollary 1 (Necessity of symmetries): *For a reservoir evolving according to a scrambling unitary, an efficient protocol to extract input-dependent dynamics, avoiding exponential concentration, must exhibit symmetries.*

In fact, in the absence of symmetries, the only way of avoiding an exponential concentration, violating Theorem 1 hypothesis, is to consider an observable of the form of Eq. (4), as dictated by Theorem 2. Then, only observables proportional to the identity satisfy this condition, yet their expectation values are obviously input-independent and useless for QRC. Consequently, symmetries become a necessary requirement for scalable reservoir computing, as formalized in Corollary 1.

Analytical QRC example avoiding concentration— To illustrate the theoretical framework developed above, we now present a tailored example where we can construct analytically the proper observables for a simple task of series discrimination:

Definition 1 (Series discrimination task): *Consider an n -qubit reservoir described by Eq. (1), initialized in a random state. We inject one of two possible time series of single-qubit states into the reservoir: $\{|0\rangle\langle 0|\}_k$ or*

$\{|1\rangle\langle 1|\}_k$. The goal of the task is to distinguish between these two cases without directly measuring the input-encoded qubit.

Considering the reservoir's unitary evolution as a scrambler, efficiently solving the task requires that the scrambler produces a response independent of the initial condition, thereby satisfying the echo state property, and transmits the input information across the reservoir qubits, preventing exponential concentration. As discussed, this condition can only be achieved in the presence of symmetries, and we will now identify a symmetry operator suited for this particular problem. More formally, we state:

Proposition 1 (Efficient task resolution): *For the discrimination task in Definition 1, when the reservoir implements unitary scrambling dynamics, a model that is symmetric under the operator $S = \sum_{i=1}^n \sigma_i^z$, can solve the task efficiently, both in required time steps and resources needed for output extraction.*

A proof of Proposition 1 is provided in the Supplementary Material [43], where the reservoir evolution is analytically computed. Furthermore, for a reservoir model to solve a task efficiently, the mere presence of symmetries is insufficient, as they must be properly matched to the problem structure. To illustrate this, if we repeat the calculations from the proof of Proposition 1, using a reservoir model symmetric under $S' = \sum_{i=1}^n \sigma_i^x$, we find that the reservoir will converge to the same final state for both input series, failing to discriminate between them.

QRC with Ising reservoir – The developed framework establishes the limitations to scalable QRC due to concentration and, more importantly, strategies to overcome them. Beyond general considerations and synthetic models, we now apply our theory to a widely used quantum reservoir model, the first proposed in the literature. Specifically, we consider the known analogue reservoir with Hamiltonian given by a fully transverse-field Ising model [7, 19–21, 35]:

$$H = \sum_{i>j=1}^n J_{ij} \sigma_i^x \sigma_j^x + \frac{1}{2} \sum_{i=1}^n (h + h_i) \sigma_i^z, \quad (5)$$

where σ_i^a ($a = x, y, z$) are the Pauli matrices, J_{ij} are the coupling strengths uniformly sampled from the set $[-J_s/2, J_s/2]$, h is a uniform magnetic field, and h_i is an on-site disorder randomly sampled from the interval $[-W, W]$. In the following, all the hyperparameters will be expressed in the J_s units. The inputs of Eq. (1) are single-qubit states parameterized by a classical real parameter:

$$\rho_k^I = |\psi_{s_k}\rangle\langle\psi_{s_k}|,$$

where $|\psi_{s_k}\rangle = \sqrt{s_k}|0\rangle + \sqrt{1-s_k}|1\rangle$ with s_k a real number belonging to range $[0, 1]$.

It has been shown that the model satisfies the echo state property, which is a necessary condition for a func-

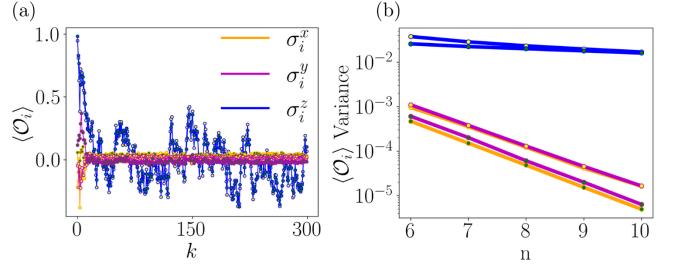


FIG. 2. (a) Expectation values evolution of single-qubit Pauli observables, responding to a random input sequence for a 7-qubit realization. (b) Average variances of the expectation values series, over 1000 random realizations of the Hamiltonian and initial reservoir states, varying the number of qubits, depicted in a logarithmic scale. 500 random inputs have been used to wash out the dependence on the initial conditions, while with the following 500, the variances have been computed. In the two plots, we represented the qubit on which the input is injected and a different qubit in the network, respectively with yellow and green points.

tioning reservoir computer, only within specific Hamiltonian regimes. The optimal operating conditions emerge in the ergodic phase, where the Hamiltonian's chaotic dynamics naturally induce scrambling behavior [19]. Consequently, the hyperparameters W and h must be carefully selected to satisfy this condition. As recalled in the Supplementary Material [43], by fixing the hyperparameters $W=10^{-2}$ and $h=10^1$ we fall in this scenario, and, from now on, we maintain these values.

Consistent with our analytical results, only observables immune to exponential concentration are suitable for the construction of the output layer. This requires identifying the possible symmetries of the Hamiltonian H of Eq. (5). In this particular case, the only exact symmetry is the parity $P = \prod_{i=1}^n \sigma_i^z$, which, despite being immune to scrambling effects, tends to vanish due to its global action on the qubit set [25, 30].

However, in the deep thermal phase ($h \gg J_s$), the Hamiltonian is well-approximated by its leading-order term $H \simeq h/2 \sum_{i=1}^n \sigma_i^z$. This dominant component identifies the operators that effectively approximate dynamical symmetries that can help identify a set of suitable observables, while the subleading contributions ensure the unitary evolution generates scrambling dynamics. Under this approximation, the dimension of the symmetric subspaces is $D_l = \binom{n}{l}$. Let us first consider $l = c$ or $l = n - c$, with c a constant independent of n . Although these subspaces are immune to exponential concentration as their size grows polynomially with n , they display negligible populations (for a random input injection) and are not useful for computational purposes. Out of these negligible subspaces, concentration is expected to emerge from our theoretical analysis in all the remaining symmetry sectors due to their exponential dimension growth. In this case, Theorem 2 helps to find the observables that

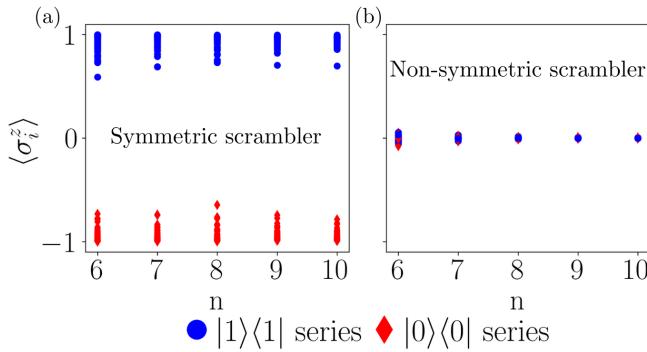


FIG. 3. Numerical verification of Theorem 3. Reservoir response as a function of the number of qubits, n , for two cases: (a) Symmetric scrambler, generated by the Ising Hamiltonian in its thermal phase; (b) Asymmetric scrambler, obtained by sampling the reservoir unitary evolution from a Haar distribution. In both cases, 1000 input sequences of either $|1\rangle\langle 1|$ or $|0\rangle\langle 0|$ states were injected into random initial reservoir states. For a fixed n , 100 different realizations are shown for each input series. Each point represents the expectation value $\langle \sigma_i^z \rangle$, computed for a randomly chosen qubit distinct from the one receiving the input.

escape concentration, as for instance σ_i^z . The results presented in Fig. 2 (a) confirm these considerations. In fact, only the expectation values of σ_i^z , which approximate the symmetries of the Hamiltonian, exhibit clearly distinguishable non-zero values, while σ_i^x and σ_i^y are very close to zero.

To quantify exponential concentration as a function of n , we computed the variance of the expectation value series. The averaged variances are depicted in Fig. 2 (b). As expected, for the σ_i^z observables, the variances persist at a system-size-independent order of magnitude, whereas for the σ_i^x and σ_i^y cases, they are exponentially suppressed as the system size increases. It implies that an output layer composed of single-qubit observables can be efficiently computed by considering only the σ_i^z matrices. Still, the expectation values of different σ_i^z tend to converge towards the same value. This is due to the permutation symmetry of the approximated Hamiltonian, and to the fact that Eq. (3) is a good approximation of the reservoir state.

Interestingly, this approximated H naturally displays the symmetry of Proposition 1 and can therefore realize the series discrimination task of Definition 1. In Fig. 3, we numerically calculate the expectation values of σ_i^z , where i refers to a randomly chosen qubit distinct from the one where the input was applied. As predicted by our theoretical results, the Ising model of Eq. (5) (symmetric scrambler) can discriminate between the two cases, while this is not possible for a unitary that is randomly sampled from a Haar distribution (non-symmetric scrambler) that tends to concentrate.

Discussion and conclusion We have demonstrated that

exponential concentration effects pose a fundamental challenge to quantum reservoir computing, particularly when using scrambling dynamics. While scramblers efficiently propagate information across quantum systems, we demonstrate that this process exponentially suppresses the distinguishability of output observables. Interestingly, our results link the issues already known about learning scrambling dynamics to the possibility of using them as machine learning models [44].

Crucially, we have shown that incorporating symmetries into the reservoir Hamiltonian prevents this exponential concentration, enabling scalable and efficient QRC. Our analytical results establish a direct connection between symmetries and the set of observables that remain distinguishable, while numerical simulations validate these predictions.

While our analysis focuses on spin-based erase-and-write dynamics, the symmetry-enhanced framework is extensible to feedback-controlled systems [10], real-time learning protocols [21, 45], hybrid classical-quantum architectures [18, 46–48], and quantum extreme learning paradigms [8].

Our work establishes symmetry engineering as an indispensable paradigm for mitigating quantum reservoir computing's inherent scalability limits. This opens the possibility of future research where optimal symmetry structures are systematically tailored for specific time series tasks of interest.

Finally, we remark that, in addition to the resource requirements for extracting the output signal, another fundamental aspect of QRC is its memory behavior. In Ref. [49], we demonstrated that avoiding an exponential memory decay, over time, necessitates the use of non-Markovian maps. Consequently, we highlight that symmetric non-Markovian reservoirs can exhibit long-range temporal correlations with the inputs, also enabling an efficient information extraction.

During the preparation of the manuscript, we became aware of a related independent work [50] that also investigates exponential concentration in quantum reservoir computing, but with a different focus. In our work, we propose symmetries as a strategy to solve exponential concentration effects that originate from the chaotic nature of the reservoir evolution. In contrast, in Ref. [50], exponential concentration is studied as a function of the iterative application of the reservoir computing algorithm.

ACKNOWLEDGMENTS

We acknowledge the Spanish State Research Agency, through the María de Maeztu project CEX2021-001164-M funded by the MCIU/AEI/10.13039/501100011033, through the COQUSY project PID2022-140506NB-C21 and -

C22 funded by MCIU/AEI/10.13039/501100011033, MINECO through the QUANTUM SPAIN project, and EU through the RTRP - NextGenerationEU within the framework of the Digital Spain 2025 Agenda. The CSIC Interdisciplinary Thematic Platform (PTI+) on Quantum Technologies in Spain (QTEP+) is also acknowledged. The project that gave rise to these results received the support of a fellowship from the “la Caixa” Foundation (ID 100010434). The fellowship code is LCF/BQ/DI23/11990081.

* sannia@ifisc.uib-csic.es

- [1] K. Nakajima and I. Fischer, *Reservoir Computing* (Springer, 2021).
- [2] H. Jaeger and H. Haas, Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication, *Science* **304**, 78–80 (2004).
- [3] W. Maass and H. Markram, On the computational power of circuits of spiking neurons, *Journal of Computer and System Sciences* **69**, 593–616 (2004).
- [4] G. Van der Sande, D. Brunner, and M. C. Soriano, Advances in photonic reservoir computing, *Nanophotonics* **6**, 561–576 (2017).
- [5] G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose, Recent advances in physical reservoir computing: A review, *Neural Networks* **115**, 100–123 (2019).
- [6] K. Nakajima, Physical reservoir computing—an introductory perspective, *Japanese Journal of Applied Physics* **59**, 060501 (2020).
- [7] K. Fujii and K. Nakajima, Harnessing disordered ensemble quantum dynamics for machine learning, *Physical Review Applied* **8** (2017).
- [8] P. Mujal, R. Martínez-Peña, J. Nokkala, J. García-Beni, G. L. Giorgi, M. C. Soriano, and R. Zambrini, Opportunities in quantum reservoir computing and extreme learning machines, *Advanced Quantum Technologies* **4** (2021).
- [9] J. Chen, H. I. Nurdin, and N. Yamamoto, Temporal information processing on noisy quantum computers, *Physical Review Applied* **14** (2020).
- [10] K. Kobayashi, K. Fujii, and N. Yamamoto, Feedback-driven quantum reservoir computing for time-series analysis, *PRX Quantum* **5** (2024).
- [11] A. Senanian, S. Prabhu, V. Kremenetski, S. Roy, Y. Cao, J. Kline, T. Onodera, L. G. Wright, X. Wu, V. Fatemi, and P. L. McMahon, Microwave signal processing using an analog quantum reservoir computer, *Nature Communications* **15** (2024).
- [12] J. Nokkala, R. Martínez-Peña, G. L. Giorgi, V. Parigi, M. C. Soriano, and R. Zambrini, Gaussian states of continuous-variable quantum systems provide universal and versatile reservoir computing, *Communications Physics* **4** (2021).
- [13] L. C. G. Govia, G. J. Ribeill, G. E. Rowlands, H. K. Krovi, and T. A. Ohki, Quantum reservoir computing with a single nonlinear oscillator, *Physical Review Research* **3** (2021).
- [14] J. García-Beni, G. L. Giorgi, M. C. Soriano, and R. Zambrini, Scalable photonic platform for real-time quantum reservoir computing, *Physical Review Applied* **20** (2023).
- [15] A. Sannia, G. L. Giorgi, S. Longhi, and R. Zambrini, Liouvillian skin effect in quantum neural networks, *Optica Quantum* **3**, 189 (2025).
- [16] S. Ghosh, A. Opala, M. Matuszewski, T. Paterek, and T. C. H. Liew, Quantum reservoir processing, *npj Quantum Information* **5** (2019).
- [17] G. Llodrà, C. Charalambous, G. L. Giorgi, and R. Zambrini, Benchmarking the role of particle statistics in quantum reservoir computing, *Advanced Quantum Technologies* **6** (2022).
- [18] M. Spagnolo, J. Morris, S. Piacentini, M. Antesberger, F. Massa, A. Crespi, F. Ceccarelli, R. Osellame, and P. Walther, Experimental photonic quantum memristor, *Nature Photonics* **16**, 318–323 (2022).
- [19] R. Martínez-Peña, G. L. Giorgi, J. Nokkala, M. C. Soriano, and R. Zambrini, Dynamical phase transitions in quantum reservoir computing, *Physical Review Letters* **127** (2021).
- [20] Q. H. Tran and K. Nakajima, Learning temporal quantum tomography, *Physical Review Letters* **127** (2021).
- [21] P. Mujal, R. Martínez-Peña, G. L. Giorgi, M. C. Soriano, and R. Zambrini, Time-series quantum reservoir computing with weak and projective measurements, *npj Quantum Information* **9** (2023).
- [22] A. Sannia, R. Martínez-Peña, M. C. Soriano, G. L. Giorgi, and R. Zambrini, Dissipation as a resource for quantum reservoir computing, *Quantum* **8**, 1291 (2024).
- [23] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, *Nature Communications* **9** (2018).
- [24] M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo, A review of barren plateaus in variational quantum computing (2024), arXiv:2405.00781 [quant-ph].
- [25] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, *Nature Communications* **12** (2021).
- [26] A. Sannia, F. Tacchino, I. Tavernelli, G. L. Giorgi, and R. Zambrini, Engineered dissipation to mitigate barren plateaus, *npj Quantum Information* **10** (2024).
- [27] S. Thanasilp, S. Wang, N. A. Nghiem, P. Coles, and M. Cerezo, Subtleties in the trainability of quantum machine learning models, *Quantum Machine Intelligence* **5** (2023).
- [28] E. R. Anschuetz and B. T. Kiani, Quantum variational algorithms are swamped with traps, *Nature Communications* **13** (2022).
- [29] S. Thanasilp, S. Wang, M. Cerezo, and Z. Holmes, Exponential concentration in quantum kernel methods, *Nature Communications* **15** (2024).
- [30] W. Xiong, G. Facelli, M. Sahebi, O. Agnel, T. Chotibut, S. Thanasilp, and Z. Holmes, On fundamental aspects of quantum extreme learning machines (2024), arXiv:2312.15124 [quant-ph].
- [31] H. Shen, P. Zhang, Y.-Z. You, and H. Zhai, Information scrambling in quantum neural networks, *Physical Review Letters* **124** (2020).
- [32] Y. Wu, P. Zhang, and H. Zhai, Scrambling ability of quantum neural network architectures, *Physical Review Research* **3** (2021).

- [33] C. Ortiz Marrero, M. Kieferová, and N. Wiebe, Entanglement-induced barren plateaus, *PRX Quantum* **2** (2021).
- [34] R. J. Garcia, K. Bu, and A. Jaffe, Quantifying scrambling in quantum neural networks, *Journal of High Energy Physics* **2022** (2022).
- [35] W. Xia, J. Zou, X. Qiu, and X. Li, The reservoir learning power across quantum many-body localization transition, *Frontiers of Physics* **17** (2022).
- [36] M. Vetrano, G. Lo Monaco, L. Innocenti, S. Lorenzo, and G. M. Palma, State estimation with quantum extreme learning machines beyond the scrambling time, *npj Quantum Information* **11** (2025).
- [37] J. J. Meyer, M. Mularski, E. Gil-Fuster, A. A. Mele, F. Arzani, A. Wilms, and J. Eisert, Exploiting symmetry in variational quantum machine learning, *PRX Quantum* **4** (2023).
- [38] M. Larocca, F. Sauvage, F. M. Sbahi, G. Verdon, P. J. Coles, and M. Cerezo, Group-invariant quantum machine learning, *PRX Quantum* **3** (2022).
- [39] M. Cucchi, S. Abreu, G. Ciccone, D. Brunner, and H. Kleemann, Hands-on reservoir computing: a tutorial for practical implementation, *Neuromorphic Computing and Engineering* **2**, 032002 (2022).
- [40] F. Hu, G. Angelatos, S. A. Khan, M. Vives, E. Türeci, L. Bello, G. E. Rowlands, G. J. Ribeill, and H. E. Türeci, Tackling sampling noise in physical systems for machine learning applications: Fundamental limits and eigen-tasks, *Physical Review X* **13** (2023).
- [41] I. B. Yıldız, H. Jaeger, and S. J. Kiebel, Re-visiting the echo state property, *Neural Networks* **35**, 1 (2012).
- [42] D. A. Roberts and B. Yoshida, Chaos and complexity by design, *Journal of High Energy Physics* **2017** (2017).
- [43] For further details, see Supplemental Material.
- [44] Z. Holmes, A. Arrasmith, B. Yan, P. J. Coles, A. Albrecht, and A. T. Sornborger, Barren plateaus preclude learning scramblers, *Physical Review Letters* **126** (2021).
- [45] G. Franceschetto, M. Płodzień, M. Lewenstein, A. Acín, and P. Mujal, Harnessing quantum back-action for time-series processing (2024), arXiv:2411.03979 [quant-ph].
- [46] J. Nokkala, G. L. Giorgi, and R. Zambrini, Retrieving past quantum features with deep hybrid classical-quantum reservoir computing, *Machine Learning: Science and Technology* **5**, 035022 (2024).
- [47] F. Wudarski, D. O'Connor, S. Geaney, A. A. Asanjan, M. Wilson, E. Strbac, P. A. Lott, and D. Venturelli, Hybrid quantum-classical reservoir computing for simulating chaotic systems (2024), arXiv:2311.14105 [quant-ph].
- [48] J. Settino, L. Salatino, L. Mariani, M. Channab, L. Bozzolo, S. Vallisa, P. Barillà, A. Policicchio, N. L. Gullo, A. Giordano, C. Mastrianni, and F. Plastina, Memory-augmented hybrid quantum reservoir computing (2024), arXiv:2409.09886 [quant-ph].
- [49] A. Sannia, R. R. Rodríguez, G. L. Giorgi, and R. Zambrini, Non-markovianity and memory enhancement in quantum reservoir computing (2025), arXiv:2505.02491 [quant-ph].
- [50] W. Xiong, Z. Holmes, A. Angrisani, Y. Suzuki, T. Chotibut, and S. Thanaisilp, Role of scrambling and noise in temporal information processing with quantum systems (2025), arXiv:2505.10080 [quant-ph].
- [51] Z. Puchala and J. Miszczak, Symbolic integration with respect to the haar measure on the unitary groups, *Bulletin of the Polish Academy of Sciences Technical Sciences* **65**, 21–27 (2017).
- [52] A. A. Mele, Introduction to Haar Measure Tools in Quantum Information: A Beginner’s Tutorial, *Quantum* **8**, 1340 (2024).

SUPPLEMENTARY MATERIAL FOR "EXPONENTIAL CONCENTRATION AND SYMMETRIES IN QUANTUM RESERVOIR COMPUTING"

PROOFS OF THE ANALYTICAL RESULTS PRESENTED IN THE MAIN TEXT

Proof of Lemma 1

Proof:

Given the commuting irreducible symmetry operators $\{S_i\}_{i=1}^K$ satisfying $[S_i, H] = 0$ and $[S_i, S_j] = 0$, we can define a basis of simultaneous eigenvectors, denoted as $|\psi_l^s\rangle$. Here, the index l , running up to L , identifies the joint eigenvalue sector of all S_i (therefore, this index identifies each set of common eigenvectors of all symmetries). The superindex s instead identifies states within the l -th degenerate subspace (whose dimension is D_l). On this basis, the Hamiltonian H can be expressed as a direct sum:

$$H = \bigoplus_{l=1}^L H_l,$$

where each H_l is a Hermitian operator that can be written as a linear combination of projectors of the form $|\psi_l^s\rangle\langle\psi_l^s|$.

Consequently, the reservoir unitary evolution becomes:

$$U = e^{-iH\Delta t} = \prod_{l=1}^L e^{-iH_l\Delta t} = \prod_{l=1}^L U_l,$$

where each U_l is sampled from a 2-design distribution, consistent with the scrambling hypothesis [42].

Under this condition, using Haar integral identities, the distribution of U_l satisfies the following superoperator equality [51, 52]:

$$\int d\mu(U_l) U_l(\cdot) U_l^\dagger = \text{Tr}\{\mathbb{P}_l \cdot\} \frac{\mathbb{P}_l}{D_l} + \prod_{l'=1, l' \neq l}^L \mathbb{P}_{l'}(\cdot) \prod_{l'=1, l' \neq l}^L \mathbb{P}_{l'},$$

where with $d\mu(\cdot)$ we indicate a distribution volume element and \mathbb{P}_l is the projector operator onto the eigenspace of H_l , defined as: $\mathbb{P}_l = \sum_{s=1}^{D_l} |\psi_l^s\rangle\langle\psi_l^s|$.

Using this identity, we can compute the average reservoir state after the unitary evolution:

$$\int d\mu(U) U (\rho_k^I \otimes \text{Tr}_m\{\rho_k^R\}) U^\dagger = \int \prod_l d\mu(U_l) U_L \cdots U_1 \rho_k^I \otimes \text{Tr}_m\{\rho_k^R\} U_1^\dagger \cdots U_L^\dagger = \bigoplus_{l=1}^L \alpha_l^k \frac{\mathbb{P}_l}{D_l},$$

where

$$\alpha_l^k = \text{Tr}\{\mathbb{P}_l \rho_k^I \otimes \text{Tr}_m\{\rho_k^R\}\}.$$

■

Proof of Theorem 1

Proof:

From Eq. (3) of Lemma 1, we find that $\text{Tr}\{\mathcal{O}_i \rho_k^R\}$ is a value sampled from a probability distribution whose mean is:

$$\text{Tr}\left\{\mathcal{O}_i \left(\bigoplus_{l=1}^L \alpha_l^k \frac{\mathbb{P}_l}{D_l} \right) \right\} = \sum_{l=1}^L \alpha_l^k \frac{\text{Tr}\{\mathcal{O}_i \mathbb{P}_l\}}{D_l},$$

which is a $O(e^{-cn})$ under the Theorem hypothesis.

To conclude the proof of the theorem, it is now necessary to determine the variance of the distribution. Therefore, we need to evaluate the following quantity:

$$\text{Var}_U[\text{Tr}\{\mathcal{O}_i \rho_k^R\}] = \int d\mu(U) \text{Tr}\{\mathcal{O}_i \rho_k^R\}^2 - \left(\int d\mu(U) \text{Tr}\{\mathcal{O}_i \rho_k^R\} \right)^2.$$

As we already know that the second term in the sum is an $O(e^{-2cn})$, we proceed to evaluate the first one:

$$\int d\mu(U) \text{Tr}\{\mathcal{O}_i \rho_k^R\}^2 = \int d\mu(U) \text{Tr}\{\mathcal{O}_i^{\otimes 2} \rho_k^{R \otimes 2}\} = \int \prod_l d\mu(U_l) \text{Tr}\{\mathcal{O}_i^{\otimes 2} U_L^{\otimes 2} \cdots U_1^{\otimes 2} \tilde{\rho}_k^{\otimes 2} U_1^{\dagger \otimes 2} \cdots U_L^{\dagger \otimes 2}\}.$$

where have used the property that the trace of the tensor product of two matrices is the product of their traces and the abbreviation $\tilde{\rho}_k = \rho_{k-1}^I \otimes \text{Tr}_m\{\rho_{k-1}^R\}$.

We now recall that, under the 2-design hypothesis, the second-order moments of the unitaries U_l satisfy the following identity [51, 52]:

$$\int d\mu(U_l) U_l^{\otimes 2}(\cdot) U_l^{\dagger \otimes 2} = \frac{\text{Tr}\{\mathbb{P}_{2l}\} - D_l^{-1} \text{Tr}\{\mathbb{F}_{2l}\}}{D_l^2 - 1} \mathbb{P}_{2l} + \frac{\text{Tr}\{\mathbb{F}_{2l}\} - D_l^{-1} \text{Tr}\{\mathbb{P}_{2l}\}}{D_l^2 - 1} \mathbb{F}_{2l} + \prod_{l'=1, l' \neq l}^L \mathbb{P}_{2l'}(\cdot) \prod_{l'=1, l' \neq l}^L \mathbb{P}_{2l'},$$

where \mathbb{P}_{2l} is the projector operator: $\mathbb{P}_{2l} = \sum_{s,s'=1}^{D_l} |\psi_l^s\rangle \langle \psi_l^{s'}|$, and \mathbb{F}_{2l} is the Flip operator: $\mathbb{F}_{2l} = \sum_{s,s'=1}^{D_l} |\psi_l^s\rangle \langle \psi_l^{s'}| \langle \psi_l^{s'}| \langle \psi_l^s|$.

From a direct substitution of Eq. (3), we can easily find the leading order term of the desired quantity:

$$\int d\mu(U_l) \text{Tr}\{\mathcal{O}_i \rho_k^R\}^2 \simeq \sum_{l=1}^L (\alpha_l^k)^2 \text{Tr}\{\mathbb{P}_l \mathcal{O}_i\}^2 / D_l^2 + \sum_{l=1}^L \text{Tr}\{\mathbb{P}_l \tilde{\rho}_k \mathbb{P}_l \tilde{\rho}_k\} \text{Tr}\{\mathbb{P}_l \mathcal{O}_i \mathbb{P}_l \mathcal{O}_i\} / D_l^2,$$

concluding that $\int d\mu(U_l) \text{Tr}\{\mathcal{O}_i \rho_k^R\}^2$ and, consequently, $\text{Var}_U[\text{Tr}\{\mathcal{O}_i \rho_k^R\}]$ are $O(e^{-2cn})$. At this point, Chebyshev's inequality concludes the proof. ■

Proof of Theorem 2

Proof:

Combining Eqs. (3) and (4) of the main text, we find an expression for the $\text{Tr}\{\mathcal{O}_i \rho_k^R\}$ mean:

$$\text{Tr}\left\{\left(\bigoplus_{l=1}^L \alpha_l^k \frac{\mathbb{P}_l}{D_l}\right)\left(\bigoplus_l \beta_l \mathbb{P}_l\right)\right\} = \sum_l \alpha_l^k \beta_l = O(1),$$

concluding the proof. ■

Proof of Proposition 1

Proof:

First, consider a scrambler that lacks symmetries. According to Theorem 1, the only observables that avoid exponential concentration are those proportional to the identity. This implies that symmetries are a necessary condition for efficiently extracting a non-trivial reservoir response, which is essential for properly solving the task of interest, as well as any other task.

For a given n -qubit reservoir, we consider a unitary evolution with n symmetry blocks:

$$U = \prod_{l=1}^n U_l,$$

where the Hilbert space on which U_l non-trivially acts is generated by computational basis states containing exactly l ones, making its action global in the set of qubits. It corresponds to the fact that U is symmetric with respect to the operator $S = \sum_{i=1}^n \sigma_i^z$.

For the sake of definiteness, we analyze the reservoir evolution when a series of $|1\rangle\langle 1|$ states is injected. Applying Lemma 1 and performing a direct calculation, we can determine the evolution of the coefficients α_l^k , which define the mean of the distribution from which the reservoir state is sampled:

$$\begin{cases} \alpha_n^{k+1} = \alpha_n^k + \frac{1}{n}\alpha_{n-1}^k, \\ \alpha_l^{k+1} = \frac{l}{n}\alpha_l^k + \left(1 - \frac{l-1}{n}\right)\alpha_{l-1}^k, \quad (0 < l < n), \\ \alpha_0^{k+1} = 0. \end{cases} \quad (\text{S1})$$

From Eq. (S1), we readily observe that the α_l^k coefficients converge exponentially fast, as a function of the time steps, to the values:

$$\alpha_n^\infty = 1; \quad \alpha_l^\infty = 0 \quad \forall l \neq n,$$

with a characteristic convergence time whose upper bound is $O(n^2)$.

In other words, the final reservoir state can be considered as sampled from a distribution whose mean is given by:

$$\bigotimes_{i=1}^n |1\rangle\langle 1|_i. \quad (\text{S2})$$

However, all distribution samples must be compatible with the symmetry expectation values in Eq. (S1). From Lemma 1, we know that $\alpha_l^k = \text{Tr}\{\mathbb{P}_l \rho_k^I \otimes \text{Tr}_m\{\rho_k^R\}\}$, which implies that the reservoir state can only converge to the state defined in Eq. (S2). This property can also be verified by computing the variance of the distribution, following a similar procedure to the one used in the proof of Theorem 1, and confirming that it is equal to zero.

Similarly, for a series of $|0\rangle\langle 0|$ states, the reservoir state will converge to:

$$\bigotimes_{i=1}^n |0\rangle\langle 0|_i.$$

Consequently, measuring σ_i^z for any system qubit immediately allows for discriminating the injected time series, proving Proposition 1. ■

ADDITIONAL NUMERICAL RESULTS

Numerical verification of the Echo state property

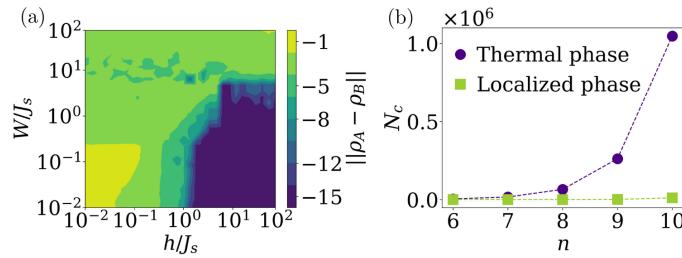


FIG. S1. (a) Convergence of a 7-qubit reservoir state after the injection of 500 random inputs into two random initial conditions. (b) Number of observables that converged after the injection of 1000 inputs, as a function of the number of qubits. The results have been averaged over 100 random realizations, and possible statistical errors have been found to be negligible.

Numerical verification of the echo state property for the fully transverse-field Ising model described in the main text. In Fig. S1 (a), we injected 500 random inputs into two random initial conditions and we calculated the trace distance of the two final states, denoted as ρ_A and ρ_B . We considered a 7-qubit system, varying the hyperparameter values, and averaging over 100 realizations of Hamiltonians and input series.

Moreover, once the Hamiltonian phases were identified, we computed the number of degrees of freedom that properly converged, say N_c , fixing a threshold. In particular, similarly to the previous analysis, we injected 1000 random inputs to a pair of random initial states, considering the cases of thermal ($W=10^{-1}$, $h=10^1$) and many-body localized phases

($W=10^2$, $h=10^1$). Averaging over 100 realizations as in the previous case, we determined N_c as the number elements $\text{Re}\{(\rho_A - \rho_B)_{i,j}\}$ and $\text{Im}\{(\rho_A - \rho_B)_{i,j}\}$ whose absolute value is below 10^{-10} . This analysis is justified by the fact that in the limit of an infinite input sequence, the echo state property reads: $\rho_A - \rho_B = 0$. In Fig. S1 (b), we can clearly observe that, in the thermal phase, N_c tends to exponentially increase with the system size, while, in the localized one, we do not observe a significant change.