Optimal training of finitely-sampled quantum reservoir computers for forecasting of chaotic dynamics

Osama Ahmed[1], Felix Tennie[1], Luca Magri[1,2,3]
[1] *Imperial College London, Department of Aeronautics, Exhibition Road, London, UK*
[2] *The Alan Turing Institute, London, UK*
[3] *Politecnico di Torino, DIMEAS, Corso Duca degli Abruzzi, Torino, Italy*

In the current Noisy Intermediate Scale Quantum (NISQ) era, the presence of noise deteriorates the performance of quantum computing algorithms. On the other hand, Quantum Reservoir Computing (QRC) is a type of Quantum Machine Learning algorithm, which can benefit from different types of tuned noise. In this paper, we analyse the effect that finite-sampling noise has on the chaotic time-series prediction capabilities of QRC and Recurrence-free Quantum Reservoir Computing (RF-QRC). In doing so, we study a three-dimensional Lorenz-63 system and a nine-dimensional turbulent chaotic shear flow model with extreme events. First, we show that finite sampling noise degrades the prediction capabilities of both QRC and RF-QRC while affecting QRC more due to correlated noise. Second, we present a theoretical description of modeling sampling noise in RF-QRC consisting of an ensemble of quantum systems governed by the input time series. Third, we optimize the training of the finite-sampled quantum reservoir computing framework using two methods: (a) Singular Value Decomposition (SVD) applied to the data matrix containing noisy reservoir activation states; and (b) data-filtering techniques to remove the high-frequencies from the noisy reservoir activation states. We show that denoising reservoir activation states improves the SNR and results in a lower training loss with a constant number of samples. Finally, we demonstrate that the training and denoising of the noisy reservoir activation signals in RF-QRC are massively parallelizable on multiple QPUs as compared to the QRC architecture with recurrent connections. This work opens opportunities for using quantum reservoir computing with finite samples for time-series forecasting on near-term quantum hardware.

Introduction

Despite various noise sources affecting the performance of quantum algorithms in NISQ devices, finite sampling noise is a major source of uncertainty in various Quantum Machine Learning (QML) algorithms. It constitutes a fundamental limit to learning in different QML applications [1, 2]. Finite sampling noise roots in the laws of quantum mechanics and, therefore, it must also be taken into account on future fault-tolerant Quantum computers (FTQC) [3]. The calculation of finite expectation values in variational quantum algorithms often results in vanishing gradients and a nearly flat loss landscape, which is also known as the *barren plateaus* [4]. To circumvent this issue, Quantum Extreme Learning Machines (QELM) and Quantum Reservoir Computing (QRC) [5–7] are promising frameworks because they do not require the evaluation of gradients for loss minimization. QRC is inspired by classical reservoir computers [8] - a class of recurrent neural networks (RNNs), which are accurate for time series forecasting [9, 10]. QELM, on the other hand, does not involve recurrence, it is simpler to train, but it has limited applications.

Our recently proposed Recurrence-free QRC (RF-QRC) [11] combines both of these frameworks by avoiding a recurrence built in the quantum circuit, similarly to QELM. This simplifies the training and the recurrence is included as a classical post-processing step. Despite the promising applications of both QRC and RF-QRC, the effect of finite-sampling noise on their time-series prediction capabilities has not been explored yet. In this work, we analyse the effect of finite sampling noise on QRC and RF-QRC frameworks. We focus our analysis on the finite sampling noise for two reasons: (a) even in the presence of various environmental and quantum hardware noises, sampling noise is the dominating noise source in different learning tasks [1], and (b) in some cases, QRC can instead benefit from certain types of tuned noise such as amplitude and phase-damping noise [12]. Therefore, the motivation for this work is to analyse the sampling noise in QRC as well as in RF-QRC and to present methods to reduce its effects.

Quantum Reservoir Computing with finite samples

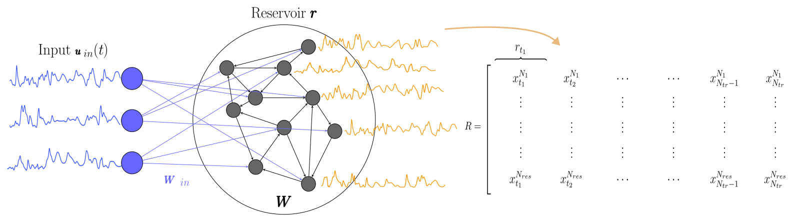


Figure 1: Training phase in reservoir computing. An input time series $\boldsymbol{u}_{in}(t)$ is mapped to a reservoir via the $\boldsymbol{W}_{in}$ matrix. In the reservoir, each neuron *echoes* with the input time series to generate a series of reservoir activation state signals. These are concatenated in the reservoir state matrix $(\boldsymbol{R})$, and used for finding the optimal output weight matrix $\boldsymbol{W}_{out}$.

The training procedure to obtain the weight matrix $\boldsymbol{W}_{out}$ in reservoir computing involves a simple linear ridge regression (Eq. 1). Here, $\boldsymbol{R} \in \mathbb{R}^{N_r \times N_{tr}}$ is a matrix of concatenated reservoir activation signals corresponding to each neuron for N_{tr} time steps of training (Fig. 1).

$$(\boldsymbol{R}\boldsymbol{R}^T + \beta \boldsymbol{I})\,\boldsymbol{W}_{out} = \boldsymbol{R}\boldsymbol{U}_d^{\mathrm{T}}, \tag{1}$$

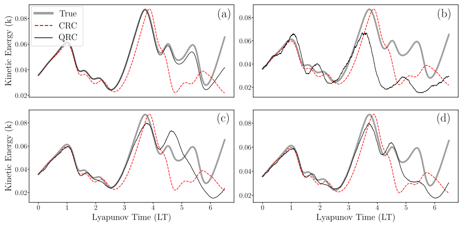
In Fig. 2a, the turbulent flow with extreme events [13] time-series prediction is shown for a different number of sampled circuits S (shots). This indicates that a certain minimum number of finite samples is required to improve the forecasting abilities of QRC beyond classical reservoir computers. In RF-QRC, the effect of finite sampling noise can be modeled as a time-dependent uncorrelated additive noise source with a constant factor of $1/\sqrt{S}$ due to the central limit theorem. In order to describe finite sampling noise, we define the following quantities, assuming that $\boldsymbol{R}$ is the actual (sampling noise-free) reservoir state matrix and that $\boldsymbol{Z}$ is a stochastic variable

$$\overline{\boldsymbol{R}} := \boldsymbol{R} + \frac{1}{\sqrt{S}}\boldsymbol{Z}(t), \qquad \boldsymbol{\Sigma}_{ij}(t) := \mathrm{Cov}[\boldsymbol{Z}_i(t), \boldsymbol{Z}_j(t)], \qquad \boldsymbol{V} := \mathbb{E}\left[\boldsymbol{\Sigma}_{ij}\right] = \mathrm{diag}(\mathrm{mean}(\overline{\boldsymbol{R}})) - \overline{\boldsymbol{R}}\,\overline{\boldsymbol{R}}^T \tag{2}$$

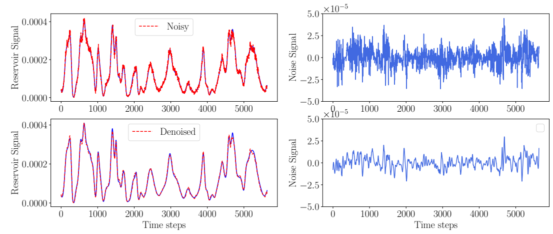
With single qubit expectation values following a binomial distribution [14], we can model $\boldsymbol{Z}(t)$ as a centered multinomial stochastic process. Without loss of generality, $\boldsymbol{Z}(t)$ can be transformed to have a zero mean ($\mathbb{E}\left[\boldsymbol{Z}(t)\right] = 0$). This stochastic noise matrix can be modeled by considering second-order moments, and because our loss function is quadratic, we can neglect higher-order moments. Furthermore, $\boldsymbol{V}$ is the covariance matrix and can be written in terms of the noisy data matrix by considering the second-order cumulants of multinomial distributions. In Fig. 2c, we compare the SNR ratio for QRC and RF-QRC for the nine-dimensional turbulent shear flow model. Our results show that the presence of correlated noise in QRC leads to more noisy estimates of the reservoir activation signals than RF-QRC.

Noise suppression using SVD and signal filtering

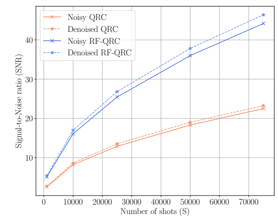
The unbiased estimation of the expectation values in quantum computation is limited by the Cramér-Rao bound [15]. For an ensemble of quantum systems, governed by an input time series, the corresponding expectation values form a reservoir signal with an added finite-sampling noise. We conjecture that the SNR of these noisy reservoir signals could be improved by using classical signal processing tools such as SVD [16, 17]. In Fig. 2d, we display the results of the training error for a 10-qubits ($N_{res} = 1024$) reservoir size trained with a chaotic turbulent shear flow [11, 13] time series with finite samples. We show that SVD improves the SNR and results in a lower training error when compared to the noisy reservoir matrix of the



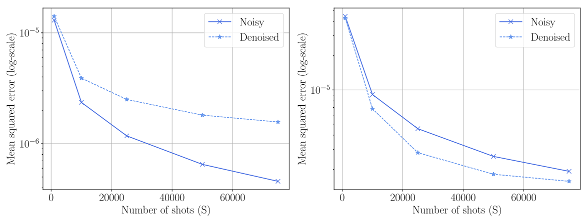
(a) Predictions with (a) ideal distribution (b) 0.5×10^5 (c) 2×10^5 (d) 4×10^5 shots.



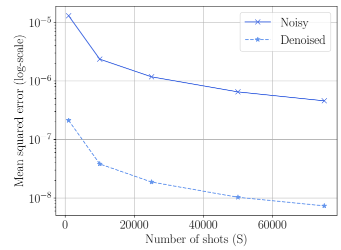
(b) Noisy and denoised reservoir states with noise signals for RF-QRC (0.5×10^5 shots)



(c) SNR for QRC and RF-QRC.



(d) Mean squared error for noisy and denoised activation states with SVD. For all 1024 included reservoir states (left). For only 500 retained reservoir states (right).



(e) Mean squared error for noisy and denoised activation states with signal filtering.

Figure 2: Turbulent chaotic shear flow time series analysis with RF-QRC, 10 qubits systems.

same reduced size. However, using SVD for denoising requires knowledge of the full reservoir matrix $\boldsymbol{R}$, whose dimension scales exponentially with the number of qubits. This renders the SVD analysis approach unfeasible. We propose a method for suppressing the noise of the reservoir activation states by applying low-pass filters to each reservoir activation state individually. We emphasize that in the case of RF-QRC and the absence of recurrence, reservoir activation states are only driven by the input time series, which is known a priori. In an experimental setting, these analyses could also be extended by employing a physical filter to the noisy signal estimates [18] and by using multiple parallel QPUs. Fig. 2e shows a comparison of the mean-squared training error for noisy and denoised reservoir activation states using polynomial regression. By contrast to denoising based on SVD, these results demonstrate that using individual filters for suppressing noise always results in lower training errors.

Conclusion

Quantum reservoir computing has shown promising potential for time-series forecasting of chaotic signals when emulated classically with the assumption of ideal (noise-free) expectation values. To realize any quantum advantage and for real-world applications in weather and climate forecasting, a high-dimensional reservoir and many qubits on quantum hardware are required. The performance of quantum hardware is, however, limited by the presence of environmental and sampling noise. In this work, we study the effect of sampling noise on Lorenz-63 and a turbulent chaotic shear flow model, which exhibits extreme events. The objective of this work is four-fold. First, we compare the effects of finite-sampling noise on quantum reservoir architectures with and without recurrence. We show that the framework of RF-QRC is more resilient to sampling noise than QRC with correlated noise. Second, we present a mathematical framework for modeling uncorrelated noise in RF-QRC based on finite expectations. Third, we show two methods based on SVD and signal filtering to suppress noise in reservoir activation signals. Our results indicate that suppressing noise improves the training accuracy as highlighted by smaller mean-squared training errors. We note that the methods of denoising applied in this work are general and the same analysis could be extended further by employing different advanced techniques for noise filtering to further improve the performance. Fourth, we demonstrate that employing RF-QRC on multiple parallel QPUs coupled with denoising techniques is very feasible. This work opens up opportunities to employ quantum reservoir computing on quantum hardware for chaotic time-series forecasting.

References

1. Hu F, Angelatos G, Khan SA, et al. Tackling sampling noise in physical systems for machine learning applications: Fundamental limits and eigentasks. Physical Review X 2023;13:041020.
2. Mujal P, Martínez-Peña R, Giorgi GL, Soriano MC, and Zambrini R. Time-series quantum reservoir computing with weak and projective measurements. npj Quantum Information 2023;9:16.
3. Preskill J. Fault-tolerant quantum computation. In: *Introduction to quantum computation and information*. World Scientific, 1998:213–69.
4. McClean JR, Boixo S, Smelyanskiy VN, Babbush R, and Neven H. Barren plateaus in quantum neural network training landscapes. Nature communications 2018;9:4812.
5. Fujii K and Nakajima K. Quantum Reservoir Computing: A Reservoir Approach Toward Quantum Machine Learning on Near-Term Quantum Devices. In: *Reservoir Computing: Theory, Physical Implementations, and Applications*. Ed. by Nakajima K and Fischer I. Singapore: Springer Singapore, 2021:423–50. DOI: 10.1007/978-981-13-1687-6_18. URL: https://doi.org/10.1007/978-981-13-1687-6_18.
6. Pfeffer P, Heyder F, and Schumacher J. Hybrid quantum-classical reservoir computing of thermal convection flow. Physical Review Research 2022;4:033176.
7. Mujal P, Martínez-Peña R, Nokkala J, et al. Opportunities in Quantum Reservoir Computing and Extreme Learning Machines. Publication Title: arXiv e-prints ADS Bibcode: 2021arXiv210211831M. 2021. DOI: 10.48550/arXiv.2102.11831. URL: https://ui.adsabs.harvard.edu/abs/2021arXiv210211831M (visited on 12/27/2023).
8. Jaeger H. The" echo state" approach to analysing and training recurrent neural networks-with an erratum note'. Bonn, Germany: German National Research Center for Information Technology GMD Technical Report 2001;148.
9. Racca A and Magri L. Data-driven prediction and control of extreme events in a chaotic flow. Physical Review Fluids 2022;7:104402.
10. Doan NAK, Polifke W, and Magri L. Physics-Informed Echo State Networks for Chaotic Systems Forecasting. In: *Computational Science – ICCS 2019*. Ed. by Rodrigues JMF, Cardoso PJS, Monteiro J, et al. Cham: Springer International Publishing, 2019:192–8.
11. Ahmed O, Tennie F, and Magri L. Prediction of chaotic dynamics and extreme events: A recurrence-free quantum reservoir computing approach. arXiv preprint arXiv:2405.03390 2024.
12. Fry D, Deshmukh A, Chen SYC, Rastunkov V, and Markov V. Optimizing Quantum Noise-induced Reservoir Computing for Nonlinear and Chaotic Time Series Prediction. arXiv:2303.05488 [quant-ph]. 2023. URL: http://arxiv.org/abs/2303.05488 (visited on 03/15/2023).
13. Moehlis J, Faisst H, and Eckhardt B. A low-dimensional model for turbulent shear flows. New Journal of Physics 2004;6:56–6.
14. Schuld M and Petruccione F. Machine learning with quantum computers. Springer, 2021.
15. Yu M, Liu Y, Yang P, et al. Quantum Fisher information measurement and verification of the quantum Cramér–Rao bound in a solid-state qubit. npj Quantum Information 2022;8:56.
16. Schanze T. Removing noise in biomedical signal recordings by singular value decomposition. current directions in biomedical engineering 2017;3:253–6.
17. Jha SK and Yadava R. Denoising by singular value decomposition and its application to electronic nose data processing. IEEE Sensors Journal 2010;11:35–44.
18. Khan SA, Kaufman R, Mesits B, Hatridge M, and Türeci HE. Practical trainable temporal post-processor for multi-state quantum measurement. arXiv preprint arXiv:2310.18519 2023.