## School on Quantum Chaos, ICTP-IFT, São Paulo — 2023

# Majorization-optimized quantum machine learning

G. G. Carlo (CNEA-Conicet, Tandar Laboratory, Buenos Aires) in collaboration with L. Domingo and F. Borondo from Universidad Autónoma de Madrid

# **Summary**

Quantum reservoir computing (QRC) is a quantum machine learning (QML) algorithm that uses a quantum reservoir to process datasets and extract information which is later fed to a classical machine learning model.

The reservoir could be implemented as a quantum circuit in a noisy intermediate-scale quantum (NISQ) computer. A recently developed criterion based on the majorization principle can be applied to select optimal quantum reservoirs, rendering better results than other common models with significantly less gates.

The presence of noise difficults QRC, correcting or mitigating the induced errors is costly. But, **can we benefit from noise?** Surprisingly, we will show that under some specific circumstances, quantum **noise can be used to improve the performance of QRC**. Certain noise types can be beneficial to machine learning, while others should be prioritized for correction. This gives practical prescriptions for successful implementations in nowadays hardware.

# Summary

Quantum reservoir computing (QRC) is a quantum machine learning (QML) algorithm that uses a quantum reservoir to process datasets and extract information which is later fed to a classical machine learning model.



The reservoir could be implemented as a quantum circuit in a noisy intermediate-scale quantum (NISQ) computer. A recently developed criterion based on the majorization principle can be applied to select optimal quantum reservoirs, rendering better results than other common models with significantly less gates.

The presence of noise difficults QRC, correcting or mitigating the induced errors is costly. But, **can we benefit from noise?** Surprisingly, we will show that under some specific circumstances, quantum **noise can be used to improve the performance of QRC**. Certain noise types can be beneficial to machine learning, while others should be prioritized for correction. This gives practical prescriptions for successful implementations in nowadays hardware.

# **Summary**

Quantum reservoir computing (QRC) is a quantum machine learning (QML) algorithm that uses a quantum reservoir to process datasets and extract information which is later fed to a classical machine learning model.

The reservoir could be implemented as a quantum circuit in a noisy intermediate-scale quantum (NISQ) computer. A recently developed criterion based on the majorization principle can be applied to select optimal quantum reservoirs, rendering better results than other common models with significantly less gates.

The presence of noise difficults QRC, correcting or mitigating the induced errors is costly. But, **can we benefit from noise?** Surprisingly, we will show that under some specific circumstances, quantum **noise can be used to improve the performance of QRC**. Certain noise types can be beneficial to machine learning, while others should be prioritized for correction. This gives practical prescriptions for successful implementations in nowadays hardware.

# **Plan of the talk**

- Preliminaries
- Introduction
  - > Quantum Circuits (QC) and Majorization
  - > Quantum Reservoir Computing (QRC)
- QRC model
- Performance of QRC Majorization criterion
- Effects of noise in QRC
  - > Quantum noise models
  - Beneficial and not
- Conclusions

# **Preliminaries**



Google Sycamore 53 qubits (2018) Quantum supremacy (?) (2019)



**Preliminaries** 

Google Sycamore 53 qubits (2018) Quantum supremacy (?) (2019)

IBM Osprey 433 qubits (2022)



#### Qubits number growth exploded in the last few years!



Preli	minar Falcon:	Hummingbird:	Eagle:	Osprey:
	27 qubits	65 qubits	127 qubits	433 qubits
	ingune Second	1314 Quantum	Elf# Quantum Exptr	IBM Quantum Osprey
	2019	2020	2021	2022
	Falcon	Hummingbird	Eagle	Osprey
	27 Qubits	65 Qubits	127 Qubits	433 Qubits



# **Preliminaries**



# **Preliminaries**

**CBPF** CENTRO BRASILEIRO DE PESQUISAS FÍSICAS





Recently, we were able to use a classical supercomputer that simulates up to 35 qubits from Senai Cimatec, "Kuatomu".

-Raúl O. Vallejos -Fernando de Melo -Alexandre B. Tacla -Nina Machado O'Neill -GGC

A quantum circuit consists of a series of quantum logic gates (unitaries) applied to n qubits Q.



computational basis

A quantum circuit consists of a series of quantum logic gates (unitaries) applied to n qubits Q. Measurement

preparation of the input state



Direction of the computation

A quantum circuit consists of a series of quantum logic gates (unitaries) applied to n qubits Q. Measurement computational basis

preparation of the input state



Direction of the computation

 $\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$  $H \vdash$  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (NOT gate)  $- \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$  $\begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$ 0 0 0 0 (CNOT gate)

These are some examples of the most commonly used single and two qubit quantum gates.

computational basis

A quantum circuit consists of a series of quantum logic gates (unitaries) applied to n qubits Q. Measurement



Direction of the computation

We use random quantum circuits where gates from a given family (we call G1={CNOT,H,X} to the example on the left) and the gubits to which they are applied are selected at random.

Sampling from random quantum circuits is strongly believed to be a hard task for classical simulations. and as such it would be a clear demonstration of quantum advantage [A. W. Harrow and A. Montanaro, Nature (London) 549, 203 (2017); A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani, Nat. Phys. 15, 159 (2019)].

Besides the non universal and classically simulatable G1 family, we also use G2={CNOT,H,S] (Clifford, non universal and simulatable) and G3={CNOT,H,T} (universal=approx. any quantum gate at desired precision).

Besides the non universal and classically simulatable G1 family, we also use G2={CNOT,H,S] (Clifford, non universal and simulatable) and G3={CNOT,H,T} (universal=approx. any quantum gate at desired precision).

• We also consider a family of two qubit gates G called matchgate circuits (MG). When acting on nearest-neighbor lines only they are classically simulatable, otherwise they are universal.

$$G(A,B) = \begin{pmatrix} p & 0 & 0 & q \\ 0 & w & x & 0 \\ 0 & y & z & 0 \\ r & 0 & 0 & s \end{pmatrix}, \qquad A = \begin{pmatrix} p & q \\ r & s \end{pmatrix}, \qquad B = \begin{pmatrix} w & x \\ y & z \end{pmatrix}$$

A and B have the same determinant and are randomly chosen according to the Haar measure in the unitary group U(2).

Besides the non universal and classically simulatable G1 family, we also use G2={CNOT,H,S] (Clifford, non universal and simulatable) and G3={CNOT,H,T} (universal=approx. any quantum gate at desired precision).

• We also consider a family of two qubit gates G called matchgate circuits (MG). When acting on nearest-neighbor lines only they are classically simulatable, otherwise they are universal.

$$G(A,B) = \begin{pmatrix} p & 0 & 0 & q \\ 0 & w & x & 0 \\ 0 & y & z & 0 \\ r & 0 & 0 & s \end{pmatrix}, \qquad A = \begin{pmatrix} p & q \\ r & s \end{pmatrix}, \qquad B = \begin{pmatrix} w & x \\ y & z \end{pmatrix}$$

A and B have the same determinant and are randomly chosen according to the Haar measure in the unitary group U(2).

• And finally 3 families of diagonal circuits (non universal/non simulatable): D2 gates applied to pairs of qubits, D3 gates are applied to 3 qubits, and Dn gates to all the qubits.  $\begin{pmatrix} e^{i\phi_1} & 0 & \cdots & 0 \\ 0 & e^{i\phi_2} & \cdots & 0 \end{pmatrix}$ 

$$\binom{n}{2}, \binom{n}{3}$$
 and 1

5

PHYSICAL REVIEW A 104, 012602 (2021)

#### Principle of majorization: Application to random quantum circuits

Raúl O. Vallejos<sup>\*</sup> and Fernando de Melo<sup>†</sup> Centro Brasileiro de Pesquisas Físicas (CBPF), Rua Doutor Xavier Sigaud 150, 22290-180 Rio de Janeiro, Brazil

Gabriel G. Carlo 0

Departamento de Física, Comisión Nacional de Energía Atómica, Avenida del Libertador 8250, (1429) Buenos Aires, Argentina

(Received 17 February 2021; accepted 1 June 2021; published 7 July 2021)

We test the principle of majorization J. I. Latorre and M. A. Martín-Delgado, Phys. Rev. A **66**, 022305 (2002). in random circuits. Three classes of circuits were considered: (i) universal, (ii) classically simulatable, and (iii) neither universal nor classically simulatable. The studied families are: (CNOT, H, T), (CNOT, H, NOT), (CNOT, H, S) (Clifford), matchgates, and IQP (instantaneous quantum polynomial-time). We verified that all the families of circuits satisfy on average the principle of decreasing majorization. In most cases the asymptotic state (number of gates  $\rightarrow \infty$ ) behaves like a random vector. However, clear differences appear in the *fluctuations* of the Lorenz curves associated with asymptotic states. The fluctuations of the Lorenz curves discriminate between universal and nonuniversal classes of random quantum circuits, and they also detect the complexity of some nonuniversal but not classically efficiently simulatable quantum random circuits. We conclude that majorization can be used as an indicator of complexity of quantum dynamics, as an alternative to, e.g., entanglement spectrum and out of-time-order correlators.

# **Introduction - Quantum Circuits and Majorization**

PHYSICAL REVIEW A 104, 012602 (2021)

#### Principle of majorization: Application to random quantum circuits

Raúl O. Vallejos<sup>\*</sup> and Fernando de Melo O<sup>†</sup> Centro Brasileiro de Pesquisas Físicas (CBPF), Rua Doutor Xavier Sigaud 150, 222

#### Gabriel G. Carlo 0

Departamento de Física, Comisión Nacional de Energía Atómica, Avenida del Libertador 8

(Received 17 February 2021; accepted 1 June 2021; published

We test the principle of majorization J. I. Latorre and M. A. Martín-Delgado, Phy in random circuits. Three classes of circuits were considered: (i) universal, (ii) clas neither universal nor classically simulatable. The studied families are: {CNOT, H, T} S} (Clifford), matchgates, and IQP (instantaneous quantum polynomial-time). We v circuits satisfy on average the principle of decreasing majorization. In most cases t of gates  $\rightarrow \infty$ ) behaves like a random vector. However, clear differences appear in t curves associated with asymptotic states. The fluctuations of the Lorenz curves dis and nonuniversal classes of random quantum circuits, and they also detect the comp but not classically efficiently simulatable quantum random circuits. We conclude tl as an indicator of complexity of quantum dynamics, as an alternative to, e.g., enta of-time-order correlators.

### arXiv:2304.04894v1

#### Majorization-based benchmark of the complexity of quantum processors

Alexandre B. Tacla,<sup>1</sup> Nina Machado O'Neill,<sup>1</sup> Gabriel G. Carlo,<sup>2</sup> Fernando de Melo,<sup>1</sup> and Raúl O. Vallejos<sup>1</sup>

<sup>1</sup>Centro Brasileiro de Pesquisas Físicas, Rua Dr. Xavier Sigaud, 150, Rio de Janeiro, RJ, Brazil <sup>2</sup>Departamento de Física, Comisión Nacional de Energía Atómica, Avenida del Libertador 8250, (1429) Buenos Aires, Argentina

Here we investigate the use of the majorization-based indicator introduced in [R. O. Vallejos, F. de Melo, and G. G. Carlo, Phys. Rev. A 104, 012602 (2021)] as a way to benchmark the complexity within reach of quantum processors. By considering specific architectures and native gate sets of currently available technologies, we numerically simulate and characterize the operation of various quantum processors. We characterize their complexity for different native gate sets, qubit connectivity and increasing number of gates. We identify and assess quantum complexity by comparing the performance of each device against benchmark lines provided by randomized Clifford circuits and Haar-random pure states. In this way, we are able to specify, for each specific processor, the number of native quantum gates which are necessary, on average, for achieving those levels of complexity. Lastly, we study the performance of the majorization-based characterization in the presence of distinct types of noise. We find that the majorization-based benchmark holds as long as the circuits' output states have, on average, high purity ( $\gtrsim 0.9$ ). In such cases, the indicator showed no significant differences from the noiseless case.

curves associated with asymptotic states. The fluctuations of the Lorenz curves discriminate between universal and nonuniversal classes of random quantum circuits, and they also detect the complexity of some nonuniversal but not classically efficiently simulatable quantum random circuits. We conclude that majorization can be used as an indicator of complexity of quantum dynamics, as an alternative to, e.g., entanglement spectrum and out-

 $\mathbf{p}, \mathbf{q} \in \mathbb{R}^N$ , we say that  $\mathbf{p}$  is majorized by  $\mathbf{q}$  (or  $\mathbf{q}$  majorizes  $\mathbf{p}$ ), denoted by  $\mathbf{p} \prec \mathbf{q}$ , if



Here, the superscript  $\downarrow$  denotes that the vector components are sorted in non-increasing order. As we will

(1) any  $p_i \ge 0$  and  $\sum_{i=1}^{N} p_i = 1$ , it follows that  $(1/N, 1/N, \dots, 1/N) \prec (p_1, p_2, \dots, p_N) \prec (1, 0, \dots, 0)$ 

These partial sums are called cummulants (Fq(k) or Fp(k))
Their plots vs (k/N) are called Lorenz curves

curves associated with asymptotic states. The fluctuations of the Lorenz curves discriminate between universal and nonuniversal classes of random quantum circuits, and they also detect the complexity of some nonuniversal but not classically efficiently simulatable quantum random circuits. We conclude that majorization can be used as an indicator of complexity of quantum dynamics, as an alternative to, e.g., entanglement spectrum and out-



Decreasing majorization is verified in all averaged cases (500 realizations) since a Lorenz curve above others for all k/N majorizes these latter.

In this Figure we show Lorenz curves for the G1 family of random circuits.

Most coincide asymptotically with Haar-n.

curves associated with asymptotic states. The fluctuations of the Lorenz curves discriminate between universal and nonuniversal classes of random quantum circuits, and they also detect the complexity of some nonuniversal but not classically efficiently simulatable quantum random circuits. We conclude that majorization can be used as an indicator of complexity of quantum dynamics, as an alternative to, e.g., entanglement spectrum and out-

FIG. 3. Fluctuations (standard deviations) of Lorenz curves for asymptotic states generated by various families of random circuits of eight qubits. For each family we considered 5000 samples of 500 gates, except for diagonal circuits which have fixed lengths.

of all strings, i.e.,  $2^n$  measurements. For the entanglement entropy we need to measure only half of the qubits, but we need the full density matrix of them. Therefore, we need  $(2^{n/2})^2 = 2^n$  measurements again. Note, however, that in this case we still need to diagonalize the matrix, which grows



# **Introduction - Quantum Circuits and Majorization**

#### Takeaway

Fluctuations of the Lorenz curves allow to classify circuit families in terms of their universality and/or complexity (smallest and closest to Haar-n).



# **Introduction - Quantum Reservoir Computing**

npj Quantum Information

www.nature.com/npjqi

ARTICLE OPEN Quantum reservoir processing

Sanjib Ghosh<sup>1</sup>, Andrzej Opala<sup>2</sup>, Michał Matuszewski<sup>2</sup>, Tomasz Paterek<sup>1,3</sup> and Timothy C. H. Liew<sup>1,3</sup>

The concurrent rise of artificial intelligence and quantum information poses an opportunity for creating interdisciplinary technologies like quantum neural networks. Quantum reservoir processing, introduced here, is a platform for quantum information processing developed on the principle of reservoir computing that is a form of an artificial neural network. A quantum reservoir processor can perform qualitative tasks like recognizing quantum states that are entangled as well as quantitative tasks like estimating a nonlinear function of an input quantum state (e.g., entropy, purity, or logarithmic negativity). In this way, experimental schemes that require measurements of multiple observables can be simplified to measurement of one observable on a trained quantum reservoir processor.

npj Quantum Information (2019)5:35; https://doi.org/10.1038/s41534-019-0149-8

Quantum reservoir computing and quantum extreme learning machines are two emerging approaches that have demonstrated their potential both in classical and quantum machine learning tasks. They exploit the quantumness of physical systems combined with an easy training strategy, achieving an excellent performance. The increasing interest in these unconventional computing approaches is fueled by the availability of diverse quantum platforms suitable for implementation and the theoretical progresses in the study of complex quantum systems. In this review article, recent proposals and first experiments displaying a broad range of possibilities are reviewed when quantum inputs, quantum physical substrates and quantum tasks are considered. The main focus is the performance of these approaches, on the advantages with respect to classical counterparts and opportunities.

# ervoir Computing



www.advquantumtech.com

# **Opportunities in Quantum Reservoir Computing and Extreme Learning Machines**

Pere Mujal, Rodrigo Martínez-Peña, Johannes Nokkala, Jorge García-Beni, Gian Luca Giorgi, Miguel C. Soriano, and Roberta Zambrini\*

Adv. Quantum Technol. 2021, 4, 2100027

# Introductio

#### npj Quantum Information

#### ARTICLE OPEN Quantum reservoir proce

Sanjib Ghosh <sup>1</sup>, Andrzej Opala<sup>2</sup>, Michał Matuszews

Mach. Learn.: Sci. Technol. 1 (2020) 045027



#### Abstract Excited states of molecules lie in the heart of photochemistry and chemical reactions. The recent development in quantum computational chemistry leads to inventions of a variety of algorithms that calculate the excited states of molecules on near-term quantum computers, but they require more computational burdens than the algorithms for calculating the ground states. In this study, we propose a scheme of supervised quantum machine learning which predicts the excited-state properties of molecules only from their ground state wavefunction resulting in reducing the computational cost for calculating the excited states. Our model is comprised of a quantum reservoir and a classical machine learning unit which processes the measurement results of single-qubit Pauli operators with the output state from the reservoir. The quantum reservoir effectively transforms the single-qubit operators into complicated multi-qubit ones which contain essential information of the system, so that the classical machine learning unit may decode them appropriately. The number of runs for quantum computers is saved by training only the classical machine learning unit, and the whole model requires modest resources of quantum hardware that

appropriately. The number of runs for quantum computers is saved by training only the classical machine learning unit, and the whole model requires modest resources of quantum hardware that may be implemented in current experiments. We illustrate the predictive ability of our model by numerical simulations for small molecules with and without noise inevitable in near-term quantum computers. The results show that our scheme reproduces well the first and second excitation energies as well as the transition dipole moment between the ground states and excited states only from the ground states as inputs. We expect our contribution will enhance the applications of quantum computers in the study of quantum chemistry and quantum materials.

#### Predicting excited states from ground state wavefunction by supervised quantum machine learning









Our QR are random quantum circuits sampled from the families tested with the Majorization criterion.

Our QR are random quantum circuits sampled from the families tested with the Majorization criterion.

We solve a quantum chemistry problem: predicting the energy of the excited states of molecules, given the ground state (quantum data).

Our QR are random quantum circuits sampled from the families tested with the Majorization criterion.

We solve a quantum chemistry problem: predicting the energy of the excited states of molecules, given the ground state (quantum data).

Representing n qubits demands a classical vector of size 2<sup>n</sup> . NISQ devices of around 100 qubits make the classical analogue totally intractable.

Our QR are random quantum circuits sampled from the families tested with the Majorization criterion.

We solve a quantum chemistry problem: predicting the energy of the excited states of molecules, given the ground state (quantum data).

Representing n qubits demands a classical vector of size  $2^n$ . NISQ devices of around 100 qubits make the classical analogue totally intractable.

Example molecules LiH and H<sub>2</sub>O for bond lengths  $R_{\text{LiH}} \in [0.5, 3.5]$  a.u.,  $R_{\text{OH}} \in [0.5, 1.5]$  a.u., and  $\phi_{\text{HOH}} = 104.45^{\circ}$ . The electronic Hamiltonian H(R), the ground  $\psi_0(R)$  state and the energies  $E_1, E_0$  are calculated in the qubit space (8 and 10) using the Jordan-Wigner transformation.

Our QR are random quantum circuits sampled from the families tested with the Majorization criterion.

We solve a quantum chemistry problem: predicting the energy of the excited states of molecules, given the ground state (quantum data).

Representing n qubits demands a classical vector of size 2<sup>n</sup> . NISQ devices of around 100 qubits make the classical analogue totally intractable.

Example molecules LiH and H<sub>2</sub>O for bond lengths  $R_{\text{LiH}} \in [0.5, 3.5]$  a.u.,  $R_{\text{OH}} \in [0.5, 1.5]$  a.u., and  $\phi_{\text{HOH}} = 104.45^{\circ}$ . The electronic Hamiltonian H(R), the ground  $\psi_0(R)$  state and the energies  $E_1, E_0$  are calculated in the qubit space (8 and 10) using the Jordan-Wigner transformation.

We train the model with the real vector  $X(R)=\{X_1,Y_1,Z_1,\dots,X_n,Y_n,Z_n\}$  from local Pauli operators  $P_i$  (i qubit) measurements on the ground states evolved with random circuits (400) for different R values, to predict  $\Delta E_1 = E_1 - E_0$  (the target function) for **unseen** configurations R.

We use the ridge regression, a regularized linear model which minimizes the mean squared error:

$$MSE_{R} = \frac{1}{N_{s}} \sum_{i=0}^{N_{s}} \left[ W \cdot X(\vec{R}_{i}) - \Delta E(\vec{R}_{i}) \right]^{2} + \alpha ||W||^{2}$$
We use the ridge regression, a regularized linear model which minimizes the mean squared error:

$$MSE_{R} = \frac{1}{N_{s}} \sum_{i=0}^{N_{s}} \left[ W \cdot X(\vec{R}_{i}) - \Delta E(\vec{R}_{i}) \right]^{2} + \alpha ||W||^{2}$$

Ns is the number of samples in the training set, W is the matrix of the linear model,  $\alpha$  (10<sup>-7</sup>) is the regularization parameter (prevents overfitting), and  $\|\cdot\|$  is the L2 norm.

We use the ridge regression, a regularized linear model which minimizes the mean squared error:

$$MSE_{R} = \frac{1}{N_{s}} \sum_{i=0}^{N_{s}} \left[ W \cdot X(\vec{R}_{i}) - \Delta E(\vec{R}_{i}) \right]^{2} + \alpha ||W||^{2}$$

Ns is the number of samples in the training set, W is the matrix of the linear model,  $\alpha$  (10<sup>-7</sup>) is the regularization parameter (prevents overfitting), and  $|| \cdot ||$  is the L2 norm.

We have used Ns=90 in our tests which amounts to 30% of the whole 300 R dataset, in a different region (to extrapolate).

We use the ridge regression, a regularized linear model which minimizes the mean squared error:

$$MSE_{R} = \frac{1}{N_{s}} \sum_{i=0}^{N_{s}} \left[ W \cdot X(\vec{R}_{i}) - \Delta E(\vec{R}_{i}) \right]^{2} + \alpha ||W||^{2}$$

Ns is the number of samples in the training set, W is the matrix of the linear model,  $\alpha$  (10<sup>-7</sup>) is the regularization parameter (prevents overfitting), and  $\|\cdot\|$  is the L2 norm.

We have used Ns=90 in our tests which amounts to 30% of the whole 300 R dataset, in a different region (to extrapolate).

We implement the calculations using Qiskit and Sklearn (our code is available on GitHub).





Solid lines LiH molecule and dashed lines H<sub>2</sub>O (this latter is a harder task thus requires more gates). The performance of the different circuits is similar in both cases, and follows a Majorization classification.



Solid lines LiH molecule and dashed lines H<sub>2</sub>O (this latter is a harder task thus requires more gates). The performance of the different circuits is similar in both cases, and follows a Majorization classification.

G1 and G2 are the less complex circuits according to the majorization indicator, with poor performance in QML, worsening as the number of gates increase



Solid lines LiH molecule and dashed lines H<sub>2</sub>O (this latter is a harder task thus requires more gates). The performance of the different circuits is smilar in both cases, and follows a Majorization classification.

G1 and G2 are the less complex circuits according to the majorization indicator, with poor performance in QML, worsening as the number of gates increase

-G3 is the best performer, stabilizes around 200 gates



Solid lines LiH molecule and dashed lines H<sub>2</sub>O (this latter is a harder task thus requires more gates). The performance of the different circuits is smilar in both cases, and follows a Majorization classification.

G1 and G2 are the less complex circuits according to the majorization indicator, with poor performance in QML, worsening as the number of gates increase

G3 is the best performer, stabilizes around 200 gates

Matchgates are slightly worse than G3 but optimal performance comes earlier at around 20-50 gates



Solid lines LiH molecule and dashed lines H<sub>2</sub>O (this latter is a harder task thus requires more gates). The performance of the different circuits is smilar in both cases, and follows a Majorization classification.

G1 and G2 are the less complex circuits according to the majorization indicator, with poor performance in QML, worsening as the number of gates increase

G3 is the best performer, stabilizes around 200 gates

Matchgates are slightly worse than G3 but optimal performance comes earlier at around 20-50 gates

D3 and Dn circuits are similar, although performance of the D2 circuits is significantly worse



Solid lines LiH molecule and dashed lines H<sub>2</sub>O (this latter is a harder task thus requires more gates). The performance of the different circuits is smilar in both cases, and follows a Majorization classification.

G1 and G2 are the less complex circuits according to the majorization indicator, with poor performance in QML, worsening as the number of gates increase

G3 is the best performer, stabilizes around 200 gates

Matchgates are slightly worse than G3 but optimal performance comes earlier at around 20-50 gates

D3 and Dn circuits are similar, although performance of the D2 circuits is significantly worse

Random Ising model is the worst when implemented with gates (Trotter), G3 needs less than 2% of gates

#### Lets think about a toy model of 2 qubits

The state of the system after the action of each circuit can be written as a linear combination of multiPauli operators in  $\{1 \otimes 1, 1 \otimes X, 1 \otimes Y, 1 \otimes Z, \dots, X \otimes Z, Y \otimes Z, Z \otimes Z\}$ .

#### Lets think about a toy model of 2 qubits

The state of the system after the action of each circuit can be written as a linear combination of multiPauli operators in  $\{1 \otimes 1, 1 \otimes X, 1 \otimes Y, 1 \otimes Z, \dots, X \otimes Z, Y \otimes Z, Z \otimes Z\}$ .

Since the Pauli space in the 2-qubit system is a 16-dimensional space, we use a dimensionality reduction technique called UMAP.

L. McInnes, J. Healy, and J. Melville, "UMAP: uniform manifold approximation and projection for dimension reduction," (2020), arXiv:1802.03426

UMAP (Uniform Manifold Approximation and Projection) is a novel manifold learning technique for dimension reduction. UMAP is constructed from a theoretical framework based in Riemannian geometry and algebraic topology. The result is a practical scalable algorithm that is applicable to real world data. The UMAP algorithm is competitive with t-SNE for visualization quality, and arguably preserves more of the global structure with superior run time performance. Furthermore, UMAP has no computational restrictions on embedding dimension, making it viable as a general purpose dimension reduction technique for machine learning.

#### ANALYSIS

nature biotechnology

Dimensionality reduction for visualizing single-cell data using UMAP

Etienne Becht<sup>1</sup>, Leland McInnes<sup>2</sup><sup>©</sup>, John Healy<sup>2</sup>, Charles-Antoine Dutertre<sup>1</sup>, Immanuel W H Kwok<sup>1</sup>, Lai Guan Ng<sup>1</sup>, Florent Ginhoux<sup>1</sup><sup>©</sup> & Evan W Newell<sup>1,3</sup><sup>©</sup>

VOLUME 37 NUMBER 1 JANUARY 2019

#### Lets think about a toy model of 2 qubits

The state of the system after the action of each circuit can be written as a linear combination of multiPauli operators in  $\{1 \otimes 1, 1 \otimes X, 1 \otimes Y, 1 \otimes Z, \dots, X \otimes Z, Y \otimes Z, Z \otimes Z\}$ .

Since the Pauli space in the 2-qubit system is a 16-dimensional space, we use a dimensionality reduction technique called UMAP.

L. McInnes, J. Healy, and J. Melville, "UMAP: uniform manifold approximation and projection for dimension reduction," (2020), arXiv:1802.03426

Dimensionality reduction for visualizing single-cell data using UMAP

nature biotechnology

Etienne Becht<sup>1</sup>, Leland McInnes<sup>2</sup><sup>0</sup>, John Healy<sup>2</sup>, Charles-Antoine Dutertre<sup>1</sup>, Immanuel W H Kwok<sup>1</sup>, Lai Guan Ng<sup>1</sup>, Florent Ginhoux<sup>1</sup><sup>0</sup> & Evan W Newell<sup>1,3</sup><sup>0</sup>

We design 4000 random circuits and see how this fills the Pauli space, compared to the uniform distribution.

#### ANALYSIS



We see that the G1 and G2 circuits fill a subset of the Pauli space. As the number of circuit gates increase, they concentrate in a dense region, biasing results.

G3 fills the Pauli space uniformly and resembles uniform sampling, improving with the number of gates.

Matchgate circuits are similar to G3, except for a small gap, this leads to a slightly worse performance.

Diagonal circuits (which for 2 gubits all coincide) also fill the whole Pauli space, but not so uniformly; this agrees with a slighlty worse performance. 15

Quantum processors are affected by the environment

$$\bullet \quad \epsilon(\rho) = \sum_{m=1} M_m \rho M_m^{\dagger}, \quad \rho = U \, \rho' \, U^{\dagger}$$

We describe the state of the qubit system by a density matrix  $\rho$  after applying a gate U, which transforms into  $\epsilon(\rho)$  after the action of quantum noise. The sum in m gives a non-unitary evolution that can be thought of as coming from a trace performed on the environment (Operator sum or Kraus representation).

Quantum processors are affected by the environment

$$\bullet \quad \epsilon(\rho) = \sum_{m=1} M_m \rho M_m^{\dagger}, \quad \rho = U \, \rho' \, U^{\dagger}$$

We describe the state of the qubit system by a density matrix  $\rho$  after applying a gate U, which transforms into  $\epsilon(\rho)$  after the action of quantum noise. The sum in m gives a non-unitary evolution that can be thought of as coming from a trace performed on the environment (Operator sum or Kraus representation).

$$M_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \quad M_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix} \blacktriangleleft$$

The amplitude damping channel (energy dissipation) is a model of
the decay of an excited two-level atom due to the spontaneous emission of a photon with probability p.

Quantum processors are affected by the environment

$$\bullet \quad \epsilon(\rho) = \sum_{m=1} M_m \rho M_m^{\dagger}, \quad \rho = U \, \rho' \, U^{\dagger}$$

We describe the state of the qubit system by a density matrix  $\rho$  after applying a gate U, which transforms into  $\epsilon(\rho)$  after the action of quantum noise. The sum in m gives a non-unitary evolution that can be thought of as coming from a trace performed on the environment (Operator sum or Kraus representation).

$$M_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \quad M_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix} \blacktriangleleft$$

The amplitude damping channel (energy dissipation) is a model of
the decay of an excited two-level atom due to the spontaneous emission of a photon with probability p.

The phase damping channel models the loss of quantum Information without loss of energy.

Quantum processors are affected by the environment

$$\bullet \quad \epsilon(\rho) = \sum_{m=1} M_m \rho M_m^{\dagger}, \quad \rho = U \, \rho' \, U^{\dagger}$$

We describe the state of the qubit system by a density matrix  $\rho$  after applying a gate U, which transforms into  $\epsilon(\rho)$  after the action of quantum noise. The sum in m gives a non-unitary evolution that can be thought of as coming from a trace performed on the environment (Operator sum or Kraus representation).

$$M_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \quad M_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix} \blacktriangleleft$$

The amplitude damping channel (energy dissipation) is a model of
the decay of an excited two-level atom due to the spontaneous emission of a photon with probability p.

The phase damping channel models the loss of quantum Information without loss of energy.

$$M_0 = \sqrt{1-p} \mathbb{I}, \ M_1 = \sqrt{\frac{p}{3}} X, \ M_2 = \sqrt{\frac{p}{3}} Y, M_3 = \sqrt{\frac{p}{3}} Z$$

The depolarizing channel introduces a Pauli error X, Y or Z with the same probability p/3.



The same chemistry problem for LiH and G3 gates applying idle noise and perfect gates



The same chemistry problem for LiH and G3 gates applying idle noise and perfect gates

For p = 0.0005 (green curve) all noisy amplitude damping reservoirs perform better than the noiseless ones for less than 135 gates



The same chemistry problem for LiH and G3 gates applying idle noise and perfect gates

For p = 0.0005 (green curve) all noisy amplitude damping reservoirs perform better than the noiseless ones for less than 135 gates

Performance of the noisy depolarizing and phase damping reservoirs is always worse than that of the noiseless one, even for small p



The same chemistry problem for LiH and G3 gates applying idle noise and perfect gates

For p = 0.0005 (green curve) all noisy amplitude damping reservoirs perform better than the noiseless ones for less than 135 gates

Rerformance of the noisy depolarizing and phase damping reservoirs is always worse than that of the noiseless one, even for small p

For large numbers of gates all noisy reservoirs decrease their performance, seemingly going to the same growing behavior

We can understand this behavior in the following way: the n qubit state can be written in terms of multiPauli operators that are the direct product of single qubit ones as

 $\rho = \sum_{i} a_{i} P_{i}, \text{ with } a_{i} = \frac{1}{2^{n}} \operatorname{tr}(P_{i}\rho),$  $\epsilon(\rho) = \sum_{i} b_{i} P_{i}, \text{ with } b_{i} = \frac{1}{2^{n}} \operatorname{tr}[P_{i}\epsilon(\rho)].$ 

We can understand this behavior in the following way: the n qubit state can be written in terms of multiPauli operators that are the direct product of single qubit ones as

The action of the noise channels on the single qubit Pauli operators is

 $\rho = \sum_{i} a_{i} P_{i}, \text{ with } a_{i} = \frac{1}{2^{n}} \operatorname{tr}(P_{i}\rho),$  $\epsilon(\rho) = \sum_{i} b_{i} P_{i}, \text{ with } b_{i} = \frac{1}{2^{n}} \operatorname{tr}[P_{i}\epsilon(\rho)].$ 

	Amplitude damping	Depolarizing	Phase damping
$\epsilon(X)$	$\sqrt{1-p} X$	$(1-\frac{4}{3}p)X$	(1-p) X
$\epsilon(Y)$	$\sqrt{1-p} Y$	$(1-\frac{4}{3}p)Y$	(1-p) Y
$\epsilon(Z)$	(1-p) Z	$(1 - \frac{4}{3}p)Z$	Z
$\epsilon(\mathbb{I})$	$\mathbb{I} + pZ$	$\mathbb{I}$	I

 $\epsilon(P_i)$  is always proportional to  $P_i$ , except for  $\epsilon(I)$  in the amplitude damping channel case (it is said to be non unital)

We can understand this behavior in the following way: the n qubit state can be written in terms of multiPauli operators that are the direct product of single gubit ones as  $\rho = \sum_{i} a_{i} P_{i}, \text{ with } a_{i} = \frac{1}{2^{n}} \operatorname{tr}(P_{i}\rho),$  $\epsilon(\rho) = \sum_{i} b_{i} P_{i}, \text{ with } b_{i} = \frac{1}{2^{n}} \operatorname{tr}[P_{i}\epsilon(\rho)].$ 

The action of the noise channels on the single qubit Pauli operators is

	Amplitude damping	Depolarizing	Phase damping
$\epsilon(X)$	$\sqrt{1-p} X$	$(1 - \frac{4}{3}p)X$	(1-p) X
$\epsilon(Y)$	$\sqrt{1-p} Y$	$(1-\frac{4}{3}p)Y$	(1-p) Y
$\epsilon(Z)$	(1-p) Z	$(1 - \frac{4}{3}p)Z$	Z
$\epsilon(\mathbb{I})$	$\mathbb{I} + pZ$	$\mathbb{I}$	$\mathbb{I}$

 $\epsilon(P_i)$  is always proportional to  $P_i$ , except for  $\epsilon(I)$  in the amplitude damping channel case (it is said to be non unital)

For any Pauli operator P<sub>i</sub>, the coefficient in the Pauli space with the depolarizing and phase damping channels is attenuated

$$b_i = \frac{1}{2^n} \operatorname{tr}[P_i \ \epsilon(\rho)] = \frac{1}{2^n} \alpha_i \ tr(P_i \rho) = \alpha_i \ a_i, \quad 0 \le \alpha_i \le 1,$$

We can understand this behavior in the following way: the n gubit state can be written in terms of multiPauli operators that are the direct product of single gubit ones as  $\rho = \sum_{i} a_{i} P_{i}, \quad \text{with } a_{i} = \frac{1}{2^{n}} \operatorname{tr}(P_{i}\rho),$  $\epsilon(\rho) = \sum_{i} b_{i} P_{i}, \quad \text{with } b_{i} = \frac{1}{2^{n}} \operatorname{tr}[P_{i}\epsilon(\rho)].$ 

The action of the noise channels on the single qubit Pauli operators is

	Amplitude damping	Depolarizing	Phase damping
$\epsilon(X)$	$\sqrt{1-p} X$	$(1 - \frac{4}{3}p)X$	(1-p) X
$\epsilon(Y)$	$\sqrt{1-p} Y$	$(1-\frac{4}{3}p)Y$	(1-p) Y
$\epsilon(Z)$	(1-p) Z	$(1 - \frac{4}{3}p)Z$	Z
$\epsilon(\mathbb{I})$	$\mathbb{I} + pZ$	$\mathbb{I}$	I

 $\epsilon(P_i)$  is always proportional to  $P_i$ , except for  $\epsilon(I)$  in the amplitude damping channel case (it is said to be non unital)

For any Pauli operator P<sub>i</sub>, the coefficient in the Pauli space with the depolarizing and phase damping channels is attenuated

$$b_i = \frac{1}{2^n} \operatorname{tr}[P_i \ \epsilon(\rho)] = \frac{1}{2^n} \alpha_i \ tr(P_i \rho) = \alpha_i \ a_i, \quad 0 \le \alpha_i \le 1,$$

It can be easily seen that for amplitude damping (acting on j qubit for example), when there is a non null component  $a_k$  associated to  $P_k = P_0 \otimes \cdots \otimes P_{i-1} \otimes I \otimes P_{i+1} \otimes \cdots \otimes P_n$ ,  $b_i = (1-p) a_i + p a_k$  (associated to  $P_i = (1-p) a_i + p a_k$ )  $P_0 \otimes \cdots \otimes P_{i-1} \otimes Z \otimes P_{i+1} \otimes \cdots \otimes P_n$ ). Even if a=0, b is not, adding an extra component in the Pauli space.



Two qubits toy model revisited

Distribution of the Pauli coefficients at the end of the random circuit with 10 gates and an error rate p = 0.2



Two qubits toy model revisited

Distribution of the Pauli coefficients at the end of the random circuit with 10 gates and an error rate p = 0.2

All noise models attenuate the non-zero coefficients

Shadowed area shows a region where the noiseless simulation (and depolarizing and phase damping) give zero expectation values



Two qubits toy model revisited

Distribution of the Pauli coefficients at the end of the random circuit with 10 gates and an error rate p = 0.2

All noise models attenuate the non-zero coefficients

Shadowed area shows a region where the noiseless simulation (and depolarizing and phase damping) give zero expectation values

Amplitude damping circuit has non-zero expectation values for the same operators, introducing non-zero terms in the Pauli distribution (similar effect as having more quantum gates in the circuit)



4000 simulations of this toy model confirm that the amplitude damping channel fills the Pauli space faster than the other circuits, including the noiseless QR, equivalently as having more quantum gates

The QRs with higher complexity according to the **majorization criterion** provide better results in **QML** tasks. The optimal family of quantum circuits fills uniformly the Pauli space of operators.

The QRs with higher complexity according to the **majorization criterion** provide better results in **QML** tasks. The optimal family of quantum circuits fills uniformly the Pauli space of operators.

This turns the **majorization criterion** into a powerful tool at the time to evaluate the quantum circuits and device suitability for QML.

The QRs with higher complexity according to the **majorization criterion** provide better results in **QML** tasks. The optimal family of quantum circuits fills uniformly the Pauli space of operators.

This turns the **majorization criterion** into a powerful tool at the time to evaluate the quantum circuits and device suitability for QML.

Under certain circumstances, **noise could be beneficial to QML.** We show that for error rates  $p \le 0.0005$ , an **amplitude damping** channel presence renders better performance than noiseless QRs for ML tasks.

The QRs with higher complexity according to the **majorization criterion** provide better results in **QML** tasks. The optimal family of quantum circuits fills uniformly the Pauli space of operators.

This turns the **majorization criterion** into a powerful tool at the time to evaluate the quantum circuits and device suitability for QML.

Under certain circumstances, **noise could be beneficial to QML**. We show that for error rates  $p \le 0.0005$ , an **amplitude damping** channel presence renders better performance than noiseless QRs for ML tasks.

The **depolarizing channel** strongly attenuates the Pauli components, so our prescription is that its **correction should be a priority**.

The behavior of **Lorenz curves fluctuations** could be predicted using **Random Matrix** models, a field widely used and developed in Quantum Chaos.

The behavior of **Lorenz curves fluctuations** could be predicted using **Random Matrix** models, a field widely used and developed in Quantum Chaos.

Our quantum machine learning method is promising in terms of predicting **excited states of complex and quantum chaotic systems** which are difficult to evaluate in general.
## Conclusions

The behavior of **Lorenz curves fluctuations** could be predicted using **Random Matrix** models, a field widely used and developed in Quantum Chaos.

Our quantum machine learning method is promising in terms of predicting **excited states of complex and quantum chaotic systems** which are difficult to evaluate in general.

Also, the very demanding task of calculating **localization measures** in phase space could be simplified by directly feeding the quantum states to our procedure.

## Conclusions

The behavior of **Lorenz curves fluctuations** could be predicted using **Random Matrix** models, a field widely used and developed in Quantum Chaos.

Our quantum machine learning method is promising in terms of predicting **excited states of complex and quantum chaotic systems** which are difficult to evaluate in general.

Also, the very demanding task of calculating **localization measures** in phase space could be simplified by directly feeding the quantum states to our procedure.

- Taking advantage of noise in quantum reservoir computing L. Domingo, G. Carlo, and F. Borondo, Sci Rep 13, 8790 (2023).
- 2. Optimal quantum reservoir computing for the noisy intermediate-scale quantum era L. Domingo, G. Carlo, and F. Borondo, Phys. Rev. E 106, L043301 (2022).

## Conclusions

LAVANGUARDIA

#### UNIVERSIDADES INVESTIGACIÓN

# El ruido es "un aliado inesperado" para la computación cuántica, dice un estudio de la UAM

• Madrid, 7 jun (EFE).- Un grupo de investigadores de la Universidad Autónoma de Madrid (UAM) ha comprobado que el uso del ruido en los ordenadores cuánticos puede ayudar a mejorar los resultados de algunos algoritmos, lo que abre "nuevas posibilidades" en el desarrollo de la computación cuántica.

### And in Spanish newspapers!..

- Taking advantage of noise in quantum reservoir computing L. Domingo, G. Carlo, and F. Borondo, Sci Rep 13, 8790 (2023).
- 2. Optimal quantum reservoir computing for the noisy intermediate-scale quantum era L. Domingo, G. Carlo, and F. Borondo, Phys. Rev. E 106, L043301 (2022).

### Majorization-optimized quantum machine learning

## THANK YOU!