

Breaking concentration barriers for quantum extreme learning on digital quantum processors*

Timothée Dao^{†1,2}, Ege Yilmaz^{†3,1}, Ibrahim Shehzad⁴, Christophe Pere⁵, Kumar Ghosh⁶, Isabelle Wittmann¹, Thomas Brunschwiler¹, Giorgio Cortiana⁶, Corey O’Meara⁶, Stefan Woerner¹, and Francesco Tacchino¹

¹IBM Research Europe – Zurich, 8803 Rüschlikon, Switzerland

²Institute for Theoretical Physics, ETH Zürich, 8093 Zürich, Switzerland

³Department of Mathematics and Computer Science, University of Basel, 4051 Basel, Switzerland

⁴IBM Quantum, T. J. Watson Research Center, Yorktown Heights, New York 10598, USA

⁵PINQ2, Sherbrooke, QC, Canada

⁶E.ON Digital Technology GmbH, Hanover, Germany

Reservoir computing (RC) provides a powerful paradigm for processing nonlinear and temporal data using fixed dynamical systems combined with simple trainable readouts [1, 2]. Quantum extensions of RC leverage the exponentially large Hilbert space of quantum systems to enhance expressivity [3–5], offering a promising route toward quantum advantage in machine learning. However, scaling such approaches on digital quantum processors is fundamentally challenged by hardware noise and by concentration phenomena [6–8], which cause output observables to become statistically indistinguishable across inputs as system size increases. These effects, closely related to barren plateaus [9] and kernel concentration [10], imply that even though models remain trainable in principle, their features become unresolvable under realistic sampling budgets.

In this work, we present a theory-informed Quantum Extreme Learning Machine (QELM) [6, 11–15] architecture – a precursor of full quantum RC architectures – designed specifically to address concentration effects on state-of-the-art superconducting quantum processors and demonstrate its performance at previously unexplored scales. QELMs belong to the family of quantum feature-based models [16, 17]: given an input vector $\mathbf{u}$, a quantum state

$$\rho(\mathbf{u}) = \left(\prod_{j=1}^{\ell} U_{\text{RES}} U^{(j)}(\mathbf{u}) \right) \rho_0 \left(\prod_{j=1}^{\ell} U_{\text{RES}} U^{(j)}(\mathbf{u}) \right)^{\dagger}$$

is prepared through alternating data encoding and reservoir evolution layers, and features are extracted as expectation values $r_i(\mathbf{u}) = \text{Tr}[\hat{R}_i \rho(\mathbf{u})]$ of a fixed observable set $\mathcal{R} = \{\hat{R}_i\}_{i=1}^M$. A classical linear readout is then trained offline [18]. This structure enables efficient execution on cloud quantum hardware, avoids iterative parameter updates on the quantum device, and scales linearly with dataset size.

Our implementation is based on circuits inspired by periodically driven (kicked) Ising dynamics [19, 20], arranged in a brickwork topology on a ring of qubits initialized in Bell pairs. The dynamics are governed by a static Hamiltonian $\hat{H}_0 = J \sum_{i=0}^{N-2} \hat{Z}_i \hat{Z}_{i+1} + h \sum_{n=0}^{N-1} \hat{Z}_i$ and a periodically applied kick $\hat{H}_K = b \sum_{n=0}^{N-1} \hat{X}_i$, where J denotes the spin-spin coupling, h the longitudinal field, and b the transverse kick strength. Each layer consists of parallel two-qubit blocks parameterized by these coupling and field strengths, with alternating *encoding layers*, where inputs modulate transverse fields [21, 22], and *dynamical layers* implementing fixed reservoir evolution. This architecture, illustrated in Fig. 1 (left), combines strong entanglement generation with structured locality, making it both expressive and hardware-efficient. Importantly, all quantum parameters are drawn from a fixed distribution and remain fixed after initialization, while only the classical readout is optimized.

*Full paper available at <https://arxiv.org/abs/2603.13005>

[†]These authors contributed equally to this work.

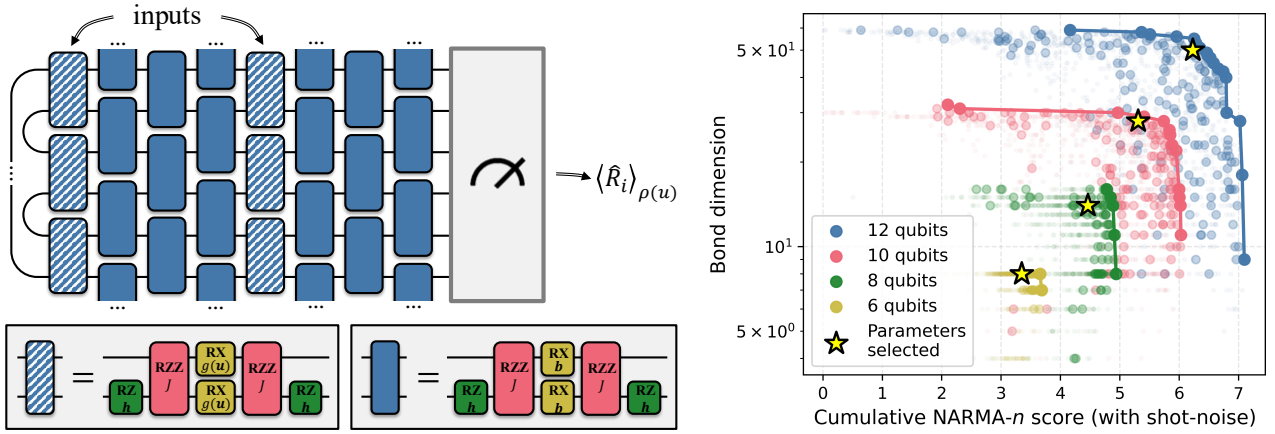


Figure 1: **Left:** Schematic of the QELM architecture. Input vectors $\mathbf{u}$ are encoded through alternating encoding and dynamics layers on a ring of qubits initialized in Bell pairs. Encoding layers apply data-dependent transverse kicks ($b = g(u[i])$), while dynamics layers implement fixed kicked-Ising evolution with parameters (J, h, b) drawn from a fixed distribution at initialization. Measured expectation values form the feature vector $\mathbf{r}(\mathbf{u})$, which feeds a classical linear readout. **Right:** Pareto-front projection showing the trade-off between NARMA performance under realistic shot noise and bond dimension across QELM architectures of different sizes (6, 8, 10, and 12 qubits). Each dot represents a sampled hyperparameter configuration, colored by system size, with solid lines indicating the corresponding Pareto fronts. The yellow star marks the selected hyperparameters (Table in full paper), which consistently lie near the Pareto-optimal region across model sizes, demonstrating a universal operating regime that generalizes well across architectures.

A central contribution of this work is the systematic identification of an operating regime that balances expressivity, complexity, and statistical resolvability. We develop a multi-objective hyperparameter tuning framework [23] based on three key metrics: (i) predictive performance on NARMA time-series tasks [24, 25], measured by the cumulative coefficient of determination; (ii) output variability across inputs, capturing sensitivity to data and diagnosing concentration effects; and (iii) the bond dimension of matrix product state approximations of the output, used as a proxy for quantum complexity. By jointly optimizing these objectives, we identify parameter regions that avoid both under-expressive and over-scrambling regimes [6, 7, 25], as shown in Fig. 1 (right). Remarkably, hyperparameters selected at small system sizes (8–10 qubits) generalize robustly across larger architectures and different tasks, providing evidence for a universal operating regime [26–28]. This transferability represents a significant practical contribution: it enables efficient hyperparameter optimization on classically simulable systems while maintaining performance at large scale, thereby circumventing the prohibitive cost of direct optimization on hundreds of qubits. This property is crucial for scaling quantum machine learning applications beyond the regime accessible to classical simulation.

To further mitigate the impact of noise and concentration, we introduce a local eigentask framework for feature extraction [29–31]. Eigentasks correspond to orthogonal linear combinations of measurement outcomes ranked by their noise-to-signal ratio (NSR), identifying directions in feature space that are maximally robust to sampling noise. While global eigentask analysis scales exponentially with system size, we develop a local variant operating on one- and two-qubit subsets, yielding polynomial complexity and compatibility with randomized Pauli measurement protocols [32, 33]. This approach allows us to construct adaptive, noise-aware observables that significantly improve feature quality and model performance in realistic experimental regimes.

We validate our framework through large-scale experiments on IBM superconducting quantum processors [19, 34] using up to 124 qubits and circuits with more than 5,000 two-qubit gates. On NARMA- n benchmarks [24, 25], QELMs with 24 and 72 qubits demonstrate effective memory capacities proportional to the number of qubits, maintaining predictive accuracy up to task orders matching the rolling input window size (Fig. 2, left). These results confirm that increasing system size directly enhances the temporal processing capability of the model, albeit with increasing sensitivity to noise.

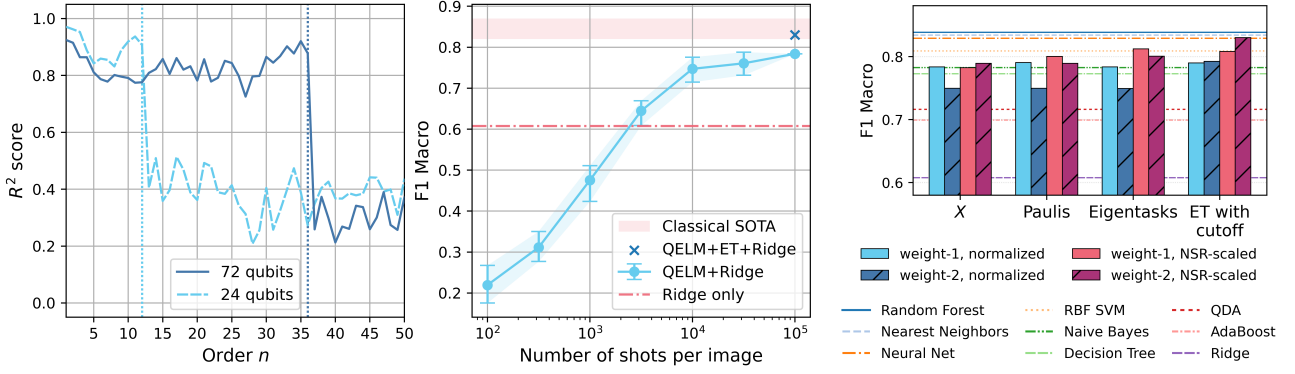


Figure 2: Hardware validation and performance analysis of large-scale QELM architectures on IBM’s `ibm_quebec` processor. **Left:** Performance on the NARMA- n benchmark for 24- and 72-qubit QELMs. The coefficient of determination R^2 is reported versus task order n ; vertical dotted lines indicate the rolling-window size $n = L = N/2$. The 72-qubit model maintains accuracy to larger n , consistent with its greater effective memory capacity. **Middle:** Classification performance on the Landsat dataset for the 124-qubit QELM. Test F1 macro scores are shown as a function of measurement shot count. The QELM+Ridge curve demonstrates systematic improvement with increasing shot precision, substantially outperforming a Ridge classifier trained directly on raw classical features (horizontal line at 0.608). At maximum shot count (10^5 shots), eigentask-based readouts (QELM+ET+Ridge, rightmost point) achieve F1 scores up to 0.83, approaching the performance range of state-of-the-art classical methods (shaded region: 0.82–0.87). Error bars denote 95% bootstrap confidence intervals. **Right:** Classification performance (F1 macro) for the 124-qubit Landsat experiment using different QELM readouts: single-basis X , full Pauli, local eigentasks (ET), and ET with NSR-based cutoff. Bars compare weight-1 vs. weight-2 features and unit-variance vs. signal-aware scaling. Horizontal lines denote classical baselines trained directly on raw features. Eigentask-based readouts with NSR cutoff achieve the highest performance, demonstrating effective noise-adaptive feature extraction at utility scale.

We further evaluate the QELM on a real-world classification task using the Statlog Landsat satellite dataset [35]. Using a 124-qubit implementation and randomized Pauli measurements [32], we extract quantum features and train a linear Ridge classifier. As shown in Fig. 2 (middle), test performance improves systematically with increasing shot count, demonstrating that the model operates in a regime where signal remains resolvable despite sampling noise. The QELM+Ridge approach substantially outperforms a Ridge classifier trained directly on raw features ($F1 = 0.608$), confirming that quantum reservoir features significantly improve linear separability [17]. At maximum shot count (10^5 shots), eigentask-based readouts (QELM+ET+Ridge) achieve F1 macro scores up to 0.83, approaching the performance of state-of-the-art classical methods such as random forests ($F1 = 0.84$) and k -nearest neighbours ($F1 = 0.83$; Fig. 2, right), thereby demonstrating the practical viability of quantum feature extraction at utility scale.

Our analysis also provides insight into the interplay between observable weight, variance, and resolvability. While higher-weight observables can encode richer correlations, they exhibit reduced variability and increased sampling noise, making them harder to resolve. The local eigentask framework naturally balances these effects by selecting informative low-weight combinations that maximize resolvable signal, enabling efficient use of measurement resources at scale.

Overall, this work demonstrates that quantum reservoir-like models can operate in nontrivial, noise-robust regimes on current-generation quantum hardware at scales beyond direct classical simulation. By combining hardware-aware circuit design, principled hyperparameter selection, and noise-adaptive feature extraction, we establish a practical and scalable approach to pre-fault-tolerant quantum machine learning. Our findings provide a foundation for extending these techniques to fully dynamical quantum reservoirs [26, 36–38] and more complex learning tasks, bringing quantum machine learning closer to real-world applicability.

References

- ¹M. Lukoševičius, H. Jaeger, and B. Schrauwen, “Reservoir computing trends”, *KI-Künstliche Intelligenz* **26**, 365–371 (2012).
- ²M. Yan, C. Huang, P. Bienstman, P. Tino, W. Lin, and J. Sun, “Emerging opportunities and challenges for the future of reservoir computing”, *Nature Communications* **15**, 2056 (2024).
- ³K. Fujii and K. Nakajima, “Quantum reservoir computing: a reservoir approach toward quantum machine learning on near-term quantum devices”, *Reservoir Computing: Theory, Physical Implementations, and Applications*, 423–450 (2021).
- ⁴P. Mújal, R. Martínez-Peña, J. Nokkala, J. García-Bení, G. L. Giorgi, M. C. Soriano, and R. Zambrini, “Opportunities in quantum reservoir computing and extreme learning machines”, *Advanced Quantum Technologies* **4**, 2100027 (2021).
- ⁵C. Gyurik, F. Wudarski, E. J. Philip, A. Sannia, H. Sadeghi, O. Kyriienko, D. Venturelli, and A. A. Gentile, “From quantum feature maps to quantum reservoir computing: an applicative perspective”, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **384** (2026).
- ⁶W. Xiong, G. Facelli, M. Sahebi, O. Agnel, T. Chotibut, S. Thanasilp, and Z. Holmes, “On fundamental aspects of quantum extreme learning machines”, *Quantum Machine Intelligence* **7**, 20 (2025).
- ⁷W. Xiong, Z. Holmes, A. Angrisani, Y. Suzuki, T. Chotibut, and S. Thanasilp, “Role of scrambling and noise in temporal information processing with quantum systems”, *arXiv preprint arXiv:2505.10080* (2025).
- ⁸A. Sannia, G. L. Giorgi, and R. Zambrini, “Exponential concentration and symmetries in quantum reservoir computing”, *arXiv preprint arXiv:2505.10062* (2025).
- ⁹M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo, “Barren plateaus in variational quantum computing”, *Nature Reviews Physics*, 1–16 (2025).
- ¹⁰S. Thanasilp, S. Wang, M. Cerezo, and Z. Holmes, “Exponential concentration in quantum kernel methods”, *Nature communications* **15**, 5200 (2024).
- ¹¹S. Ghosh, A. Opala, M. Matuszewski, T. Paterek, and T. C. Liew, “Quantum reservoir processing”, *npj Quantum Information* **5**, 35 (2019).
- ¹²L. Innocenti, S. Lorenzo, I. Palmisano, A. Ferraro, M. Paternostro, and G. M. Palma, “Potential and limitations of quantum extreme learning machines”, *Communications Physics* **6**, 118 (2023).
- ¹³A. Suprano, D. Zia, L. Innocenti, S. Lorenzo, V. Cimini, T. Giordani, I. Palmisano, E. Polino, N. Spagnolo, F. Sciarrino, G. M. Palma, A. Ferraro, and M. Paternostro, “Experimental property reconstruction in a photonic quantum extreme learning machine”, *Phys. Rev. Lett.* **132**, 160802 (2024).
- ¹⁴M. 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**, 20 (2025).
- ¹⁵M. Kawanabe, S. Cindrak, K. Luedge, J.-i. Shirakashi, T. Shibuya, and H. Imai, *Efficient time-series prediction on nisy devices via time-delayed quantum extreme learning machine*, 2026.
- ¹⁶V. Bokov, L. Kohl, S. Schmitt, and V. Dunjko, *Machine learning with minimal use of quantum computers: provable advantages in learning under quantum privileged information (luqpi)*, 2026.
- ¹⁷Q. Zhang, A. Simen, C. Flores-Garrigós, G. A. Barrios, P. A. Erdman, E. Solano, A. C. Kemp, V. Beltrani, V. Pathak, and H. Mohammadbagherpoor, *Quantum-enhanced satellite image classification*, 2026.
- ¹⁸J. Wang, S. Lu, S.-H. Wang, and Y.-D. Zhang, “A review on extreme learning machine”, *Multimedia Tools and Applications* **81**, 41611–41660 (2022).

- ¹⁹L. E. Fischer, M. Leahy, A. Eddins, N. Keenan, D. Ferracin, M. A. Rossi, Y. Kim, A. He, F. Pietracaprina, B. Sokolov, et al., “Dynamical simulations of many-body quantum chaos on a quantum computer”, arXiv preprint arXiv:2411.00765 (2024).
- ²⁰B. Bertini, P. W. Claeys, and T. Prosen, *Exactly solvable many-body dynamics from space-time duality*, 2025.
- ²¹A. Pérez-Salinas, A. Cervera-Lierta, E. Gil-Fuster, and J. I. Latorre, “Data re-uploading for a universal quantum classifier”, *Quantum* **4**, 226 (2020).
- ²²M. Schuld, R. Sweke, and J. J. Meyer, “Effect of data encoding on the expressive power of variational quantum-machine-learning models”, *Phys. Rev. A* **103**, 032430 (2021).
- ²³T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, *Optuna: a next-generation hyperparameter optimization framework*, 2019.
- ²⁴T. Kubota, Y. Suzuki, S. Kobayashi, Q. H. Tran, N. Yamamoto, and K. Nakajima, “Temporal information processing induced by quantum noise”, *Physical Review Research* **5**, 023057 (2023).
- ²⁵R. Martínez-Peña, G. L. Giorgi, J. Nokkala, M. C. Soriano, and R. Zambrini, “Dynamical phase transitions in quantum reservoir computing”, *Physical Review Letters* **127**, 100502 (2021).
- ²⁶M. Kornjača, H.-Y. Hu, C. Zhao, J. Wurtz, P. Weinberg, M. Hamdan, A. Zhdanov, S. H. Cantu, H. Zhou, R. A. Bravo, et al., “Large-scale quantum reservoir learning with an analog quantum computer”, arXiv preprint arXiv:2407.02553 (2024).
- ²⁷K. Kobayashi, K. Fujii, and N. Yamamoto, “Feedback-driven quantum reservoir computing for time-series analysis”, *PRX Quantum* **5**, 040325 (2024).
- ²⁸P. D. Solanki and A. Pham, *Harnessing quantum dynamics for robust and scalable quantum extreme learning machines*, 2025.
- ²⁹F. 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 Eigentasks”, en, *Physical Review X* **13**, 041020 (2023).
- ³⁰R. Martínez-Peña, J. Nokkala, G. L. Giorgi, R. Zambrini, and M. C. Soriano, “Information processing capacity of spin-based quantum reservoir computing systems”, *Cognitive Computation* **15**, 1440–1451 (2023).
- ³¹M. Gross and H.-M. Riesen, *Kernel-based optimization of measurement operators for quantum reservoir computers*, 2026.
- ³²H.-Y. Huang, R. Kueng, and J. Preskill, “Predicting many properties of a quantum system from very few measurements”, *Nature Physics* **16**, 1050–1057 (2020).
- ³³L. E. Fischer, T. Dao, I. Tavernelli, and F. Tacchino, “Dual-frame optimization for informationally complete quantum measurements”, *Phys. Rev. A* **109**, 062415 (2024).
- ³⁴L. E. Fischer, D. Bultrini, I. Tavernelli, and F. Tacchino, *Large-scale implementation of quantum subspace expansion with classical shadows*, 2025.
- ³⁵A. Srinivasan, *Statlog (landsat satellite)*, <https://archive.ics.uci.edu/dataset/146/statlog+landsat+satellite>, accessed on 2025-08-08, UCI Machine Learning Repository.
- ³⁶M. Negoro, K. Mitarai, K. Nakajima, and K. Fujii, “Toward nmr quantum reservoir computing”, *Reservoir computing: theory, physical implementations, and applications*, 451–458 (2021).
- ³⁷V. Cimini, M. M. Sohoni, F. Presutti, B. K. Malia, S.-Y. Ma, R. Yanagimoto, T. Wang, T. Onodera, L. G. Wright, and P. L. McMahon, “Large-scale quantum reservoir computing using a gaussian boson sampler”, arXiv preprint arXiv:2505.13695 (2025).
- ³⁸D. Zia, L. Innocenti, G. Minati, S. Lorenzo, A. Suprano, R. D. Bartolo, N. Spagnolo, T. Giordani, V. Cimini, G. M. Palma, A. Ferraro, F. Sciarrino, and M. Paternostro, “Quantum reservoir computing for photonic entanglement witnessing”, *Science Advances* **11**, eady7987 (2025).