

# A continuous-variable photonic memristor as a reinforcement learner

Dean Brand<sup>1,2</sup>, Francesco Petruccione<sup>2,3</sup>

<sup>1</sup>*Department of Physics, Stellenbosch University, Stellenbosch, South Africa*

<sup>2</sup>*National Institute for Theoretical and Computational Sciences (NITheCS), Stellenbosch, South Africa*

<sup>3</sup>*School of Data Science and Computational Thinking, Stellenbosch University, Stellenbosch, South Africa*

Analogue memristive crossbars are a leading route to energy-efficient learning hardware, and Portner *et al.* recently ran a valence-change-memristor crossbar as a full actor–critic reinforcement-learning agent that learns reward-driven navigation [1]. We derive a continuous-variable (CV) photonic memristor that carries this idea into the quantum-optical domain through a new mechanism. A coherent beam is weakly tapped and read by homodyne detection, the photocurrent drives a fading-memory integrator, and the integrator state sets an electro-optic phase that rotates the transmitted field into a cross-quadrature output. The memory is therefore held in a lossless optical phase while the weak measurement serves only as a sensor for a separate phase actuator, a sense-and-feed-forward, phase-encoded design distinct from the reflectivity-encoded quantum memristors demonstrated to date [2, 3]. The device reproduces the three canonical memristor fingerprints, pinched hysteresis, a frequency-driven collapse of the loop area that peaks at the memory rate, and exponential retention, under both a coherent drive and a nonclassical squeezed-vacuum drive whose minimum quadrature variance dips below the vacuum.

To establish that the device learns, we place it in the same actor–critic scaffold and run it on the navigation tasks used for the classical memristor, with its internal memory states serving as the agent’s trainable weights. Over twenty seeds it is statistically indistinguishable from the valence-change-memristor baseline, reaching an asymptotic  $6.80 \pm 0.11$  steps against  $6.76 \pm 0.10$  on the T-maze and  $14.18 \pm 0.76$  against  $14.32 \pm 1.02$  on the water maze. The device’s volatile dynamics also do genuine work in the learning: on a non-stationary stream its physical retention time can be tuned to the task timescale, and this matched forgetting beats a frozen memory, so the volatility that sets the device apart from an ideal nonvolatile memristor becomes a learning resource. Two optical knobs then carry into the quantum study that follows, the weak-measurement strength that interpolates continuously between quantum and classical operation, and a squeezed-vacuum injection that lowers the readout-noise floor below the shot-noise limit [4].

Because every element is a genuine quantum-optical component, we then run the device on streams of quantum states to elucidate which of its behaviours are genuinely quantum, taking the boundary between efficiently simulable Gaussian dynamics and non-Gaussian dynamics as the dividing line [5]. The sharpest result is a separation in measurement access: reading one copy at a time, the device inherits a copy cost that grows as  $2^{n/2}$  to read a global property of an unknown state [6, 7], whereas the same substrate extended with a two-copy interference readout, the CV SWAP test [8], decides with a constant number of copies, and we derive in closed form how efficient the photon detectors must be for the advantage to survive. A companion result has the device learn online, inside its own loop, a photon-number-parity feature that no first- or second-moment strategy can reach at any measurement budget. Every instance is classically simulated and we claim no computational speed-up; the contribution is a single learning device that performs well on the classical navigation task and on quantum data alike, with a quantified map of where each quantum advantage washes out under loss.

## References

- [1] K. Portner *et al.*, “Actor–critic networks with analogue memristors mimicking reward-based learning,” *Nat. Mach. Intell.* **7**, 1939–1953 (2025).
- [2] P. Pfeiffer, I. L. Egusquiza, M. Di Ventra, M. Sanz, and E. Solano, “Quantum memristors,” *Sci. Rep.* **6**, 29507 (2016).
- [3] M. Sanz, L. Lamata, and E. Solano, “Quantum memristors in quantum photonics,” *APL Photonics* **3**, 080801 (2018).
- [4] C. M. Caves, “Quantum-mechanical noise in an interferometer,” *Phys. Rev. D* **23**, 1693–1708 (1981).
- [5] A. Mari and J. Eisert, “Positive Wigner functions render classical simulation of quantum computation efficient,” *Phys. Rev. Lett.* **109**, 230503 (2012).
- [6] S. Chen, J. Cotler, H.-Y. Huang, and J. Li, “Exponential separations between learning with and without quantum memory,” in *2021 IEEE 62nd Annual Symp. Found. Comput. Sci. (FOCS)* (2021), arXiv:2111.05881.
- [7] H.-Y. Huang *et al.*, “Quantum advantage in learning from experiments,” *Science* **376**, 1182 (2022).
- [8] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, “Direct estimations of linear and nonlinear functionals of a quantum state,” *Phys. Rev. Lett.* **88**, 217901 (2002).