

Reinforcement Learning for Optimal Control in Quantum-Computer-Enhanced Electron Microscopy

Sebastian Egginger ¹, Ellena Pflieger¹, and Johannes Kofler ¹

¹*Department of Quantum Information and Computation at Kepler (QUICK),
Johannes Kepler University Linz, Linz, Austria*

Most imaging ends in an intensity-based detector that measures each particle individually. For particles small enough to exhibit measurable quantum behavior, such as photons or electrons, this collapses the state of every particle in isolation, ruling out entanglement or collective measurements across them – precisely the resources by which quantum metrology beats classical precision limits. Using programmable quantum sensors that can make use of state amplitudes rather than click probabilities is an active field of research. Advantages may arise from joint measurements [1], coherently summing phase information [2, 3], or processing with sophisticated quantum algorithms [4–6].

Transmission electron microscopy (TEM) naturally benefits from these advances [7]. Our project partners are building an ion trap inside a TEM to interact with electrons beyond standard detection [8]. Trapped ions are well established as both quantum computers [9] and sensors [10]. We develop the control layer for this setup, but our approach applies to any quantum computer used as a sensor [11].

I. PROBLEM AND APPROACH

We work in a standard metrology setting. A field $f(j, \theta)$ imprints a phase on each of n sensors. The index $j = 1, \dots, n$ lets the field vary spatially across the sensor array, while θ is the parameter of interest corresponding to, e.g., a magnetic-field strength, a clock misalignment, or a specimen property in a TEM. We cast the sensing process as a controllable circuit shown in Fig. 1(A): an initialization $U_I(\mathbf{x})$ prepares the probe, the encoding $U_E(\theta)$ writes θ into single-qubit rotation angles $f(j, \theta)$, and a decoding $U_D(\mathbf{x})$ rotates into the optimal measurement basis. The vector $\mathbf{x}$ collects

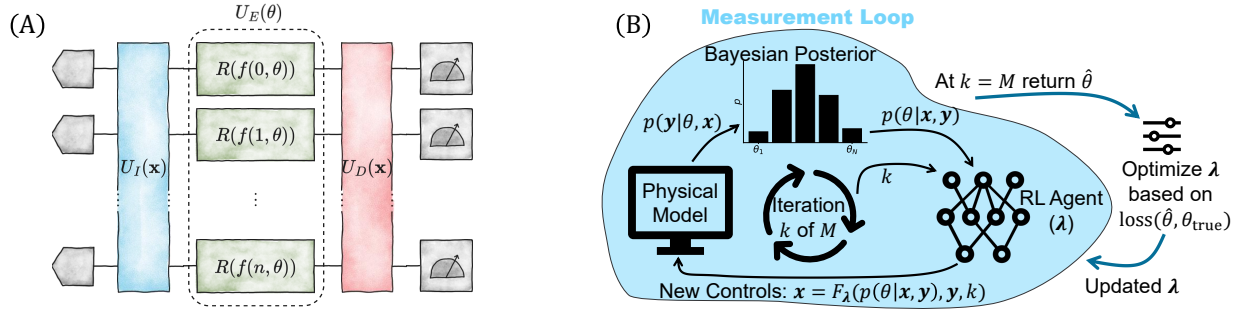


Figure 1. (A): *Sensing process as a quantum circuit.* By choosing the optimal $\mathbf{x}$ for an initialization Ansatz in $U_I(\mathbf{x})$, the ideal state should be prepared. That state then evolves in $U_E(\theta)$ encoding the unknown parameter θ . Lastly, in the decoding $U_D(\mathbf{x})$, the ideal measurement should be chosen. (B): *Schematics of the optimization procedure.* The agent outputs controls for each measurement. Using a physical model, those can be transformed into probabilities for outcomes $\mathbf{y}$. On that basis, Bayesian updates are performed to update our estimates $\hat{\theta}$ of θ . Those define a loss during training. See the main text for further explanations.

all free control parameters.

Locally, the (quantum) Fisher information (FI) is maximized by GHZ-type probes, which reach the Heisenberg limit (HL) by summing the n phases. The resulting n -times-faster signal then demands $\propto 1/n$ prior knowledge to resolve ambiguity (periodicity) [2, 3, 12]. Lacking such a prior, a static strategy cannot reach HL precision and is closer to the standard quantum limit (SQL). Our results show that an adaptive strategy recovers the HL scaling by spending some measurements to resolve the ambiguity on θ and then moving to higher precision. This is global metrology, where optimal schemes are analytically hard to obtain and are typically found numerically via Bayesian updating and reinforcement learning (RL) [13–17].

We build on the model-aware framework of *Belliardo et al.* [18], in which a differentiable particle filter and physical model make the controls trainable by gradient descent. We made multiple changes and improvements to it, with our main extension replacing their fixed proof-of-concept tasks [19] with arbitrary differentiable circuits (Fig. 1(A)). We also used circular statistics to re-engineer the estimator for the circular parameter domain, so the full range is handled without boundary artifacts. The loop itself is shown in Fig. 1(B): the controls yield an outcome $\mathbf{y}$, a discretized posterior over θ is sharpened by Bayesian updates, and its summary is input to the RL agent. The training loss compares the final estimate $\hat{\theta}$ to the (only during training) known θ_{true} . In related work [20], we recast this pipeline as a special type of Markov decision process and revisit the loss function based on classical RL literature.

II. RESULTS

We demonstrate the pipeline on the canonical uniform-phase task, $f(j, \theta) = \theta$, which offers the largest metrological advantage [3]. Initialization and decoding use a Quantum Neural Network (QNN) ansatz of single- and two-qubit rotations in an XZX pattern, which is expressive enough for Hadamard, GHZ, and Quantum Fourier Transform (QFT) primitives. The agent is a small feedforward neural network.

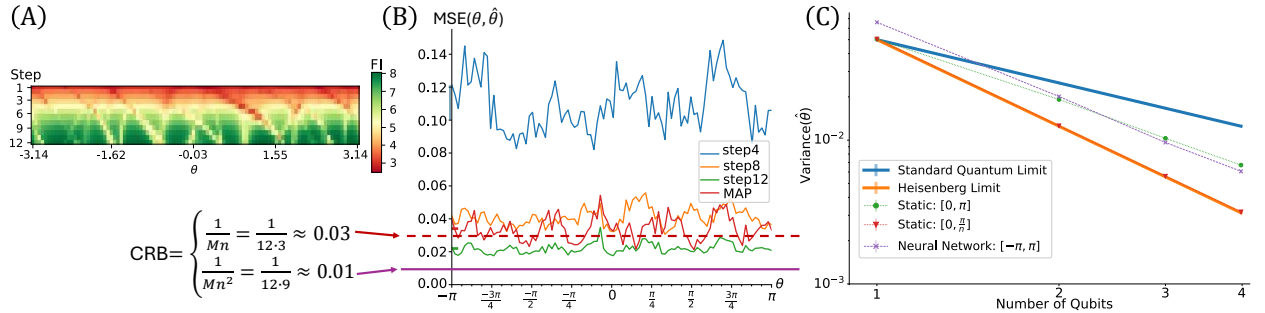


Figure 2. (A): The FI increases over the number of (up to $M = 12$) iteration steps, because the adaptive strategy first identifies the correct parameter region without entanglement (SQL: FI = 3 due to $n = 3$ qubits) and then leverages entanglement (HL: FI = 9) without periodicity ambiguity. We find small variations as θ varies. (B): The MSE decreases with every step, surpassing the SQL, but remaining above the HL. This is expected and almost independent of θ . (C): Our adaptive neural network RL strategy scales like the HL ($\frac{1}{n^2}$) with an offset due to a lack of prior knowledge. With such a prior (red), static methods also reach the HL, but without them (green) are closer to the SQL ($\frac{1}{n}$).

Figure 2 summarizes the result when we train over the full range $\theta \in [-\pi, \pi]$ with $n = 3$ sensors and $M = 12$ measurement steps. (A) Across the measurement loop, the per-step FI grows from the standard-quantum-limit value ($n = 3$) toward the Heisenberg value ($n^2 = 9$) as the adaptive controls localize θ and gradually increase entanglement, with no periodicity ambiguity. (B) Accordingly, in the evaluation, the (periodic) mean squared error (MSE) drops below the standard quantum limit and approaches – but stays above – the HL. This error is defined as $\text{MSE}(\theta, \hat{\theta}) = \frac{1}{N} \sum_{j=1}^N \min_{\delta \in \{-\pi, 0, \pi\}} (\hat{\theta}_j - \theta + \delta)^2$, where N is the batch size of equal, but stochastic runs and the minimum over δ is required by circular statistics. To put this MSE in context, the FI is inverted to acquire the Cramér-Rao bound (CRB). (C) As a function of qubit number, the variance of our adaptive strategy follows the Heisenberg scaling $\propto 1/n^2$, offset due to the missing prior. Static strategies reach the HL only when given a matching prior (red, promise that θ is between 0 and $\frac{\pi}{n}$) and otherwise sit near the SQL (green, θ is between 0 and π). In [11], we perform similar experiments for other known field patterns like $f(j, \theta) = 2^j \theta$, which corresponds to Quantum Phase Estimation (QPE), and we analyze our circuits in the context of QFT. We also analyze novel field patterns like $f(j, \theta) = \sin(j\theta)$, which correspond to estimating the frequency of spatially oscillating fields. Those, in turn, may stem from a specimen that resembles a grating. We find that this method is well-suited to address all these tasks, with performances that vary with the true value of θ .

Thus, learned adaptive control achieves HL-like scaling over the full parameter range without prior promise of the range of the unknown parameter, on a task expressed entirely as a quantum circuit. This ability is crucial to transform a quantum computer into a sensor. It also demonstrates how machine learning techniques can push experimental quantum information science.

-
- [1] H.-Y. Huang, M. Broughton, J. Cotler, S. Chen, J. Li, M. Mohseni, H. Neven, R. Babbush, R. Kueng, J. Preskill, and J. R. McClean, [Science](#) **376**, 1182 (2022).
 - [2] V. Giovannetti, S. Lloyd, and L. Maccone, [Phys. Rev. Lett.](#) **96**, 010401 (2006).
 - [3] T. J. Proctor, P. A. Knott, and J. A. Dunningham, [Phys. Rev. Lett.](#) **120**, 080501 (2018).
 - [4] A. Mokeev, B. Saif, M. D. Lukin, and J. Borregaard, [PRX Quantum](#) **7**, 010318 (2026).
 - [5] S. A. Khan, S. Prabhu, L. G. Wright, and P. L. McMahon, [npj Quantum Information](#) **10.1038/s41534-026-01235-w** (2026).
 - [6] R. R. Allen, F. Machado, I. L. Chuang, H.-Y. Huang, and S. Choi, [Quantum computing enhanced sensing](#) (2025), [arXiv:2501.07625 \[quant-ph\]](#).
 - [7] H. Okamoto and Y. Nagatani, [Applied Physics Letters](#) **104**, 062604 (2014).
 - [8] E. Pescoller, S. Beltrán-Romero, S. Egginger, N. Jungwirth, M. Zanetti, D. Hornof, M. S. Seifner, I. Březinová, P. Haslinger, T. Juffmann, J. Kofler, P. Schindler, and D. Rätzel, [Coupling free electrons to a trapped-ion quantum computer](#) (2026), [arXiv:2601.11446 \[quant-ph\]](#).
 - [9] A. Ransford, M. S. Allman, J. Arkinstall, *et al.*, [Nature](#) **10.1038/s41586-026-10676-4** (2026).
 - [10] C. D. Marciniak, T. Feldker, I. Pogorelov, R. Kaubruegger, D. V. Vasilyev, R. van Bijnen, P. Schindler, P. Zoller, R. Blatt, and T. Monz, [Nature](#) **603**, 604 (2022).
 - [11] S. Egginger, E. Pflieger, and J. Kofler, Optimized global sensing with quantum computers (in preparation).
 - [12] D. V. Vasilyev, A. Shankar, R. Kaubruegger, and P. Zoller, [Optimal multiparameter metrology: The quantum compass solution](#) (2024), [arXiv:2404.14194 \[quant-ph\]](#).
 - [13] A. Hentschel and B. C. Sanders, [Phys. Rev. Lett.](#) **107**, 233601 (2011).
 - [14] T. Fösel, P. Tighineanu, T. Weiss, and F. Marquardt, [Phys. Rev. X](#) **8**, 031084 (2018).

- [15] L. J. Fiderer, J. Schuff, and D. Braun, [PRX Quantum](#) **2**, 020303 (2021).
- [16] V. Cimini, M. Valeri, E. Polino, S. Piacentini, F. Ceccarelli, G. Corrielli, N. Spagnolo, R. Osellame, and F. Sciarrino, [Advanced Photonics](#) **5**, 016005 (2023).
- [17] V. Cimini, E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, [Quantum Science and Technology](#) **9**, 035035 (2024).
- [18] F. Belliardo, F. Zoratti, F. Marquardt, and V. Giovannetti, [Quantum](#) **8**, 1555 (2024).
- [19] F. Belliardo, F. Zoratti, and V. Giovannetti, [Phys. Rev. A](#) **109**, 062609 (2024).
- [20] E. Pfleger, S. Egginger, and J. Kofler, Rethinking adaptive Bayesian approaches as Belief-MDPs in quantum sensing (in preparation).