

LLM-based Multi-Agent Copilot for Quantum Sensor

Rong Sha^{1,2†}, Binglin Wang^{1†}, Jun Yang^{1,2*†}, Xiaoxiao Ma^{1,2†}, Chengkun Wu¹, Liang Yan¹, Chao Zhou^{1,2}, Jixun Liu^{1,2}, Guochao Wang^{1,2}, Shuhua Yan^{1,2}, Lingxiao Zhu^{1,2*}

¹National University of Defense Technology, Changsha, 410073 China.

²National Key Laboratory of Equipment State Sensing and Smart Support, National University of Defense Technology, Changsha, 410073 China.

*Corresponding author(s). E-mail(s): jyang@nudt.edu.cn;
lzx334@nudt.edu.cn;

†These authors contributed equally to this work.

Abstract

Large language models (LLM) exhibit broad utility but face limitations in quantum sensor development, stemming from interdisciplinary knowledge barriers and involving complex optimization processes. Here we present QCopilot, an LLM-based multi-agent framework integrating external knowledge access, active learning, and uncertainty quantification for quantum sensor design and diagnosis. Comprising commercial LLMs with few-shot prompt engineering and vector knowledge base, QCopilot employs specialized agents to adaptively select optimization methods, automate modeling analysis, and independently perform problem diagnosis. Applying QCopilot to atom cooling experiments, we generated 10^8 sub- μK atoms without any human intervention within a few hours, representing $\sim 100\times$ speedup over manual experimentation. Notably, by continuously accumulating prior knowledge and enabling dynamic modeling, QCopilot can autonomously identify anomalous parameters in multi-parameter experimental settings. Our work reduces barriers to large-scale quantum sensor deployment and readily extends to other quantum information systems.

1 Introduction

Atom cooling, trapping and quantum manipulation has been one of the leading platforms in quantum sensing in the last decades[1–8]. Cold-atoms based quantum sensors are increasingly transitioning from controlled laboratory environments to real-world deployments[9–18]. However, such systems—integrating laser subsystems, optical components, electronic control modules, and firmware—exhibit high complexity in their operational dynamics[19]. Furthermore, cold atoms, as the core sensing medium, display extreme susceptibility to thermal perturbations, mechanical vibrations, and electromagnetic noise[20]. This confluence of complexity and environmental sensitivity elevates deployment costs and degrades operational robustness. Developing cold-atoms based devices capable of self-optimization and self-diagnosis during prolonged field operation thus remains a critical bottleneck. Existing strategies fall into three broad categories: control-based methods (e.g., PID control, robust control[21]), black-box optimization techniques (e.g., Bayesian optimization[22–25], genetic algorithms[26], deep neural networks[27]), and reinforcement learning (RL)-based agents[28–32]. Control-based approaches offer strong interpretability via mathematical models but falter when models are inaccurate or systems exhibit nonlinearity—a common scenario in real-world experiments. Black-box methods, while model-agnostic, suffer from limited generalization due to data quality dependencies and domain specificity[33]. RL-based agents enable autonomous learning but rely on high-quality offline datasets and human-designed reward functions, stumbling over the Polanyi Paradox—the gap between tacit knowledge and explicit articulation. A unifying limitation across these methods is their inability to bridge knowledge barriers, hindering scalable utility and cumulative knowledge growth, and creating a critical trade-off between generalization and domain adaptability.

Large language model (LLM) offers a transformative solution to these challenges. Pre-trained on vast datasets, LLM internalizes cross-disciplinary knowledge, exhibits strong few-shot and zero-shot generalization, and excels in reasoning, planning, and contextual understanding[34]. Unlike RL agents, LLM-based agents operate via text instructions, prompt templates, and in-context learning, enabling flexible, human-like interaction[35, 36]. When integrated into multi-agent systems, LLMs foster collective intelligence, as demonstrated in software development[37], multi-robot planning[38], self-driving experiment optimization[39–44], and scientific discovery[45–47]. However, existing LLM-based frameworks lack integration across real-time monitoring, experimental optimization and fault diagnosis—critical for quantum experiments, which are inherently uncertain.

To address these gaps, we introduce QCopilot, a centralized multi-agent framework (Fig. 1d1) that combines LLM (e.g., Qwen3), vector knowledge bases, active learning, and uncertainty quantification. This framework enables bidirectional functionality: forward optimization of experimental systems and reverse diagnosis of anomalies. As shown in Fig. 1d2, QCopilot orchestrates specialized agents—Decision Maker, Experimenter, Analyst, Multimodal Diagnoser, Web Searcher, and Recorder—to decompose tasks, automate optimization, quantify uncertainties, and diagnose faults. By synergistically integrating its core components, the framework effectively breaks down knowledge barriers, leverages natural language-based prior knowledge to enhance

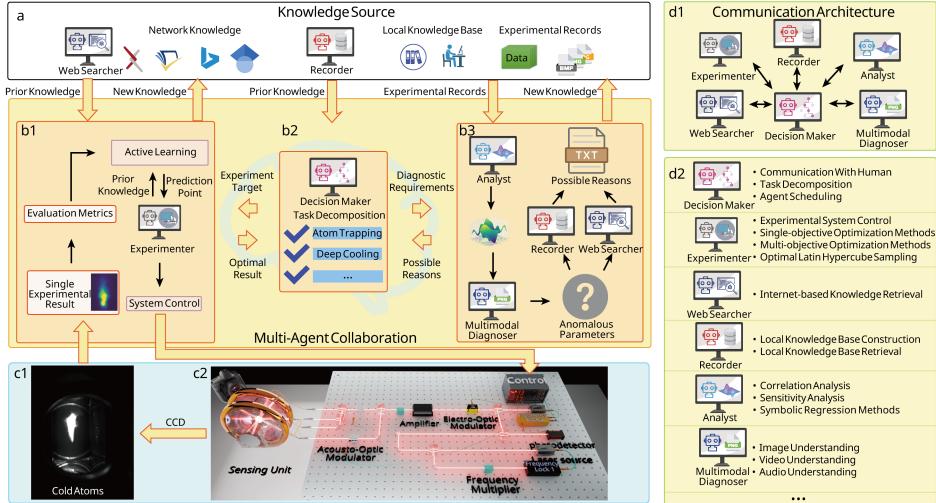


Fig. 1 Architecture of QCopilot **a**, The system employs the Web Searcher and Recorder to construct a centralized knowledge source, integrating both historical experimental data and external information. **b2**, During task decomposition, the Decision Maker utilizes prior knowledge from this centralized source to decompose complex tasks and determine optimal agent invocation workflows. **b1**, Experimental optimization workflow: The Experimenter optimizes experimental parameters through active learning techniques, driving autonomous system tuning. **b3**, Fault diagnosis workflow: The Analyst models system behavior, the Multimodal Diagnoser performs image-based anomaly analysis, and the Recorder (with the Web Searcher) retrieves potential root causes to enable targeted troubleshooting. **c1**, Cold-atom fluorescence image captured by charge coupled device (CCD). **c2**, Schematic of the experimental setup, illustrating key components of the cold-atom platform. **d1**, The framework adopts a centralized agent communication architecture to coordinate inter-agent interactions. **d2**, Inventory of existing agents, highlighting the framework's modular design that supports flexible integration of additional agents as needed.

decision-making, and continuously refines and accumulates knowledge, as illustrated in Fig. 1a and b1-b3. During the experimental optimization stage that shown in Fig. 1b1, Experimenter directly interacts with a cold atoms experimental platform (Fig. 1c1-c2) and employs active learning techniques to optimize experimental parameters. Upon completion of the optimization, updated experimental records and associated knowledge are integrated into the knowledge source. Experimental validation in atomic cooling demonstrates the generation of 10^8 sub- μK atoms within several hours, achieving a \sim 100-fold speedup over manual experimentation. In the problem diagnosis stage that displayed in Fig. 1b3, Analyst processes experimental records, which are transformed into visual data and forwarded to the Multimodal Diagnoser to identify anomalous parameters. Working in collaboration with the Recorder and Web Searcher, the system infers potential root causes, addresses the issues, and subsequently updates the knowledge base with newly acquired insights. This framework is capable of autonomously detecting anomalous parameters in multi-parameter experimental environments. As a result, QCopilot accelerates the automated preparation of cold atoms workflows, enhances generalization to environmental variations, and paves the way for the development of cost-effective and reliable cold-atoms based quantum sensors in both academic and industrial settings.

2 Results

In cold-atoms based quantum sensors, dense cold-atoms source at sub- μK temperatures is typically required for improving sensing precision and sensitivity. To achieve this, the experimental process is divided into two stages: capturing and cooling the atomic ensemble to the sub-mK level using a magneto-optical trap (MOT), followed by further reducing the temperature to the μK level through sub-Doppler cooling. As illustrated in Fig. 1c, our experimental platform combines commercially available lasers with self-developed control systems. First, we stabilize the laser frequency using modulation transfer spectroscopy and optical phase-locked loop. Then, three pairs of circularly polarized orthogonal lasers, along with a gradient magnetic field, form a MOT for cooling and trapping rubidium atoms within the vacuum chamber. In our system, polarization gradient cooling (PGC) serves as the primary mechanism for sub-Doppler cooling. The advantage of this approach is its simplicity, as it does not require any additional optical or electronic components. After loading the MOT, the magnetic field is turned off after 1s, and the laser frequency and power are adjusted accordingly. To determine the total number of atoms in the MOT, we use pixel integration from charge coupled device (CCD) images, as shown in Fig. 2b. This step aims to acquire a dense atomic cloud. During the PGC stage, we employ the molasses diffusion method to measure the temperature and atom number. Representative images are shown in Fig. 2e, while the fitted results of the atomic cloud diffusion radius are presented below. Our objective is to obtain a cold atomic ensemble with plenty of atoms.

The whole experiments are subject to significant uncertainties. Controllable variables, such as optical parameters, dictate the upper limit of optimization during the tuning process, while uncontrollable variables, such as environmental factors, restrict the system's generalization capabilities. Given the high-dimensional and sparse nature of the parameters, combined with the substantial experimental costs, researchers must carefully balance cost and performance. This often renders parameter selection heavily reliant on prior experimental experience. Furthermore, manual optimization is limited in its ability to simultaneously consider multiple objectives, significantly increasing the complexity of the process. The intricate experimental procedure and time-consuming individual sub-experiments optimizations limit the total number of trials that can be performed. The quality of the atomic cloud produced in the cooling stage directly influences key outcomes such as the signal-to-noise ratio of interference signals.

2.1 Adaptive optimization

For experimental optimization, we provide the Decision Maker with instructions regarding the list of sub-experiments, their order, and specific requirements. The Decision Maker calls the Experimenter sequentially according to the sub-experiment order. Based on the requirements of different sub-experiments, the Experimenter automatically selects appropriate optimization methods and sets experimental parameters using retrieval-augmented generation to retrieve hardware parameter reports from a knowledge base, without requiring prior physical knowledge. From the dialogue history as shown in Fig. 2a, the Experimenter employs Bayesian optimization (BO) based on batch Log Expected Improvement to efficiently search the high-dimensional parameter

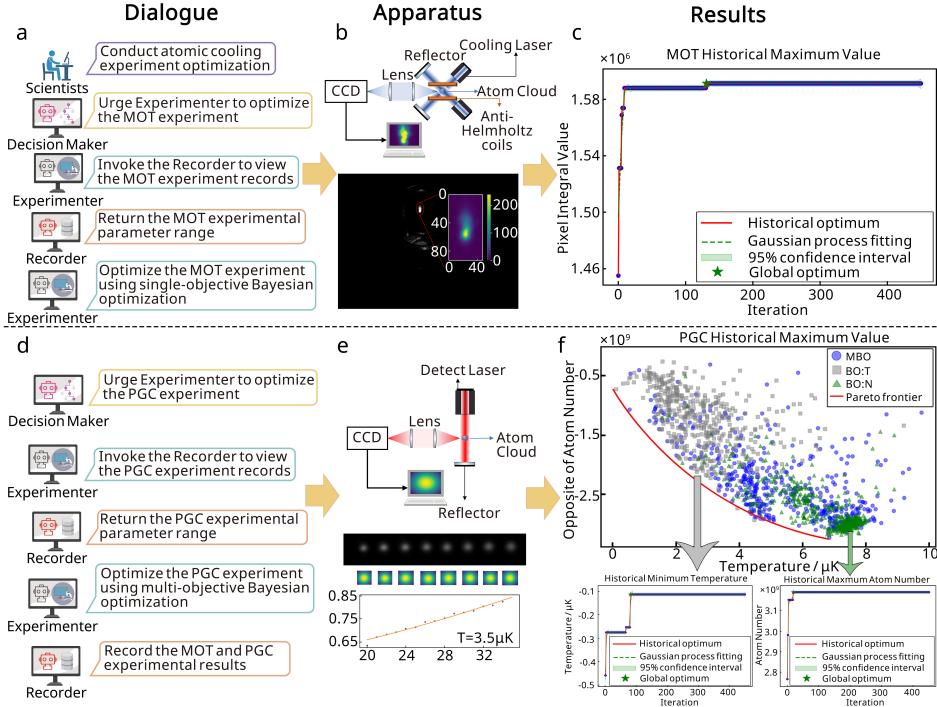


Fig. 2 Optimization of atomic cooling experiments via QCopilot **a**, Workflow for optimizing the magneto-optical trap (MOT) via the QCopilot framework. **d**, Corresponding workflow for polarization gradient cooling (PGC) optimization. **b**, Experimental procedure for quantifying MOT atom number using CCD image pixel integration. **c**, Evolution of Bayesian optimization (BO) performance in MOT optimization, tracking five key parameters, converges to optimal values within ~100 iterations. **e**, Protocol for measuring atomic temperature via the molasses diffusion method. **f**, Multi-objective optimization results for PGC: main panel shows Pareto frontier (red line) identified by multi-objective Bayesian optimization (MBO, blue points) across 500 iterations, targeting six parameters. Insets display historical optimization trends for minimizing temperature (gray squares) and maximizing atom number (green triangles) via single-objective BO. MBO efficiently balances both objectives, yielding sub- μK atoms with high density, whereas single-objective approaches prioritize one metric at the expense of the other.

space for the MOT experiment. To enhance robustness, each parameter setting during optimization was tested three times, with the average value recorded. Fig. 2c shows the historical optimal values across iterations. The ordinate is obtained by performing pixel integration on the CCD image, and the pixel integration value is proportional to the number of atoms. Optimization converged to the optimum in approximately 100 iterations.

Having determined the best parameters for the MOT, the Decision Maker invokes the Experimenter to fix these parameters for subsequent experiments. For the PGC optimization, the Experimenter uses multi-objective Bayesian optimization (MBO) based on parallel Expected Hypervolume Improvement to identify the Pareto frontier, as displayed in Fig. 2d. Since interferometric measurements generally require cold atomic ensembles below 10 μK , only data points representing temperatures below this

threshold are depicted in Fig. 2f. Each comparison experiment consisted of 500 setups, yielding acceptable results (below 10 μK) as follows: 477 Blue circles represent experimental points guided by MBO, while the red line indicate the Pareto frontier deduced by MBO. 463 Gray squares correspond to results achieved when the Decision Maker optimized solely for temperature using single-objective BO. 489 Green triangles depict results when the Decision Maker optimized solely for atom number. Notably, BO optimizes for the maximum, so the temperature was multiplied by -1 to align with the optimization goal. Conversely, in MBO, all objectives were minimized, necessitating multiplication of the atom number by -1. As illustrated in Fig. 2f, points near the origin, which satisfy both low temperature and high atom number, are concentrated due to MBO. Optimization focuses solely on temperature yielded points in regions of low temperature but insufficient atom density, whereas optimization focuses on atom number produced points with high atom number but relatively higher temperature. MBO targeting six parameters across 500 iterations efficiently identifies the Pareto frontier, which demonstrates the superior performance of MBO in meeting experimental requirements. All three methods demonstrated efficient searching, effectively balancing experimental cost and system performance.

2.2 Autonomous analysis

As illustrated in Figs. 3a-c, we perform correlation analysis, a typical method of uncertainty quantification, to analyze the correlation of parameters in the optimized experimental data from the MOT. Fig. 3a indicates that the number of atoms is strongly correlated with both the gradient magnetic field voltage X_5 and the cooling light intensity $I_{Cool}^{MOT}(X_3)$. In combination with Figs. 3b and c, it is evident that the parameter distributions for these two variables approach the upper limits of their respective ranges. The ranges are determined by the experimental hardware system. According to atomic cooling theory[48], the force experienced by an atom in the MOT can be given by:

$$F_{MOT} = -4\hbar k^2 \frac{I_{Cool}^{MOT}}{I_{sat}} \frac{-2\delta/\Gamma}{\left[1 + (2\delta/\Gamma)^2\right]^2} v - 4\hbar k^2 \frac{I_{Cool}^{MOT}}{I_{sat}} \frac{-2\delta/\Gamma}{\left[1 + (2\delta/\Gamma)^2\right]^2} \frac{g\mu_B}{k\hbar} \frac{dB}{dz} z, \quad (1)$$

where $\hbar k$ represents the photon momentum, Γ is the natural linewidth, I_{sat} denotes the saturation absorption intensity, and δ is the detuning, representing the difference between the laser frequency and the atomic resonance frequency. For transitions from the hyperfine level $|F, M_F\rangle$ to $|F', M_{F'}\rangle$, $g = g'_F M_{F'} - g_F M_F$. The quantity μ_B is the Bohr magneton, and dB/dz is the magnetic field gradient, which increases proportionally with the applied voltage of the gradient magnetic field X_5 . The Zeeman effect induces an imbalance in the radiation force, creating a restoring force. Under typical operating conditions, the atom experiences overdamped simple harmonic motion. Atoms entering the overlapping region of the laser beams are decelerated, and the position-dependent force drives cold atoms toward the center of the trap. According to the Eq. 1, it can be inferred that higher values of $I_{Cool}^{MOT}(X_3)$ lead to stronger damping

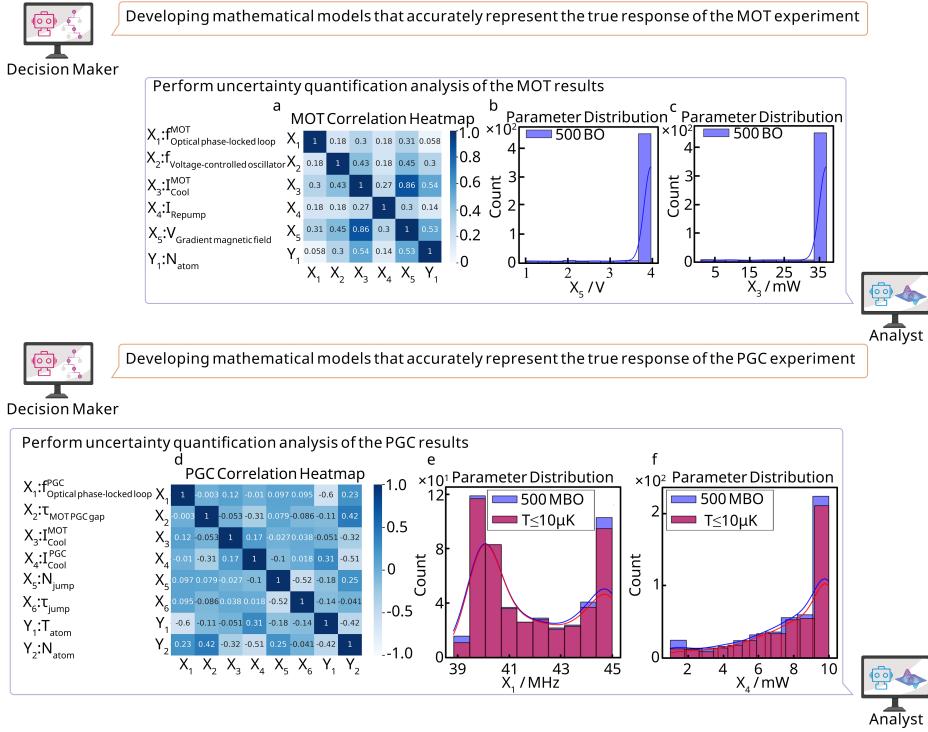


Fig. 3 Uncertainty quantification of the experimental system response The Analyst performed uncertainty quantification on 500 sets of magneto-optical trap (MOT) experimental results generated via Bayesian optimization. **a**, Correlation matrix quantifying the linear relationships between variables. **b,c**, Distributions of key parameters: voltage of the gradient magnetic field (**b**) and cooling light intensity (**c**). Subsequent analysis was conducted on 500 sets of polarization gradient cooling (PGC) results obtained from multi-objective Bayesian optimization. **d**, Correlation matrix characterizing parameter interactions in the PGC experiment. **e,f**, Distributions of the optical phase-locked loop reference frequency during PGC (**e**) and the cooling light intensity at the PGC endpoint (**f**), with probability density curves overlaid to highlight parameter ranges critical for sub- μK cooling performance.

effects, enabling more efficient cooling of atoms. Additionally, A larger X_5 generates a stronger magnetic field gradient, ultimately enhancing the trapping force. The stronger trapping force results in greater confinement of cold atoms at the MOT center. As illustrated in Fig. 3a, the correlation analysis results align well with established theoretical models.

Similarly, we perform a correlation analysis based on PGC experiment results to investigate the factors influencing both the temperature and atom number of the cold atomic ensemble. Fig. 3d presents the correlation matrix, which quantifies the relationships between various experimental parameters. Our focus is on identifying key parameters impacting temperature and atom number, as well as understanding the interplay between these two objectives. Specifically, the final atomic temperature is strongly affected by the frequency of high-detuned PGC and the endpoint of cooling light intensity attenuation. In addition, the atom number is influenced by the

re-lighting time interval between MOT and PGC, the initial power of the reactivated cooling light, and the number of steps in the power attenuation process. Furthermore, according to statistical mechanics principles, the temperature of the atomic ensemble depends on the velocity distribution of the atoms. A lower atom number often indicates that high-speed atoms have escaped, leaving only low-speed atoms in the ensemble, which explains the observed inverse relationship between temperature and atom number. Figs. 3e-f display parameter distributions along with probability density plots. The purple histograms and curves correspond to the full dataset of 500 experiments, while the red histograms and curves represent data points with temperatures below $10 \mu\text{K}$. Notably, the trends for both datasets are largely consistent, demonstrating that the Gaussian process regression model developed through MBO effectively characterizes the experimental system and accurately identifies parameter distributions meeting operational requirements. These findings enable further refinement of the optimization range, thereby enhancing the efficiency of parameter scanning.

2.3 Autonomous problem diagnosis

After the optimization experiment is completed, the system transitions to the self-sustained running stage. During this stage, all sub-experiments are executed sequentially according to the list order, utilizing the optimal parameters derived from the optimization process for parameter fixation and result preservation. If significant deviations in experimental results are detected, the framework enters the problem diagnosis stage. In this stage, sub-experiments are rerun with the same parameters, and their results, together with those from the self-sustained running stage, are submitted to a multimodal diagnostic system for analysis. Fig. 4a displays the MOT experimental results during the self-sustained running stage, whereas Fig. 4b shows the MOT experimental results during the problem diagnosis stage. The Multimodal Diagnoser, which possesses image analysis capabilities, identifies differences by comparing the two sets of images, thereby localizing the problematic sub-experiment.

After isolating the problematic sub-experiment, the Decision Maker instructs the Experimenter to conduct 50 times optimal Latin hypercube sampling. The resulting experimental data are subjected to correlation analysis. The Multimodal Diagnoser extracts correlation matrix values from the correlation analysis result graph and compares them with those obtained during the optimization stage to identify potentially problematic parameters. Fig. 4c displays the correlation analysis results from the optimization stage, while the remaining figures depict the results after artificial errors were introduced. The blue box highlights the correlation coefficient between X_3 and Y_1 , the yellow box highlights that between X_5 and Y_1 , and the red box highlights that between X_1 and Y_1 . In the first experiment, we fix the cooling light intensity (X_3) at 4 mW/cm^2 . Fig. 4d clearly shows that the originally strong correlation for cooling light intensity decreases by two orders of magnitude, while the other parameters exhibit little change. In the second set of experiments, we reduce the actual magnetic field voltage (X_5) to half its set value, drastically diminishing the trapping force. As shown in Fig. 4e, the correlations of most variables remain at the same order of magnitude, whereas the correlations for voltage-controlled oscillator frequency difference and gradient magnetic field voltage decrease by more than an order of magnitude,

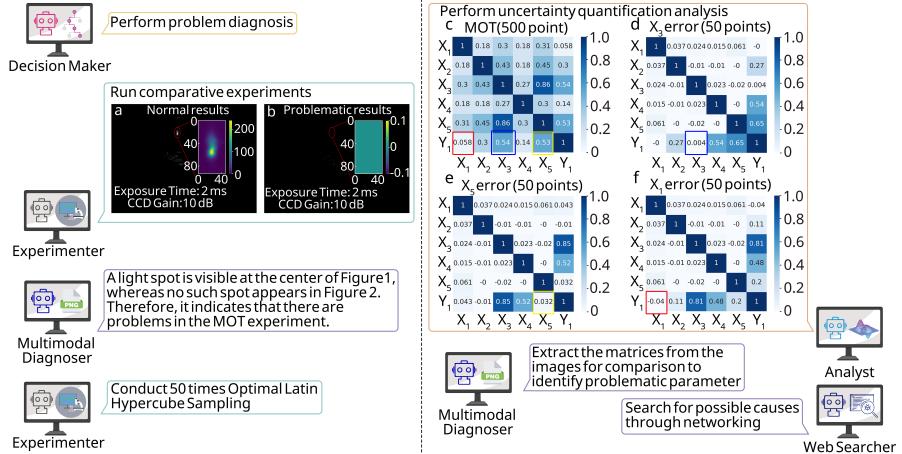


Fig. 4 Fault detection and diagnosis in the experimental system When operational anomalies occur, the Decision Maker initiates comparative experiments, and the Multimodal Diagnoser contrasts normal outcomes (**a**) with anomalous results (**b**) to pinpoint the faulty sub-experiment. For the identified sub-experiment, 50 iterations of optimal Latin hypercube sampling are executed, followed by uncertainty quantification analysis of the resulting data. **c**, Baseline correlation matrix derived from 500 Bayesian optimization runs during MOT optimization. **d-f**, Correlation matrices under specific fault conditions: cooling light intensity malfunction (**d**), gradient magnetic field voltage anomaly (**e**), and optical phase-locked loop reference frequency deviation (**f**). Parameters in **c-f** correspond to those defined in Fig. 3a. Through the Multimodal Diagnoser, the problematic parameter is precisely identified; the online retriever is then activated to analyze root causes, enabling efficient fault resolution.

thereby reducing the list of suspect parameters from five to two. In the third set of experiments, we fix the optical phase-locked loop reference frequency (X_1) to 38.6 MHz. Since this parameter is inherently weakly correlated, Fig. 4f confirms its minimal impact on the results. By leveraging the Multimodal Diagnoser, the problematic parameter is pinpointed, and the online retriever is invoked to analyze the root cause of the fault, enabling effective problem detection and diagnosis.

3 Discussion

To address the challenges of high construction costs and the difficulties associated with fault detection and diagnosis in modern quantum sensors, we propose a multi-disciplinary, coupled dynamic tuning framework built on a multi-agent system that leverages large language models. This framework is specifically designed for applications in the quantum domain. The proposed framework facilitates highly efficient autonomous operation across the entire workflow, including experimental optimization, problem detection, and diagnostic processes. By conducting only a few hundred experimental trials, the framework can identify solutions for single-objective optimization or determine the Pareto frontier for multi-objective optimization. Furthermore, after completing uncertainty quantification modeling, only 50 optimal Latin hypercube samples are required to accurately pinpoint critical problem parameters. To address

the limitations of existing methods in knowledge accumulation, the framework incorporates a local knowledge base and supports network-based retrieval, enabling effective alignment and incremental knowledge acquisition.

However, the framework currently depends on online API calls to access large language models, which restrict its offline applications. Fortunately, advances in localized inference models (e.g., qwq-32B) on consumer-grade hardware, such as a single 24GB GPU, are gradually bridging the performance gap with commercial large-scale models. By integrating further optimizations in prompt engineering, the framework holds the potential for local deployment on standard consumer laptops. This development could ultimately enable the autonomous operation of quantum sensors in field applications.

Supplementary information.

Acknowledgements. This work was supported by National Natural Science Foundation of China (Grant No. 12404556, 62103426) and Hunan provincial major sci-tech program (2023ZJ1010) and.

References

- [1] Wieman, C. E. Atom cooling, trapping, and quantum manipulation. *Rev. Mod. Phys.* **71**, S253–S262 (1999).
- [2] Becker, D. *et al.* Space-borne Bose–Einstein condensation for precision interferometry. *Nature* **562**, 391–395 (2018).
- [3] Schreck, F. & Druten, K. V. Laser cooling for quantum gases. *Nat. Phys.* **17**, 1296–1304 (2021).
- [4] Yuan, J. *et al.* Quantum sensing of microwave electric fields based on rydberg atoms. *Rep. Prog. Phys.* **86**, 106001 (2023).
- [5] Huang, J., Zhuang, M. & Lee, C. Entanglement-enhanced quantum metrology: From standard quantum limit to Heisenberg limit. *Appl. Phys. Rev.* **11**, 031302 (2024).
- [6] Salducci, C. *et al.* Quantum sensing of acceleration and rotation by interfering magnetically launched atoms. *Sci. Adv.* **10**, eadq4498 (2024).
- [7] Panda, C. D. *et al.* Measuring gravitational attraction with a lattice atom interferometer. *Nature* **631**, 515–520 (2024).
- [8] Ye, J. & Zoller, P. Essay: Quantum sensing with atomic, molecular, and optical platforms for fundamental physics. *Phys. Rev. Lett.* **132**, 190001 (2024).
- [9] Barrett, B. *et al.* Dual matter-wave inertial sensors in weightlessness. *Nat. Commun.* **7** (2016).

- [10] Bidel, Y. *et al.* Absolute marine gravimetry with matter-wave interferometry. *Nat. Commun.* **9**, 627 (2018).
- [11] Bongs, K. *et al.* Taking atom interferometric quantum sensors from the laboratory to real-world applications. *Nat. Rev. Phys.* **1**, 731–739 (2019).
- [12] Wu, X. *et al.* Gravity surveys using a mobile atom interferometer. *Sci. Adv.* **5**, eaax0800 (2019).
- [13] Stray, B. *et al.* Quantum sensing for gravity cartography. *Nature* **602**, 590–594 (2022).
- [14] d'Armagnac De Castanet, Q. *et al.* Atom interferometry at arbitrary orientations and rotation rates. *Nat. Commun.* **15**, 6406 (2024).
- [15] Deng, S. *et al.* Cold atom microwave clock based on intracavity cooling in China space station. *npj Microgravity* **10**, 66 (2024).
- [16] Zhai, C. *et al.* Airborne absolute gravity measurements based on quantum gravimeter. *Acta Phys. Sin.* **74**, 070302 (2025).
- [17] Li, C.-Y. *et al.* Drift-free continuous gravity measurement and application analysis of a high-precision atom gravimeter. *Phys. Rev. Appl.* **24**, 014045 (2025).
- [18] Li, J. *et al.* Realization of a cold atom gyroscope in space. *Natl. Sci. Rev.* **12**, nwaf012 (2025).
- [19] Degen, C. L., Reinhard, F. & Cappellaro, P. Quantum sensing. *Rev. Mod. Phys.* **89**, 035002 (2017).
- [20] Geraci, A. A. & Derevianko, A. Sensitivity of atom interferometry to ultralight scalar field dark matter. *Phys. Rev. Lett.* **117**, 261301 (2016).
- [21] Saywell, J. C. *et al.* Enhancing the sensitivity of atom-interferometric inertial sensors using robust control. *Nat. Commun.* **14**, 7626 (2023).
- [22] Wigley, P. B. *et al.* Fast machine-learning online optimization of ultra-cold-atom experiments. *Sci. Rep.* **6**, 25890 (2016).
- [23] Nakamura, I., Kanemura, A., Nakaso, T., Yamamoto, R. & Fukuhara, T. Non-standard trajectories found by machine learning for evaporative cooling of ^{87}Rb atoms. *Opt. Express* **27**, 20435 (2019).
- [24] Barker, A. J. *et al.* Applying machine learning optimization methods to the production of a quantum gas. *Mach. Learn.: Sci. Technol.* **1**, 015007 (2020).
- [25] Ma, X. *et al.* Accelerated bayesian optimization in deep cooling atoms (2025). [arXiv:2412.11793](https://arxiv.org/abs/2412.11793).

- [26] Weidner, C. A., Yu, H., Kosloff, R. & Anderson, D. Z. Atom interferometry using a shaken optical lattice. *Phys. Rev. A* **95**, 043624 (2017).
- [27] Tranter, A. D. *et al.* Multiparameter optimisation of a magneto-optical trap using deep learning. *Nat. Commun.* **9**, 4360 (2018).
- [28] Melnikov, A. A. *et al.* Active learning machine learns to create new quantum experiments. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 1221–1226 (2018).
- [29] Cimini, V. *et al.* Deep reinforcement learning for quantum multiparameter estimation. *Adv. Photonics* **5**, 016005 (2023).
- [30] Chih, L.-Y. & Holland, M. Reinforcement learning for rotation sensing with ultracold atoms in an optical lattice. *Phys. Rev. Res.* **6**, 043191 (2024).
- [31] Reinschmidt, M., Fortágh, J., Günther, A. & Volchkov, V. V. Reinforcement learning in cold atom experiments. *Nat. Commun.* **15**, 8532 (2024).
- [32] Liang, C. *et al.* Multi-parameter optimization of polarization gradient cooling for 87rb atoms based on reinforcement learning. *Opt. Express* **32**, 40364–40374 (2024).
- [33] Wang, H. *et al.* Scientific discovery in the age of artificial intelligence. *Nature* **620**, 47–60 (2023).
- [34] Singh, I. *et al.* ProgPrompt: program generation for situated robot task planning using large language models. *Auton. Robot* **47**, 999–1012 (2023).
- [35] Boiko, D. A., MacKnight, R., Kline, B. & Gomes, G. Autonomous chemical research with large language models. *Nature* **624**, 570–578 (2023).
- [36] Gao, S. *et al.* Empowering biomedical discovery with AI agents. *Cell* **187**, 6125–6151 (2024).
- [37] Dong, Y., Jiang, X., Jin, Z. & Li, G. Self-collaboration code generation via chatgpt. *ACM Trans. Softw. Eng. Methodol.* **33** (2024).
- [38] Chen, Y., Arkin, J., Zhang, Y., Roy, N. & Fan, C. Scalable multi-robot collaboration with large language models: Centralized or decentralized systems? (2024). Paper presented at the 2024 IEEE International Conference on Robotics and Automation (ICRA), 4311-4317 2024.
- [39] Sanders, L. M. *et al.* Biological research and self-driving labs in deep space supported by artificial intelligence. *Nat. Mach. Intell.* **5**, 208–219 (2023).
- [40] M. Bran, A. *et al.* Augmenting large language models with chemistry tools. *Nat. Mach. Intell.* **6**, 525–535 (2024).

- [41] Luo, X. *et al.* Large language models surpass human experts in predicting neuroscience results. *Nat. Hum. Behav.* (2024).
- [42] Ruan, Y. *et al.* An automatic end-to-end chemical synthesis development platform powered by large language models. *Nat. Commun.* **15**, 10160 (2024).
- [43] Jablonka, K. M., Schwaller, P., Ortega-Guerrero, A. & Smit, B. Leveraging large language models for predictive chemistry. *Nat. Mach. Intell.* **6**, 161–169 (2024).
- [44] Wu, Z. *et al.* Leveraging language model for advanced multiproperty molecular optimization via prompt engineering. *Nat. Mach. Intell.* 1359–1369 (2024).
- [45] Tshitoyan, V. *et al.* Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature* **571**, 95–98 (2019).
- [46] Melko, R. G. & Carrasquilla, J. Language models for quantum simulation. *Nat. Comput. Sci.* **4**, 11–18 (2024).
- [47] Ghafarollahi, A. & Buehler, M. J. Sciagents: Automating scientific discovery through bioinspired multi-agent intelligent graph reasoning. *Adv. Mater.* 2413523 (2024).
- [48] Foot, C. J. *Atomic Physics* No. 7 in Oxford Master Series in Physics Atomic, Optical, and Laser Physics (Oxford Univ. Press, Oxford, 2011).