

Optimal multi-parameter quantum metrology for frequencies of magnetic field

Zhenhua Long(龙振华) and Shengshi Pang(庞盛世)

Citation: Chin. Phys. B, 2025, 34 (**8**): 080301. DOI: 10.1088/1674-1056/add4e7
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Quantum-enhanced time-of-arrival measuring method based on partially entangled state

Peng-Xian Li(李芃鲜), Lu-Ping Xu(许录平), Gui-Ting Hu(胡桂廷), and Wen-Long Gao(高文珑)
Chin. Phys. B, 2025, 34 (**7**): 070303. DOI: 10.1088/1674-1056/adc7ee

Sudden change of interferometric power for X-shape states

Dian Zhu(朱典), Fu-Lin Zhang(张福林), and Jing-Ling Chen(陈景灵)
Chin. Phys. B, 2025, 34 (**2**): 020305. DOI: 10.1088/1674-1056/ad9732

Precision bounds for quantum phase estimation using two-mode squeezed Gaussian states

Jian-Dong Zhang(张建东), Chuang Li(李闯), Lili Hou(侯丽丽), and Shuai Wang(王帅)
Chin. Phys. B, 2025, 34 (**1**): 010304. DOI: 10.1088/1674-1056/ad8dc0

Enhancing quantum metrology for multiple frequencies of oscillating magnetic fields by quantum control

Xin Lei(雷昕), Jingyi Fan(范静怡), and Shengshi Pang(庞盛世)
Chin. Phys. B, 2024, 33 (**6**): 060304. DOI: 10.1088/1674-1056/ad3430

Micron-sized fiber diamond probe for quantum precision measurement of microwave magnetic field

Wen-Tao Lu(卢文韬), Sheng-Kai Xia(夏圣开), Ai-Qing Chen(陈爱庆), Kang-Hao He(何康浩), Zeng-Bo Xu(许增博), Yi-Han Chen(陈艺涵), Yang Wang(汪洋), Shi-Yu Ge(葛仕宇), Si-Han An(安思瀚), Jian-Fei Wu(吴建飞), Yi-Han Ma(马艺菡), and Guan-Xiang Du(杜关祥)
Chin. Phys. B, 2024, 33 (**8**): 080305. DOI: 10.1088/1674-1056/ad5321

Optimal multi-parameter quantum metrology for frequencies of magnetic field

Zhenhua Long(龙振华)[1] and Shengshi Pang(庞盛世)[1,2,†]

[1] *School of Physics, Sun Yat-sen University, Guangzhou* 510275, *China*
[2] *Hefei National Laboratory, University of Science and Technology of China, Hefei* 230088, *China*

(Received 18 March 2025; revised manuscript received 22 April 2025; accepted manuscript online 7 May 2025)

Multi-parameter quantum estimation has attracted considerable attention due to its broad applications. Due to the complexity of quantum dynamics, existing research places significant emphasis on estimating parameters in time-independent Hamiltonians. Here, our work makes an effort to explore multi-parameter estimation with time-dependent Hamiltonians. In particular, we focus on the discrimination of two close frequencies of a magnetic field by using a single qubit. We optimize the quantum controls by employing both traditional optimization methods and reinforcement learning to improve the precision for estimating the frequencies of the two magnetic fields. In addition to the estimation precision, we also evaluate the robustness of the optimization schemes against the shift of the control parameters. The results demonstrate that the hybrid reinforcement learning approach achieves the highest estimation precision, and exhibits superior robustness. Moreover, a fundamental challenge in multi-parameter quantum estimation stems from the incompatibility of the optimal control strategies for different parameters. We demonstrate that the hybrid control strategies derived through numerical optimization remain effective in enhancing the precision of multi-parameter estimation in spite of the incompatibilities, thereby mitigating incompatibilities between control strategies on the estimation precision. Finally, we investigate the trade-offs in estimation precision among different parameters for different scenarios, revealing the inherent challenges in balancing the optimization of multiple parameters simultaneously and providing insights into the fundamental distinction between quantum single-parameter estimation and multi-parameter estimation.

Keywords: quantum metrology, multi-parameter estimation, quantum control

PACS: 03.65.Ta, 03.67.–a

DOI: 10.1088/1674-1056/add4e7 **CSTR:** 32038.14.CPB.add4e7

1. Introduction

Metrology, rooted in the classical parameter estimation theory,[1] aims to enhance measurement precision and refine estimation methodologies, playing a pivotal role in physics and other scientific fields. The emergence of quantum metrology is a significant leap forward, offering immense potential to surpass the limitations of classical measurements by introducing quantum resources such as squeezed states and entangled states, and thereby markedly enhancing the precision of measurements.[2–4] In recent years, quantum metrology has undergone considerable development[5–9] and has been widely applied in a variety of areas such as atomic clocks,[10–12] atomic interferometers,[2,13] quantum biology,[14,15] quantum imaging,[16–19] quantum sensing.[4,20,21]

The primary goal of quantum metrology is to optimize the precision of estimating unknown parameters. The general approach of quantum metrology has been shown in Fig. 1. To maximize information extraction, one can optimize the initial state,[22–25] the evolution path, and the final measurement.[4,26–29] Recently, an increasing number of studies have recognized that optimizing the quantum parameterization process,[30,31] particularly by selecting appropriate control strategies, can also significantly improve the precision limits. Numerical methods,[32–34] like gradient ascent pulse engineering (GRAPE), are used but have limitations such as getting trapped in local optima. Machine learning, especially reinforcement learning (RL),[35] has shown promise in addressing complex quantum processes and high-dimensional systems,[36–38] and has been successfully applied to quantum parameter estimation.[39–41]



Fig. 1. The basic process of quantum metrology involves preparing the initial quantum state, allowing the probes to evolve to encode the parameters of interest, and extracting information about the parameters through measurements.

In the realm of parameter estimation, the Cramér–Rao bound[1,42] provides a theoretical limit on estimator variance, with the Fisher information dictating the inverse of the minimum variance achievable for the parameter estimators. Quantum metrology extends this with quantum Fisher information, providing a precision limit for parameter estimation.[22,43,44] Multi-parameter estimation is more complex,[45,46] with the Holevo bound establishing precision limits. Incompatibility is a critical issue in multi-parameter quantum metrology,[47,48]

†Corresponding author. E-mail: pangshsh@mail.sysu.edu.cn

© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

originating from the potential inconsistency between the optimal estimation strategies for distinct parameters, leading to a trade-off between the precisions of individual parameter estimates. Current research primarily focuses on mitigating these incompatibilities by optimizing the probe states[49] and the measurement schemes,[50,51] whereas the exploration of incompatibilities between control strategies remains much underdeveloped.

In quantum metrology, time-dependent Hamiltonians offer significant advantages for parameter estimation by enhancing estimation precision through quantum system dynamics. This is particularly useful in situations where the the time variation of the Hamiltonian allows for a more agile and adaptive estimation strategy.[30] However, analytically determining the optimal measurement can be challenging because of the increased complexity of the system's evolution under a time-dependent Hamiltonian. In this case, numerical methods become a more practical approach to finding the optimal measurement solutions, thereby enhancing the sensitivity of quantum metrology.[32,41]

In this work, we investigate the multi-parameter quantum estimation with time-dependent Hamiltonians by focusing on a qubit system in a magnetic field with two close frequencies in orthogonal directions. The goal is to enhance the precision of estimating the mean and the difference of the two frequencies. To achieve this, we introduce quantum control on the Hamiltonian and optimize the control to increase the estimation precision by both the traditional GRAPE method and the reinforcement learning. We further assess the robustness of these optimization techniques against the deviations of the control parameters from their optimal values. Through numerical optimization, we obtain joint control strategies that successfully enhance the precision of estimating the mean and difference of the magnetic field frequencies, thereby mitigating incompatibilities between control strategies. Finally, we explore the trade-offs in the estimation precision among different parameters in various scenarios, providing insights into the fundamental distinction between quantum single-parameter estimation and multi-parameter estimation.

This paper is organized as follows. In Section 2, we provide the theory of multi-parameter quantum estimation and present the multi-parameter estimation model, focusing on a qubit system where the magnetic field frequencies in two orthogonal directions are nearly equal in magnitude, and introduce various optimization methods, including the gradient ascent pulse engineering and the reinforcement learning. In Section 3, we conduct a detailed analysis of the results, including the optimal estimation precision, robustness, and trade-offs of our approach. Finally, the paper is concluded in Section 4.

2. Multiparameter quantum estimation and methods

2.1. Multiparameter quantum estimation

The central task of parameter estimation is to determine the estimators for the parameters of interest, establishing a mapping from measurement outcomes to the estimated parameters.[5] In practical scenarios, the unknown parameters are typically multidimensional, denoted as $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_d)$. These parameters are encoded in the quantum state $\rho_{\boldsymbol{\lambda}}$, followed by a set of positive-operator valued measurements (POVM), $\Pi = \{\Pi_k | \sum_k \Pi_k = \boldsymbol{I}, \Pi_k \geq 0\}$, on these states to obtain the outcome. According to the Born rule, $p(k|\boldsymbol{\lambda}) = \mathrm{Tr}[\rho_{\boldsymbol{\lambda}}\Pi_k]$, the probability of each measurement result is determined to estimate parameter information. To quantify estimation precision, the Cramér–Rao bound[1,42] constrains all possible mean squared error matrices of any unbiased estimator as

$$V \geq F_{\mathrm{cl}}^{-1}, \tag{1}$$

where the covariance matrix V is used to describe the precision of parameter estimation, with the entries

$$V_{\mu\nu} = \sum_k p(k|\boldsymbol{\lambda})(\tilde{\lambda_\mu} - \lambda_\mu)(\tilde{\lambda_\nu} - \lambda_\nu), \tag{2}$$

and F_{cl} represents the classical Fisher information, with elements

$$F_{\mu\nu} = \sum_k p(k|\boldsymbol{\lambda})(\partial_\mu \log p(k|\boldsymbol{\lambda}))(\partial_\nu \log p(k|\boldsymbol{\lambda})). \tag{3}$$

The inequality (1) provides an upper bound on estimation precision, especially for locally unbiased estimators.

When applying the classical Cramér–Rao bound to quantum parameter estimation, selecting different measurement bases within the quantum system leads to varying probability distributions of measurement outcomes, and consequently, different precision of the estimation results. By minimizing over all possible measurement bases, the quantum Cramér–Rao bound is defined to be a lower bound on the estimation error, satisfying

$$V \geq F_{\mathrm{cl}}^{-1} \geq Q^{-1}. \tag{4}$$

The quantum Cramér–Rao theorem provides an ultimate constraint, which depends on the geometric structure of the quantum statistical model and is independent of the specific measurement bases employed. The optimal quantum measurement is implemented through POVM, when the Fisher information reaches the maximum, *i.e.*, the quantum Fisher information.[43,44]

The most celebrated quantum Cramér–Rao bound is the symmetric logarithmic derivative (SLD) bound,[52,53] defined via the SLD operators L_μ^{S},

$$\partial_{\boldsymbol{\lambda}_\mu}\rho_{\boldsymbol{\lambda}} = \frac{L_\mu^{\mathrm{S}}\rho_{\boldsymbol{\lambda}} + \rho_{\boldsymbol{\lambda}}L_\mu^{\mathrm{S}}}{2}. \tag{5}$$

The elements of the corresponding quantum Fisher information matrix are given by

$$Q^{\mathrm{S}}_{\mu\nu} = \mathrm{Tr}\left[\rho_{\boldsymbol{\lambda}} \frac{L^{\mathrm{S}}_{\mu} L^{\mathrm{S}}_{\nu} + L^{\mathrm{S}}_{\nu} L^{\mathrm{S}}_{\mu}}{2}\right]. \tag{6}$$

For pure states, we have

$$\rho_{\boldsymbol{\lambda}} = |\psi_{\boldsymbol{\lambda}}(t)\rangle\langle\psi_{\boldsymbol{\lambda}}(t)|, \tag{7}$$

and the the SLD operator can be defined as

$$L_{\boldsymbol{\lambda}}(t) = 2\left(|\partial_{\boldsymbol{\lambda}}\psi_{\boldsymbol{\lambda}}(t)\rangle\langle\psi_{\boldsymbol{\lambda}}(t)| + |\psi_{\boldsymbol{\lambda}}(t)\rangle\langle\partial_{\boldsymbol{\lambda}}\psi_{\boldsymbol{\lambda}}(t)|\right). \tag{8}$$

The elements of the quantum Fisher information matrix can be expressed as

$$\begin{aligned} Q^{\mathrm{S}}_{\mu\nu} &= \mathrm{Tr}\left[\rho_{\boldsymbol{\lambda}} \frac{L^{\mathrm{S}}_{\mu} L^{\mathrm{S}}_{\nu} + L^{\mathrm{S}}_{\nu} L^{\mathrm{S}}_{\mu}}{2}\right] \\ &= 2(\langle\partial_{\mu}\psi_{\boldsymbol{\lambda}}(t)|\psi_{\boldsymbol{\lambda}}(t)\rangle\langle\partial_{\nu}\psi_{\boldsymbol{\lambda}}(t)|\psi_{\boldsymbol{\lambda}}(t)\rangle \\ &\quad + \langle\psi_{\boldsymbol{\lambda}}(t)|\partial_{\mu}\psi_{\boldsymbol{\lambda}}(t)\rangle\langle\psi_{\boldsymbol{\lambda}}(t)|\partial_{\nu}\psi_{\boldsymbol{\lambda}}(t)\rangle \\ &\quad + \langle\partial_{\nu}\psi_{\boldsymbol{\lambda}}(t)|\partial_{\mu}\psi_{\boldsymbol{\lambda}}(t)\rangle + \langle\partial_{\mu}\psi_{\boldsymbol{\lambda}}(t)|\partial_{\nu}\psi_{\boldsymbol{\lambda}}(t)\rangle) \\ &= 4\mathrm{Re}\left(\langle\partial_{\mu}\psi_{\boldsymbol{\lambda}}(t)|\partial_{\nu}\psi_{\boldsymbol{\lambda}}(t)\rangle\right. \\ &\quad \left. - \langle\partial_{\mu}\psi_{\boldsymbol{\lambda}}(t)|\psi_{\boldsymbol{\lambda}}(t)\rangle\langle\psi_{\boldsymbol{\lambda}}(t)|\partial_{\nu}\psi_{\boldsymbol{\lambda}}(t)\rangle\right). \end{aligned} \tag{9}$$

In practical scenarios of multi-parameter quantum estimation, different estimation tasks prioritize the precision of different parameters. To address this issue, a positive semi-definite matrix W can be introduced to convert the matrix-form quantum Cramér–Rao bound into a scalar bound

$$C(t) = \mathrm{Tr}[WQ^{-1}], \tag{10}$$

where W serves as a weight matrix that represents the degree of interest in different parameters.

In multi-parameter quantum estimation, a central concern is whether the quantum Cramér–Rao bound can be attained,[45,54] which differs from single-parameter estimation where the quantum Cramér–Rao bound is always attainable. This inaccessibility arises from potential incompatibilities between the optimal measurements for different parameters.[55] The following quantum statistical models $\rho_{\boldsymbol{\lambda}}$ have been found to reach the highest precision for all unknown parameters simultaneously:[47,56–58] (i) Classical quantum statistical models: the density matrix $\rho_{\boldsymbol{\lambda}}$ can be diagonalized by a unitary transformation independent of λ, $\rho_{\boldsymbol{\lambda}} = U\Lambda_{\boldsymbol{\lambda}}U^{\dagger}$, where $\Lambda_{\boldsymbol{\lambda}}$ is a diagonal matrix; (ii) Quasi-classical quantum statistical models: all the SLD operators commute with each other at any point $\boldsymbol{\lambda}$,

$$[L^{\mathrm{S}}_{\mu}, L^{\mathrm{S}}_{\nu}] = 0, \quad \forall\mu,\nu, \text{ and } \forall\boldsymbol{\lambda}. \tag{11}$$

(iii) D-invariant quantum statistical models: the SLD-tangent space at λ is an invariant subspace of the commutation superoperator, and collective measurements on an asymptotically large number of copies $\rho_{\boldsymbol{\lambda}}^{\otimes n}$ are required to achieve the quantum Cramér–Rao bound under asymptotic conditions. (iv) Asymptotically classical quantum statistical models: all the SLD operators commute on average over $\rho_{\boldsymbol{\lambda}}$, a property known as weak commutativity

$$\mathrm{Tr}[\rho_{\boldsymbol{\lambda}}[L^{\mathrm{S}}_{\mu}, L^{\mathrm{S}}_{\nu}]] = 0, \quad \forall\mu,\nu, \text{ and } \forall\boldsymbol{\lambda}. \tag{12}$$

In this case, the SLD Cramér–Rao bound can be asymptotically saturated through collective measurements on an asymptotically large number of copies of $\rho_{\boldsymbol{\lambda}}^{\otimes n}$, and the SLD Cramér–Rao bound becomes equivalent to the well-known Holevo bound.

2.2. Model

In this study, in order to explore multi-parameter quantum metrology with time-dependent Hamiltonians, we consider a qubit system subjected to an oscillating magnetic field, where the magnetic field have two frequencies which are nearly equal in magnitude in two orthogonal directions (see Fig. 2). Our primary focus is on estimating the difference between these two frequencies. However, since their mean is generally unknown, we aim to improve the overall precision in estimating both the mean and the difference of the magnetic field frequencies in the two orthogonal directions. To ensure the attainability of the SLD Cramér–Rao bound for the overall precision,[59] we introduce an auxiliary qubit system within the SU(2) dynamics framework, as illustrated in Fig. 3. We employ a maximally entangled two-qubit state, $|\psi(0)\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, to satisfy the weak commutativity condition (12).

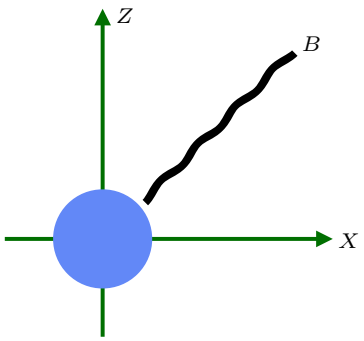


Fig. 2. Schematic diagram of the quantum system: A qubit subjected to an oscillating magnetic field with nearly similar frequencies in orthogonal spatial axes.

The Hamiltonian of the current model can be written as

$$H_0(t) = -B\left(\cos(\overline{\omega} - \delta\omega)t\sigma_x + \sin(\overline{\omega} + \delta\omega)t\sigma_z\right)\otimes I, \tag{13}$$

where B denotes the strength of the external magnetic field, and $\overline{\omega}$ and $\delta\omega$ are respectively the mean and the difference of the magnetic field frequencies in the two orthogonal directions which are unknown and to be estimated. These parameters are assumed to satisfy $\omega \gg \delta\omega$, *i.e.*, the magnetic field frequencies in both directions are of similar magnitude.

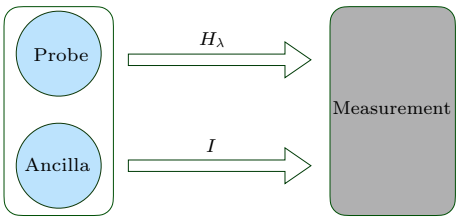


Fig. 3. Process of single-qubit multi-parameter quantum estimation with an ancilla. The multi-parameter quantum estimation with an auxiliary system comprises a probe qubit and an ancilla qubit, both initialized in a maximally entangled state. The system evolves under the Hamiltonian $H = H_\lambda \otimes I$, where H_λ is the parameter-dependent Hamiltonian and λ is the unknown parameter to be estimated. Upon completion of the evolution, a joint quantum state measurement of the probe-ancilla system is performed to extract the encoded parameter information.

Within this framework, the total Hamiltonian is expressed as

$$H(t) = H_0(t) + H_c(t), \tag{14}$$

where the control Hamiltonian $H_c(t)$, which is to be optimized for the qubit system, is assumed to be

$$H_c(t) = \sum_{i=1}^{3} V_i(t)\sigma_i \otimes I. \tag{15}$$

For a time-dependent Hamiltonian, the solution to the Schrödinger equation is often obtained using the time-evolution operator, $U(t)$, which satisfies

$$\mathrm{i}\frac{\partial U(t)}{\partial t} = H(t)U(t). \tag{16}$$

The time-evolution operator can be formally expressed as

$$U(t) = \mathscr{T}\exp\left[-\mathrm{i}\int_0^t H(\tau)\,\mathrm{d}\tau\right], \tag{17}$$

where $\mathscr{T}$ is the time-ordering operator that ensures the Hamiltonian is evaluated in the correct chronological order. Then the evolved state of the system at time t is then given by

$$|\psi(t)\rangle = U(t)\,|\psi(0)\rangle. \tag{18}$$

By carefully designing the control Hamiltonian $H_c(t)$, one can achieve precise control over the quantum system's evolution, enabling a wide range of quantum technologies, such as quantum computation, quantum simulation, and quantum metrology.

Here, we introduce a diagonal weight matrix $W = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, where a and b represent the weights assigned to the mean frequency $\overline{\omega}$ and the frequency difference $\delta\omega$, respectively. According to Eq. (10), the SLD Cramér–Rao bound for multi-parameter quantum estimation of $\overline{\omega}$ and $\delta\omega$ can be written as

$$C(t) = \frac{bQ^{\mathrm{S}}_{\overline{\omega}\overline{\omega}}(t) + aQ^{\mathrm{S}}_{\delta\omega\delta\omega}(t)}{Q^{\mathrm{S}}_{\overline{\omega}\overline{\omega}}(t)Q^{\mathrm{S}}_{\delta\omega\delta\omega}(t) - Q^{\mathrm{S}}_{\overline{\omega}\delta\omega}(t)Q^{\mathrm{S}}_{\delta\omega\overline{\omega}}(t)}, \tag{19}$$

where

$$\begin{aligned} Q^{\mathrm{S}}_{\mu\nu} &= \mathrm{Tr}[|\psi(t)\rangle\langle\psi(t)|\frac{L^{\mathrm{S}}_\mu L^{\mathrm{S}}_\nu + L^{\mathrm{S}}_\nu L^{\mathrm{S}}_\mu}{2}] \\ &= 4\mathrm{Re}\left(\langle\partial_\mu\psi_\lambda(t)|\partial_\nu\psi_\lambda(t)\rangle\right. \\ &\quad \left. - \langle\partial_\mu\psi_\lambda(t)|\psi_\lambda(t)\rangle\langle\psi_\lambda(t)|\partial_\nu\psi_\lambda(t)\rangle\right) \end{aligned} \tag{20}$$

is the elements of the quantum Fisher information matrix. Here, $Q^{\mathrm{S}}_{\overline{\omega}\overline{\omega}}(t)$ and $Q^{\mathrm{S}}_{\delta\omega\delta\omega}(t)$ are the quantum Fisher information for the parameters $\overline{\omega}$ and $\delta\omega$, respectively, and $Q^{\mathrm{S}}_{\overline{\omega}\delta\omega}(t)$ and $Q^{\mathrm{S}}_{\delta\omega\overline{\omega}}(t)$ represent the cross-correlation information between these parameters.

2.3. Method

In this study, we employ the GRAPE and hybrid reinforcement learning to optimize quantum controls for enhancing the precision of multi-parameter estimation. Although gradient-based methods are commonly used in reinforcement learning, they encounter challenges such as vanishing or exploding gradients, which hinder effective model updates, as well as high sensitivity to hyperparameters, especially the learning rate, which can significantly affect the stability and speed of the learning process. To address these issues, we incorporate heuristic optimization techniques, including genetic algorithms and particle swarm optimization, into specific modules to replace traditional gradient-based updates. Furthermore, to ensure the mathematical convergence of the optimization results obtained via hybrid reinforcement learning, we utilize the GRAPE method to yield more precise update of the control parameters.

2.3.1. Gradient ascent pulse engineering

We introduce the gradient ascent pulse engineering method, a widely employed optimization technique. The objective function $f(x)$ is a function of the parameter x, which typically represents the variables or control parameter being optimized. The core principle involves iteratively adjusting parameters by computing the gradient of the objective function at each step and gradually approaching the minimum value of the objective function $f(x)$. The fundamental steps of GRAPE are outlined as follows:

(i) Randomly select an initial set of parameters $\{x_0\}$.

(ii) Compute the gradient of the objective function at the current parameters, indicating the direction of the steepest descent.

(iii) Adjust the parameters in the opposite direction of the gradient by the update rule

$$x_{i+1} = x_i - \kappa\nabla f(x_i), \tag{21}$$

where x_i denotes the current values of the parameters, κ represents the step size, controlling the magnitude of each update,

and $\nabla f(x_i)$ represents the gradient of the objective function $f(x)$ at x_i.

(iv) Terminate the iteration if either the gradient criterion $\|\nabla f(x_i)\| < \varepsilon$ or the update magnitude criterion $\|x_{i+1} - x_i\| < \varepsilon$ is satisfied, where ε denotes a sufficiently small positive constant. If neither condition is satisfied, return to step 2 and continue iterating.

Despite its widespread application in various optimization problems, GRAPE exhibits several limitations, such as sensitivity to initial parameters and susceptibility to being trapped in local optima.

2.3.2. Reinforcement learning

Deep reinforcement learning (DRL) is a machine learning approach in which an agent learns optimal strategies through interactions with the environment. RL problems are typically modeled as Markov decision processes, which comprise five essential elements, state space S, action space A, transition probability $P(s'|s,a)$, reward function $R(s,a)$, and discount factor γ. During the reinforcement learning process, the agent begins in an initial state $s_0 \in S$, selects an action $a_0 \in A$, and interacts with the environment to receive the next state $s_1 \in S$ and reward r_1, based on the transition probabilities and reward function. Through successive interactions with the environment, the agent selects actions based on the current state at each time step, generating a Markov trajectory $(s_0, a_0, r_1, s_1, a_1, r_2, \ldots)$ and continuously refining its policy to maximize long-term cumulative rewards. The total discounted reward is defined as the weighted sum of all future rewards starting from time step t.

To identify the optimal policy π^* that maximizes the return, reinforcement learning employs the Bellman equation to solve and optimize the policy. The Bellman equation establishes a recursive relationship between the value of the current state and the values of subsequent states, providing a theoretical foundation for the optimal policy. For the optimal state value function $V^*(s)$, the Bellman equation can be given by

$$V(s) = \max_a \left[R(s,a) + \gamma \sum_{s'} P(s'|s,a) V(s') \right], \tag{22}$$

where $V(s)$ represents the maximum expected return achievable from state s, and $\max_a$ denotes the action that maximizes this return. The reward $R(s,a)$ is generally set in accordance with the objective function of the optimization problem, guiding the agent to adopt the optimal policy for achieving the desired outcome. By continuously optimizing value functions, reinforcement learning progressively converges towards the optimal policy, ultimately enabling the agent to maximize long-term cumulative rewards.

In this work, we employ the advantage Actor–Critic algorithm,[60] a widely used reinforcement learning method, to find the optimal quantum control strategy. Within the reinforcement learning framework, the quantum state corresponds to the environment state, the control Hamiltonian serves as the action, and the estimation precision acts as the reward. At each time step t, the RL agent generates the control Hamiltonian u_t based on the current quantum state ρ_t through the policy network. The action u_t interacts with the environment, driving the system to evolve into the next quantum state ρ_{t+1} and generating a reward R_t associated with the estimation precision. Through continuous interactions, the agent gradually optimizes its policy to maximize the estimation precision.

3. Results and discussion

3.1. Optimal multiparameter estimation

Here, we consider the weight matrix $W_1 = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$, which represents a uniformly average precision for both the mean and the difference of the two magnetic field frequencies. In this case, the SLD Cramér–Rao bound can be written as

$$\begin{aligned} C_{\text{multi11}}(t) &= \text{Tr}[W_1 Q^{-1}] \\ &= \frac{Q^{\text{S}}_{\overline{\omega}\overline{\omega}}(t) + Q^{\text{S}}_{\delta\omega\delta\omega}(t)}{2(Q^{\text{S}}_{\overline{\omega}\overline{\omega}}(t) Q^{\text{S}}_{\delta\omega\delta\omega}(t) - Q^{\text{S}}_{\overline{\omega}\delta\omega}(t) Q^{\text{S}}_{\delta\omega\overline{\omega}}(t))}. \end{aligned} \tag{23}$$

To better illustrate the optimization effects, the figures below present the multi-parameter estimation precision in terms of the natural logarithm to show the scaling of the estimation precision more explicitly. As shown in Fig. 4, the blue line, representing the estimation precision without optimization, shows the lowest overall precision. The orange line, corresponding to the optimized result obtained by GRAPE, demonstrates an improvement over the non-optimized scenario but still exhibits fluctuations due to its sensitivity to the initial condition and its tendency to be trapped in local minimas. In contrast, the green line, obtained using hybrid reinforcement learning, achieves the highest precision, indicating superior performance and greater potential in optimizing multi-parameter estimation precision.

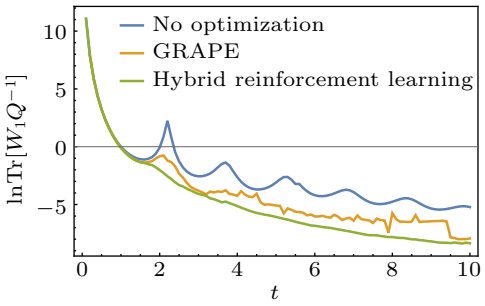


Fig. 4. The overall precision of the mean and difference of the magnetic field frequencies varies over time under different control strategies for a single qubit with an auxiliary qubit system using the maximally entangled state as the initial state. The blue solid line shows the precision variation with target time t under unoptimized conditions, while the yellow and green curves represent the outcomes after optimizing the control Hamiltonian with GRAPE and hybrid reinforcement learning, respectively. The true values are 1, 1, 0.001 for B, $\overline{\omega}$, $\delta\omega$.

One should note that the optimal control strategies derived from numerical optimization are often closely tied to the target evolution time. When the target evolution time varies, the corresponding optimal parameters must be adjusted accordingly. Figure 5 presents the optimal joint control schemes obtained within the time intervals $t = 5$ and $t = 10$.

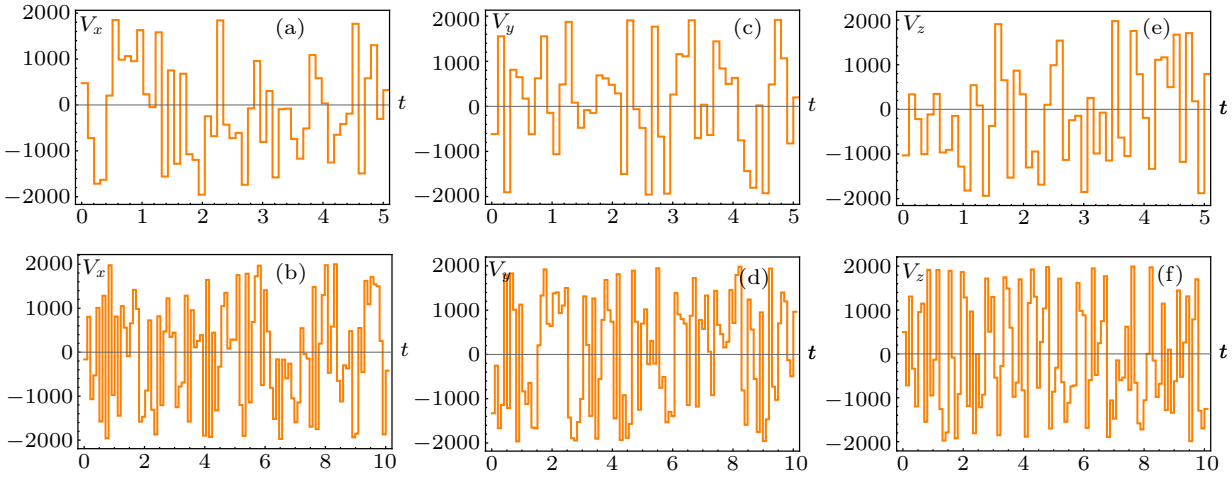


Fig. 5. The optimal control strategies for the σ_x, σ_y, σ_z components of the Hamiltonian over the time intervals from 0 to the target times $t = 5$ and $t = 10$: (a) σ_x-direction with $t = 5$, (b) σ_x-direction with $t = 10$, (c) σ_y-direction with $t = 5$, (d) σ_y-direction with $t = 10$, (e) σ_z-direction with $t = 5$, (f) σ_z-direction with $t = 10$.

3.2. Robustness

Robustness is a critical concern in quantum parameter estimation. The optimal controls often rely on the parameters to be estimated which are unknown a priori. When applying the optimal control based on the estimated values of the unknown parameters, deviations between the estimated and true values of the parameters may lead to discrepancies between the theoretical optimal estimation precision and the actual performance. Consequently, it is essential to evaluate the robustness of the derived optimal control strategies against the deviations between the true values and the estimated values of the parameters involved in the control Hamiltonian.

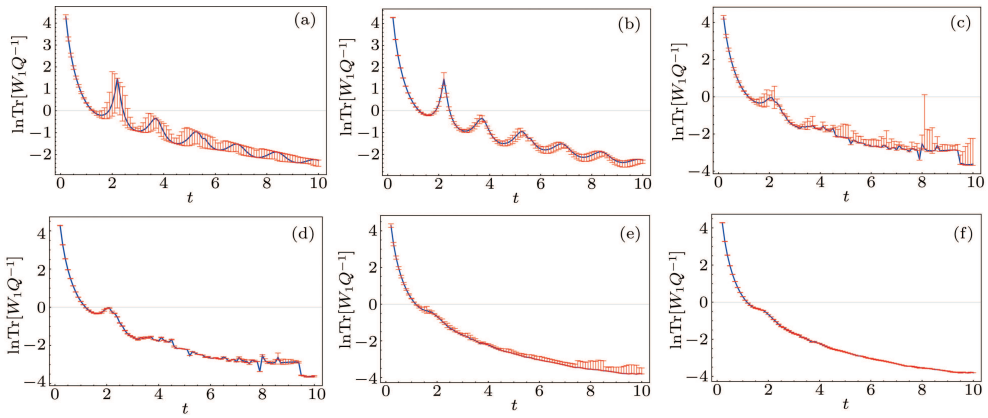


Fig. 6. The overall precision of the mean and difference of the magnetic field frequencies varies as the parameter $\overline{\omega}$ or $\delta\omega$ deviates from its true values. The blue solid line represents the precision limit at each time moment when the parameters take their true values, where $\overline{\omega}$ is 1 and $\delta\omega$ is 0.001. The red error bars indicate the maximum deviation of the SLD Cramér–Rao bound at different times when the mean frequency $\overline{\omega}$ deviates within a certain range $\Delta\overline{\omega} = [-0.1, 0.1]$ or the frequency difference $\delta\omega$ deviates within a certain range $\Delta\delta\omega = [-0.01, 0.01]$. (a) Scheme without optimization. The estimation precision limit varies from the parameter $\overline{\omega}$, when $\delta\omega$ is fixed. (b) Scheme without optimization. The estimation precision limit varies from the parameter $\delta\omega$, when $\overline{\omega}$ is fixed. (c) Scheme with GRAPE. The estimation precision limit varies from the parameter $\overline{\omega}$, when $\delta\omega$ is fixed. (d) Scheme with GRAPE. The estimation precision limit varies from the parameter $\delta\omega$, when $\overline{\omega}$ is fixed. (e) Scheme with the hybrid reinforcement learning. The estimation precision limit varies from the parameter $\overline{\omega}$, when $\delta\omega$ is fixed. (f) Scheme with the hybrid reinforcement learning. The estimation precision limit varies from the parameter $\delta\omega$, when $\overline{\omega}$ is fixed.

In this subsection, we investigate the robustness of the optimized quantum metrology scheme when the parameters deviate from their true values. As shown in Fig. 6, we illustrate the range of estimation precision in response to the deviation of the parameter for different evolution times. The results indicate that when the mean frequency $\overline{\omega}$ deviates from its true value within the range $\Delta\overline{\omega} = [-0.1, 0.1]$, the largest deviations in estimation precision occur in the absence of Hamiltonian control, while the GRAPE method reduces the fluctuations, and hybrid reinforcement learning achieves the best performance. For the estimated frequency difference $\delta\omega$ deviating from its true value within the range $\Delta\delta\omega = [-0.01, 0.01]$, the overall fluctuations in estimation

precision remain relatively small, with hybrid reinforcement learning exhibiting the highest robustness. Furthermore, in Fig. 7, we present a contour plot of $\ln[\mathrm{Tr}[W_1 Q^{-1}]]$ to illustrate how the overall precision of the mean and difference of the magnetic field frequencies varies when both the mean frequency $\overline{\omega}$ and the frequency difference $\delta\omega$ deviate simultaneously. In the absence of Hamiltonian control, the precision fluctuates by up to approximately 114%. The GRAPE method improves estimation precision but exhibits relatively large fluctuations, reaching approximately 278%. In contrast, hybrid reinforcement learning not only achieves the highest precision but also significantly reduces fluctuations, constraining deviations to around 27%. These findings demonstrate that hybrid reinforcement learning not only enhances the parameter estimation precision but also improves the robustness of the control strategy.

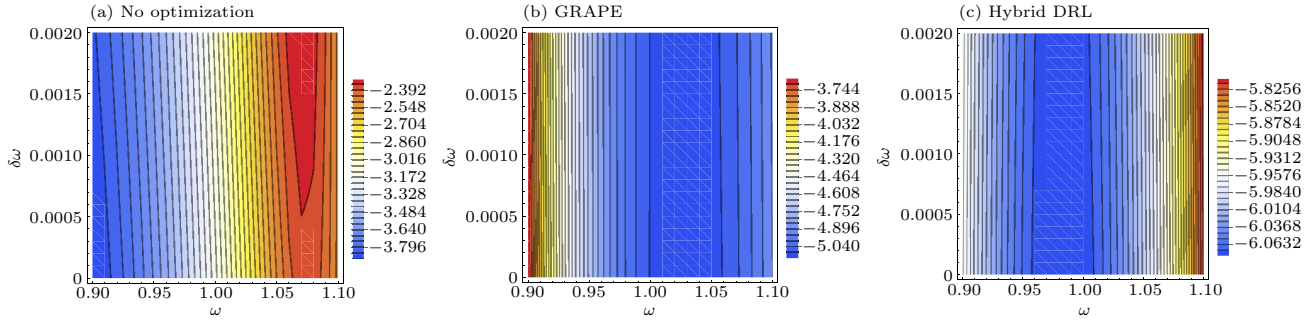


Fig. 7. $\ln(\mathrm{Tr}[W_1 Q^{-1}])$ as both the mean frequency $\overline{\omega}$ and the frequency difference $\delta\omega$ deviate simultaneously at target time $t = 5$. (a) Scheme without any optimization. (b) Scheme with GRAPE. (c) Scheme with the hybrid reinforcement learning.

3.3. Incompatibility

In this subsection, we focus on the incompatibility between the control Hamiltonians for different parameters. In single-parameter quantum estimation, by introducing the optimal control Hamiltonian, the upper bound of the Fisher information can be achieved as[30]

$$Q_\lambda = \left[\int_0^T (\mu_{\max}(t) - \mu_{\min}(t))\,\mathrm{d}t\right]^2, \tag{24}$$

where $\mu_{\max}$ and $\mu_{\min}$ are the maximum and minimum eigenvalues of $\partial_\lambda H_\lambda$, respectively. However, in multi-parameter quantum estimation, the optimal control Hamiltonian derived for single-parameter estimation generally cannot simultaneously optimize the estimation precision of multiple parameters as the control Hamiltonians for different parameters can be different. Therefore, it is necessary to design Hamiltonian control strategies that can enhance the overall precision of different parameters, resolving trade-offs between the respective control strategies for the unknown parameters. An optimal control method is proposed recently for scenarios where the off-diagonal elements of the Fisher information matrix are zero, that is $\langle \partial_\mu H(t)\, \partial_\nu H(t)\rangle = 0$.[59] However, for more general models, the off-diagonal elements of the Fisher information matrix are non-zero. The latter is also the case in our model, where

$$\langle \partial_{\overline{\omega}} H(t) \partial_{\delta\omega} H(t)\rangle = \frac{1}{2} B^2 t^2 \langle \cos[2(\overline{\omega} - \delta\omega)t]\, I + \cos[2(\overline{\omega} + \delta\omega)t]\, I\rangle, \tag{25}$$

which is generally not zero. So, we employ numerical optimization methods to investigate this scenario.

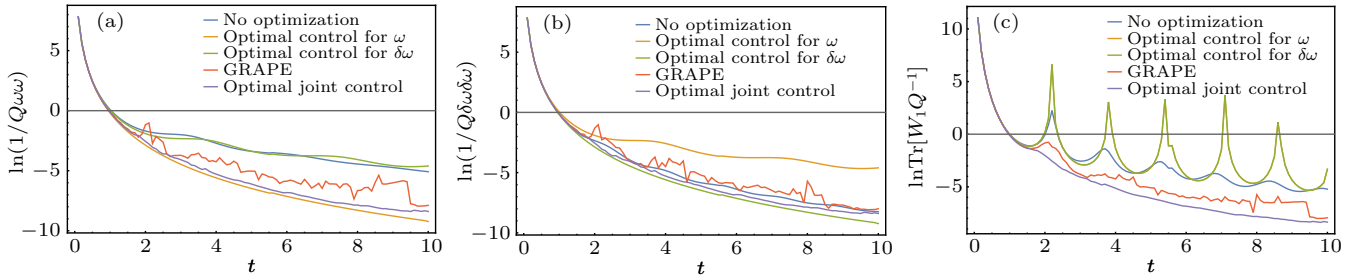


Fig. 8. The precision of parameter estimation over time under different control Hamiltonians strategies. The blue line denotes the scenario without any applied control Hamiltonian, the orange line represents the optimal control Hamiltonian for the mean frequency $\overline{\omega}$, the green line corresponds to the optimal control Hamiltonian for the frequency difference $\delta\omega$, the red line corresponds to results optimized using GRAPE, and the purple line depicts the control Hamiltonian obtained through the hybrid optimization algorithm. (a) The logarithm of the SLD Cramér–Rao bound $\ln\left[1/\mathcal{Q}^{\mathrm{S}}_{\overline{\omega}\overline{\omega}}(t)\right]$ for the mean frequency $\overline{\omega}$ as a function of target time t. (b) The logarithm of the SLD Cramér–Rao bound $\ln\left[1/\mathcal{Q}^{\mathrm{S}}_{\delta\omega\delta\omega}(t)\right]$ for the frequency difference $\delta\omega$ as a function of target time t. (c) The logarithm of the SLD Cramér–Rao bound $C_{\mathrm{multi11}}(t)$ for multi-parameter quantum estimation the mean frequency $\overline{\omega}$ and the frequency difference $\delta\omega$ as a function of target time t.

The comparative results of different methods are presented in Fig. 8. It is evident that optimal control strategies derived for single-parameter estimation do not consistently perform well in multi-parameter estimation, due to the fundamental incompatibility between optimal strategies for the estimation of different parameters. The results demonstrate that hybrid reinforcement learning significantly enhances the precision of both the mean and difference of the magnetic field frequencies and therefore addresses the challenges of control strategy incompatibilities in estimating the mean frequency $\overline{\omega}$ and frequency difference $\delta\omega$.

In the preceding section, we investigated the scenario where the weight matrix is $W_1 = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$, indicating equal emphasis on both parameters. However, in certain situations, when focusing on one or a few specific unknown parameters, different weight matrix W needs to be selected to represent the varying interest in different parameters. For example, $W_{21} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ emphasize the first parameter, while $W_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ focuses on the second parameter. In our model, when focusing on the mean frequency $\overline{\omega}$, the weight matrix should be W_{21} and the SLD Cramér–Rao bound is given by

$$\begin{aligned} C_{\text{multi10}}(t) &= \text{Tr}\left[W_{21}Q^{-1}\right] \\ &= \frac{Q_{\delta\omega\delta\omega}(t)}{Q_{\overline{\omega}\overline{\omega}}(t)Q_{\delta\omega\delta\omega}(t) - Q_{\overline{\omega}\delta\omega}(t)Q_{\delta\omega\overline{\omega}}(t)}. \end{aligned} \tag{26}$$

Similarly, when concentrating on the frequency difference $\delta\omega$, the weight matrix is W_{22} and the SLD Cramér–Rao bound becomes

$$\begin{aligned} C_{\text{multi01}}(t) &= \text{Tr}\left[W_{22}Q^{-1}\right] \\ &= \frac{Q_{\overline{\omega}\overline{\omega}}(t)}{Q_{\overline{\omega}\overline{\omega}}(t)Q_{\delta\omega\delta\omega}(t) - Q_{\overline{\omega}\delta\omega}(t)Q_{\delta\omega\overline{\omega}}(t)}. \end{aligned} \tag{27}$$

Furthermore, if one parameter is known, it suffices to estimate the remaining unknown parameter. In this scenario, the problem becomes a single-parameter estimation, and the SLD Cramér–Rao bounds for the two parameters are respectively

$$C_{\text{single1}}(t) = \frac{1}{Q_{\overline{\omega}\overline{\omega}}(t)}, \tag{28}$$

$$C_{\text{single2}}(t) = \frac{1}{Q_{\delta\omega\delta\omega}(t)}. \tag{29}$$

We define the following metrics to compare the precision of multi-parameter estimation and single-parameter estimation

$$C_{\text{m1}}|_{W=W_1}(t) = C_{\text{multi11}}(t), \tag{30}$$

$$C_{\text{m2}}|_{W=W_{21}+W_{22}}(t) = \frac{1}{2}(C_{\text{multi10}}(t) + C_{\text{multi01}}(t)), \tag{31}$$

$$C_{\text{s}}(t) = \frac{1}{2}(C_{\text{single1}}(t) + C_{\text{single2}}(t)), \tag{32}$$

where C_{m1} stands for the precision when both parameters are unknown and estimated with equal weights, C_{m2} denotes the average of the precision when the weights are biased towards one of the parameters, and C_{s} represents the average of the precision for single-parameter estimation of the two parameters. The optimized results of the estimation precision, as depicted in Fig. 9, satisfy the relation $C_{\text{m1}} \geq C_{\text{m2}} \geq C_{\text{s}}$. These findings demonstrate that the single-parameter estimation achieves higher precision than multi-parameter estimation, which is expected due to potential incompatibilities in estimating different parameters. In multi-parameter estimation, the parameters are generally correlated, with some parameters depending on the others, whereas in single-parameter estimation, the estimation of each parameter is unaffected by the other parameters. In multi-parameter quantum estimation, when all unknown parameters are estimated simultaneously, there exists a potential incompatibility between the optimal measurements for different parameters, leading to trade-offs between the precisions for different parameters. This difference reflects the distinction between multi-parameter quantum estimation and single-parameter quantum estimation.

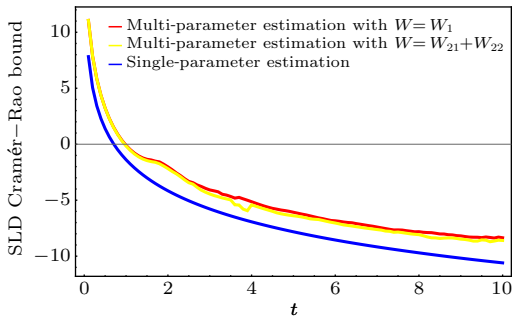


Fig. 9. The SLD Cramér–Rao bound as a function of target time t for multi-parameter estimation with different weight matrices and single-parameter estimation. The solid red line denotes the highest precision achieved when both the mean and difference of the magnetic field frequencies are given equal consideration in multi-parameter estimation, the yellow one represents the average of the highest precision for multiparameter estimation focusing only on each parameter separately, and the blue one represents the average of the highest precision for single-parameter estimation.

4. Conclusion

In this work, we explore the optimization methods for multi-parameter estimation of two close frequencies of a magnetic field by using a single qubit. We apply various methods to optimize the quantum Cramér–Rao bound. The results demonstrate that the hybrid reinforcement learning approach achieves the highest precision of optimization, significantly enhances estimation precision, and exhibits superior robustness with stable performance against deviations between the optimal and practical values of the control parameters. Additionally, we address the issue of control strategy incompatibility in the general case where the off-diagonal elements of the Fisher information matrix are non-zero. The findings indicate that our proposed method excels in mitigating control incompatibilities between different parameters. Furthermore, the precisions of estimating the two frequencies

both simultaneously and individually are analyzed, manifesting the distinctions between multi-parameter estimation and single-parameter estimation. Our research provides methods for enhancing the precision and robustness of multi-parameter quantum estimation and addresses control incompatibility issues, underscoring its potential application in optimal multi-parameter quantum estimation.

Acknowledgements

Project supported by the National Natural Science Foundation of China (Grant No. 12075323) and the Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0300702).

References

[1] Helstrom C W 1969 *J. Stat. Phys.* **1** 231
[2] Caves C M 1981 *Phys. Rev. D* **23** 1693
[3] Giovannetti V, Lloyd S and Maccone L 2004 *Science* **306** 1330
[4] Giovannetti V, Lloyd S and Maccone L 2011 *Nat. Photon.* **5** 222
[5] Paris M G A 2009 *International Journal of Quantum Information* **07** 125
[6] Escher B M, de Matos Filho R L and Davidovich L 2011 *Nat. Phys.* **7** 406
[7] Tsang M 2013 *New J. Phys.* **15** 073005
[8] Demkowicz-Dobrzański R 2014 *Phys. Rev. Lett.* **113** 250801
[9] Tan Q S, Huang Y, Yin X, Kuang L M and Wang X 2013 *Phys. Rev. A* **87** 032102
[10] Bužek V, Derka R and Massar S 1999 *Phys. Rev. Lett.* **82** 2207
[11] Derevianko A and Katori H 2011 *Rev. Mod. Phys.* **83** 331
[12] Ludlow A D, Boyd M M, Ye J, Peik E and Schmidt P O 2015 *Rev. Mod. Phys.* **87** 637
[13] Bollinger J J, Itano W M, Wineland D J and Heinzen D J 1996 *Phys. Rev. A* **54** R4649
[14] Michael A Taylor and Warwick P Bowen 2016 *Phys. Rep.* **615** 1
[15] Cox G 2017 *Optical Imaging Techniques in Cell Biology*, 2nd edn. (CRC Press)
[16] Kolobov M I 1999 *Rev. Mod. Phys.* **71** 1539
[17] Lugiato L A, Gatti A and Brambilla E 2002 *J. Opt. B: Quantum Semiclass. Opt.* **4** S176
[18] Brida G, Genovese M and Ruo Berchera I 2010 *Nat. Photon.* **4** 227
[19] Morris P A, Aspden R S, Bell J E C, Boyd R W and Padgett M J 2015 *Nat. Commun.* **6** 5913
[20] Degen C L, Reinhard F and Cappellaro P 2017 *Rev. Mod. Phys.* **89** 035002
[21] Pirandola S, Bardhan B R, Gehring T, Weedbrook C and Lloyd S 2018 *Nat. Photon.* **12** 724
[22] Giovannetti V, Lloyd S and Maccone L 2006 *Phys. Rev. Lett.* **96** 010401
[23] Dorner U, Demkowicz-Dobrzanski R, Smith B J, Lundeen J S, Wasilewski W, Banaszek K and Walmsley I A 2009 *Phys. Rev. Lett.* **102** 040403
[24] Humphreys P C, Barbieri M, Datta A and Walmsley I A 2013 *Phys. Rev. Lett.* **111** 070403
[25] Safránek D and Fuentes I 2016 *Phys. Rev. A* **94** 062313
[26] Armen M A, Au J K, Stockton J K, Doherty A C and Mabuchi H 2002 *Phys. Rev. Lett.* **89** 133602
[27] Yue J D, Zhang Y R and Fan H 2014 *Sci. Rep.* **5933** 2045
[28] Brask J B, Chaves R and Kołodyński J 2015 *Phys. Rev. X* **5** 031010
[29] Knott P A, Proctor T J, Hayes A J, Ralph J F, Kok P and Dunningham J A 2016 *Phys. Rev. A* **94** 062312
[30] Pang S and Jordan A N 2017 *Nat. Commun.* **14695** 2041
[31] Yuan H and Fung C H F 2015 *Phys. Rev. Lett.* **115** 110401
[32] Liu J and Yuan H 2017 *Phys. Rev. A* **96** 012117
[33] Liu J and Yuan H 2017 *Phys. Rev. A* **96** 042114
[34] Lei X, Fan J and Pang S 2024 *Chin. Phys. B* **33** 060304
[35] Sutton R S and Barto A G 2018 *Reinforcement Learning, An Introduction*, 2nd edn. (Cambridge, Massachusetts: Bradford Books)
[36] Carrasquilla J and Melko R G 2017 *Nat. Phys.* **13** 431
[37] Deng D L, Li X and Das Sarma S 2017 *Phys. Rev. X* **7** 021021
[38] Hsu Y T, Li X, Deng D L and Das Sarma S 2018 *Phys. Rev. Lett.* **121** 245701
[39] Xu H, Li J, Liu L, Wang Y, Yuan H and Wang X 2019 *npj Quantum Information* **5** 1
[40] Xu H, Wang L, Yuan H and Wang X 2021 *Phys. Rev. A* **103** 042615
[41] Xiao T, Fan J and Zeng G 2022 *npj Quantum Information* **8** 1
[42] Care C M 1983 *Phys. Bull.* **34** 395
[43] Braunstein S L and Caves C M 1994 *Phys. Rev. Lett.* **72** 3439
[44] Braunstein S, Caves C and Milburn G 1996 *Ann. Phys.* **247** 135
[45] Szczykulska M, Baumgratz T and Datta A 2016 *Adv. Phys. X* **1** 621
[46] Liu J, Yuan H, Lu X M and Wang X 2019 *J. Phys. A: Math. Theor.* **53** 023001
[47] Ragy S, Jarzyna M and Demkowicz-Dobrzański R 2016 *Phys. Rev. A* **94** 052108
[48] Hou Z, Tang J F, Chen H, Yuan H, Xiang G Y, Li C F and Guo G C 2021 *Sci. Adv.* **7** eabd2986
[49] Serge Massar 2000 *Phys. Rev. A* **61** 042312
[50] Xia B, Huang J, Li H, Wang H and Zeng G 2023 *Nat. Commun.* **14** 1021
[51] Belliardo F and Giovannetti V 2021 *New J. Phys.* **23** 063055
[52] Helstrom C W 1967 *Phys. Lett. A* **25** 101
[53] Yuen H and Lax M 1973 *IEEE Transactions on Information Theory* **19** 740
[54] Matsumoto K 2002 *J. Phys. A: Math. Gen.* **35** 3111
[55] Pezzè L, Ciampini M A, Spagnolo N, Humphreys P C, Datta A, Walmsley I A, Barbieri M, Sciarrino F and Smerzi A 2017 *Phys. Rev. Lett.* **119** 130504
[56] Suzuki J 2019 *Entropy* **21** 703
[57] Suzuki J 2019 *Entropy* **21** 703
[58] Suzuki J 2016 *J. Math. Phys.* **57** 042201
[59] Hu Z, Wang S, Qiao L, Isogawa T, Li C, Yang Y, Wang G, Yuan H and Cappellaro P 2024 arXiv:2411.18896 [quant-ph]
[60] Mnih V 2016 arXiv preprint arXiv:1602.01783