

Noisy quantum parameter estimation with indefinite causal orderMin An,¹ Shihao Ru,¹ Yunlong Wang,¹ Yu Yang,¹ Feiran Wang,^{1,2} Pei Zhang,¹ and Fuli Li^{1,*}¹*Ministry of Education Key Laboratory for Nonequilibrium Synthesis and Modulation of Condensed Matter and Shaanxi Province Key Laboratory of Quantum Information and Quantum Optoelectronic Devices,**School of Physics, Xi'an Jiaotong University, Xi'an 710049, China*²*School of Science, Xi'an Polytechnic University, Xi'an 710048, China*

(Received 5 June 2023; revised 6 October 2023; accepted 11 December 2023; published 3 January 2024)

Optimal probe states and measurements are always required to achieve the best estimation precision in quantum metrology. In the actual physical environment, ubiquitous noise hinders estimation precision and amplifies the challenges of optimizing the probe states and measurements. In this study we extend a theoretical proposal, presented in *Phys. Rev. A* **103**, 032615 (2021), to encompass noisy general Pauli channels for SU(2) phase estimation. By utilizing a superposition of different causal orders of two channels, we establish theoretically a probe-state-independent criterion for the estimation precision. We show that the probe-state-independent property results from the noncommutativity of the Kraus operators of the parameter-encoding channels. Based on this criterion, one can find some parameter-encoding channels without the requirement of precisely preparing the probe state. Moreover, our scheme requires only deterministic projection measurement on the control qubit. In this way, one can simultaneously avoid optimizing both the probe states and measurements. In addition, we also show that the estimation precision and probe-state-independent property of the indefinite causal order scheme are independent of representations of the Kraus operators. This alleviates the experimental challenges in realizing quantum channels. We demonstrate experimentally the advantages of the indefinite causal order for phase estimation of an SU(2) unitary transformation in the three kinds of quantum noise channels. Our result shows that the dynamic evolution of indefinite causal order can outperform the conventional cascaded estimation scenario at high noise levels, exhibiting high robustness and feasibility in practical estimation tasks.

DOI: [10.1103/PhysRevA.109.012603](https://doi.org/10.1103/PhysRevA.109.012603)**I. INTRODUCTION**

Quantum metrology seeks to surpass the shot-noise limit and achieve the Heisenberg limit in parameter-estimation tasks by utilizing quantum resources such as entanglement and quantum superposition. However, quantum systems are susceptible to noise, which can diminish or even eliminate the benefits of quantum enhancement in estimation precision [1,2].

In order to depress quantum noise, one uses entangled probe states or entanglement-assisted schemes, which have shown advantages in phase estimation across different noise regimes [3–6]. Despite the challenges posed by noise, innovative methods have been proposed to maintain quantum enhancement. For example, delicate preparation of the probe state may be necessary to utilize noise information, although this is not always possible in real physical systems [7]. Additionally, sophisticated projection measurements are often required to achieve more precise parameter estimation, which can pose experimental difficulties and hinder the practical deployment of quantum metrology techniques. Encoding parameters within the system dynamics is also a crucial step in enhancing measurement precision and noise robustness [8–12]. Moreover, well-designed evolution dynamics have been proposed to circumvent the no-go theorem that limits

quantum advantages to a constant factor enhancement in noisy environments [13].

A new type of quantum evolution dynamics, called indefinite causal order, has recently been introduced [14,15]. This approach allows for quantum evolution in a superposition of two different causal orders and has been demonstrated experimentally through a quantum switch [16,17]. Indefinite causal order has been successfully applied in various fields, including quantum communication [18–21], quantum thermodynamics [22–24], and quantum metrology [25–28], yielding results that challenge classical understanding.

For single-parameter estimation, a strict hierarchy has been established between a general adaptive strategy and an indefinite causal order strategy in both nonasymptotic [29] and asymptotic regimes [30]. By utilizing the quantum switch, the indefinite causal order scheme can not only enhance the absolute estimation precision to super-Heisenberg scaling in a noise-free situation [31], but also potentially overcome the negative effects of depolarization in phase-estimation tasks of SU(2) systems [32] and can even turn noise into a beneficial factor to enhance parameter-estimation precision in intermediate noisy physical systems. Unlike general adaptive schemes, the indefinite causal order scheme does not require precisely preparing probe states or sophisticated projection measurements [33]. Although it may not achieve the Heisenberg limit when noise exists, it has been shown to outperform conventional SU(2) phase-estimation schemes in intermediate noisy environments within a nonasymptotic regime.

*fli@xjtu.edu.cn

In this work we focus on phase estimation of the SU(2) unitary transformation in a noisy quantum channel with the assistance of the quantum switch. Leveraging the incoherent superposition property of the quantum channels, we extend the theoretical proposal [32] to general Pauli channels. We analyze theoretically the origin of the probe-state-independent property in the estimation precision and establish a criterion to avoid precisely preparing the probe state. The probe-state-independent property is attributed to the noncommutativity of the Kraus operators of the parameter-encoding channels. We show that the representations of the Kraus operators have no impact on the estimation precision and probe-state-independent property, which alleviates the experimental challenges in realizing quantum channels. We conduct our experiments on a photonics platform, specifically utilizing a phase-flip channel, a depolarization channel, and an arbitrary Pauli channel, which are special cases of general Pauli channels. By encoding the rotation angle of the SU(2) transformation into the photonic polarization degree of freedom, we estimate the rotation angle and its variance. Our results show that the estimation precision reaches the bound setting by the quantum Fisher information. We implement the quantum switch by utilizing the path degree of freedom in a Mach-Zehnder interferometer. The quantum channels are realized by postprocessing the outcomes of different single experiments. To ensure the validity of our method, we perform quantum process tomography to verify the accuracy of the implemented channels. We also account for losses caused by imperfect optical elements in our experiments. Our experimental results reveal that the indefinite causal order scheme outperforms the conventional cascaded estimation scenario in situations with high noise levels. Notably, our approach does not require both precisely optimizing the probe state and sophisticated measurement setups. Moreover, our work can be extended to multiparameter estimation by redesigning the measurement settings to suit the specific requirements of the target parameters.

The structure of our paper is organized as follows. In Sec. II we provide an overview of the SU(2) phase-estimation task and the traditional scenario, highlighting the challenges in the existing scheme. In Sec. III B we extend the theoretical framework proposed in [32] to accommodate the general Pauli channel using the switch channel (detailed in Sec. III A). We present the experimental setup on a photonic platform in Sec. IV. The experimental results are shown in Sec. IV C.

II. PHASE ESTIMATION IN THE SU(2) SYSTEM

The Bloch representation is a useful tool for working with a qubit, a two-level quantum system. The density matrix of a qubit can be expressed as

$$\rho_{in} = \frac{1}{2}(\mathbb{1}_2 + \vec{r} \cdot \vec{\sigma}), \quad (1)$$

where \vec{r} is the Bloch vector that encodes all information about the qubit and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^T$. An SU(2) unitary operator representing a rotation by $\xi \in [0, 2\pi)$ about the normalized axis $\vec{n} = (n_x, n_y, n_z)^T$ can be expressed as

$$U_\xi = \exp\left(-i\frac{\xi}{2}\vec{n} \cdot \vec{\sigma}\right) = \cos\frac{\xi}{2}\mathbb{1}_2 - i\sin\frac{\xi}{2}\vec{n} \cdot \vec{\sigma}. \quad (2)$$

Assume we have an unknown phase ξ encoded on a probe qubit with a unitary transformation U_ξ specified by (2). We can express the encoded density matrix as

$$\begin{aligned} U_\xi \rho_{in} U_\xi^\dagger &= \frac{1}{2}(\mathbb{1}_2 + U_\xi \vec{r} \cdot \vec{\sigma} U_\xi^\dagger) \\ &= \frac{1}{2}[\mathbb{1}_2 + R_{\vec{n}}(\xi) \vec{r} \cdot \vec{\sigma}]. \end{aligned} \quad (3)$$

Utilizing the isomorphism between SU(2) and SO(3) rotations, we can represent the encoding process as a rotation around the Bloch vector \vec{r} , which is specified by Rodrigues' rotation formula

$$R_{\vec{n}}(\xi) = e^{\vec{n}\xi} = \mathbb{1}_3 + \vec{n} \sin \xi + \vec{n}^2(1 - \cos \xi). \quad (4)$$

Here $\mathbb{1}_3$ is the 3×3 identity matrix and \vec{n} denotes the anti-symmetric matrix defined as

$$\vec{n} = \begin{pmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{pmatrix}. \quad (5)$$

Our goal is to enhance the precision limit of phase estimation and noise robustness by carefully designing the dynamics of the system.

In the SU(2) phase-estimation task, the conventional scheme involves imprinting the estimated parameter ξ onto the probe state through the SU(2) dynamic evolution (2). Subsequently, a two-outcome positive-operator-valued measurement (POVM) $M_\pm = (\mathbb{1} \pm \vec{a} \cdot \vec{\sigma})/2$, $|\vec{a}| = 1$, is carried out on the encoded state (3). By establishing an unbiased estimator $\hat{\xi}$, i.e., $E(\hat{\xi}) = \xi$, we can obtain the variance of the estimated parameter ξ from the measurement probability distribution [33]

$$\begin{aligned} \text{var}(\hat{\xi}) &\geq \frac{1}{\mathcal{F}_C(\xi)} = \frac{1 - (\vec{a} \cdot \vec{r}_\xi)^2}{[\vec{a} \cdot (\vec{n} \times \vec{r}_\xi)]^2} \\ &\geq \frac{1}{\mathcal{F}_Q(\xi)} = \frac{1}{(\vec{n} \times \vec{r})^2}, \end{aligned} \quad (6)$$

where $\vec{r}_\xi = R_{\vec{n}}\vec{r}$, and $\mathcal{F}_C(\xi)$ and $\mathcal{F}_Q(\xi)$ are the classical Fisher information and quantum information, respectively (detailed in Appendix A).

Note that the lower bound of variance is influenced by the geometric relationship between the rotation axis \vec{n} and the Bloch vector of probe state \vec{r} . When \vec{n} is orthogonal to \vec{r} , the POVM required to achieve the quantum Cramér-Rao bound depends on the parameter being estimated [33]

$$\vec{a} = \vec{n} \times \vec{r}_\xi. \quad (7)$$

In this case, achieving optimal precision without prior knowledge of ξ is difficult. When noise is present, the carefully prepared probe state is corrupted, and optimizing the POVM is also difficult in most situations. To overcome this challenge, it is necessary to devise dynamic evolution that reduces the parameter's dependence on state preparation and measurement.

III. DESCRIPTION OF THE QUANTUM CHANNEL AND SWITCH CHANNEL

A. Operator-sum representation of the quantum channel

A quantum channel is a completely positive and trace-preserving map that transforms a density matrix into another

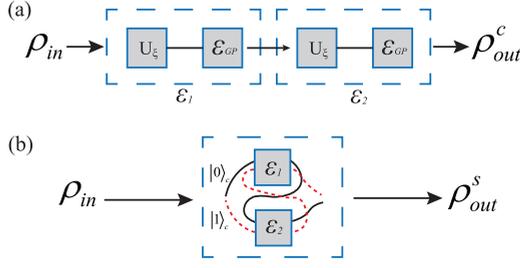


FIG. 1. The two types of quantum channels: (a) the conventional cascaded channel, where \mathcal{E}_1 and \mathcal{E}_2 are composed with a unitary transformation U_ξ and a general Pauli channel \mathcal{E}_{GP} in a cascade manner, and (b) the switch channel, where the order of two channels \mathcal{E}_1 and \mathcal{E}_2 can be controlled using the state of the control qubit.

density matrix. This map can be expressed using the operator-sum representation

$$\mathcal{E}(\rho) = \sum_i K_i \rho K_i^\dagger, \quad (8)$$

where Kraus operators $\{K_i\}$ satisfy $\sum_i K_i^\dagger K_i = \mathbb{1}_2$.

The evolution of a qubit can be regarded as a quantum channel acting on its density matrix (1), which can be expressed in the operator-sum representation as (8). The special case of unitary evolution occurs when there is only one Kraus operator, which is the unitary operator itself. For example, a general Pauli channel can be described as

$$\mathcal{E}_{GP}(\rho) = p_I \rho + p_x \sigma_x \rho \sigma_x + p_y \sigma_y \rho \sigma_y + p_z \sigma_z \rho \sigma_z, \quad (9)$$

where $p_I + p_x + p_y + p_z = 1$. The Kraus operators for the general Pauli channel are $\{\sqrt{p_I} \mathbb{1}_2, \sqrt{p_x} \sigma_x, \sqrt{p_y} \sigma_y, \sqrt{p_z} \sigma_z\}$, which reduce to specific channels such as phase-flip and depolarization channels depending on the values of the parameters $\{p_I, p_x, p_y, p_z\}$.

Consider two identical channels \mathcal{E}_1 and \mathcal{E}_2 , each composed of the unitary transformation U_ξ and general Pauli channel \mathcal{E}_{GP} sequentially. Let $K_i^{(1)}$ and $K_j^{(2)}$ be the Kraus operators of \mathcal{E}_1 and \mathcal{E}_2 , respectively, where i and j denote the index of the Kraus operators. The Kraus operators for \mathcal{E}_1 and \mathcal{E}_2 are

$$\begin{aligned} K_0^{(1)} = K_0^{(2)} &= \sqrt{p_I} U_\xi, & K_1^{(1)} = K_1^{(2)} &= \sqrt{p_x} \sigma_x U_\xi, \\ K_2^{(1)} = K_2^{(2)} &= \sqrt{p_y} \sigma_y U_\xi, & K_3^{(1)} = K_3^{(2)} &= \sqrt{p_z} \sigma_z U_\xi. \end{aligned} \quad (10)$$

Then we can consider the composition of two quantum channels \mathcal{E}_1 and \mathcal{E}_2 as a cascaded operation, denoted by $\mathcal{E}_2 \circ \mathcal{E}_1$ and illustrated in Fig. 1(a). The Kraus operators of the composite channel $\mathcal{E}_2 \circ \mathcal{E}_1$ can be obtained as $K_{ij} = K_j^{(2)} K_i^{(1)}$, where ij represents the index of the composite Kraus operators.

By coherently controlling the sequence of the two channels shown in Fig. 1(b), a switch channel, also known as a quantum switch, can be obtained. It can be described by the Kraus operators

$$K_{ij} = K_j^{(2)} K_i^{(1)} \otimes |0\rangle\langle 0|_c + K_i^{(1)} K_j^{(2)} \otimes |1\rangle\langle 1|_c. \quad (11)$$

Here $|0\rangle\langle 0|_c$ represents the control state that makes the composite channel be implemented as $\mathcal{E}_2 \circ \mathcal{E}_1$, while $|1\rangle\langle 1|_c$ represents the control state that enables the composite channel to be implemented as $\mathcal{E}_1 \circ \mathcal{E}_2$. The switch channel in the

operator-sum representation then is given as

$$\begin{aligned} \mathcal{S}(\rho_{in} \otimes \rho_c) &= \sum_{i,j} K_{ij} (\rho_{in} \otimes \rho_c) K_{ij}^\dagger \\ &= S_{00}(\rho) \otimes \langle 0|\rho_c|0\rangle |0\rangle\langle 0|_c \\ &\quad + S_{01}(\rho) \otimes \langle 0|\rho_c|1\rangle |0\rangle\langle 1|_c \\ &\quad + S_{01}^\dagger(\rho) \otimes \langle 1|\rho_c|0\rangle |1\rangle\langle 0|_c \\ &\quad + S_{11}(\rho) \otimes \langle 1|\rho_c|1\rangle |1\rangle\langle 1|_c, \end{aligned} \quad (12)$$

where ρ_c is the density matrix of the control qubit and $S_{00}(\rho_{in})$, $S_{01}(\rho_{in})$, and $S_{11}(\rho_{in})$ are three superoperators defined as

$$S_{00}(\rho_{in}) = \sum_{i,j} K_j^{(2)} K_i^{(1)} \rho_{in} K_i^{(1)\dagger} K_j^{(2)\dagger}, \quad (13)$$

$$S_{01}(\rho_{in}) = \sum_{i,j} K_j^{(2)} K_i^{(1)} \rho_{in} K_j^{(2)\dagger} K_i^{(1)\dagger}, \quad (14)$$

$$S_{11}(\rho_{in}) = \sum_{i,j} K_i^{(1)} K_j^{(2)} \rho_{in} K_j^{(2)\dagger} K_i^{(1)\dagger}. \quad (15)$$

Here we emphasize that the output of the switch channel $\mathcal{S}(\rho_{in} \otimes \rho_c)$ is independent of a specific representation of Kraus operators of the two channels \mathcal{E}_1 and \mathcal{E}_2 (see Appendix B for details). Because the phase-flip and phase-damping channels can be related by a unitary transformation, we only need to consider one of them under the same noise level in the experiment.

B. Phase estimation with two types of quantum channels

Now we consider the phase-estimation precision using quantum Fisher information for the two types of quantum channels depicted in Fig. 1.

For the conventional cascaded channel illustrated in Fig. 1(a),

$$\begin{aligned} \rho_{out}^c &= \mathcal{E}_2[\mathcal{E}_1(\rho_{in})] \\ &= \mathcal{E}_{GP}[U_\xi \mathcal{E}_{GP}(U_\xi \rho_{in} U_\xi^\dagger) U_\xi^\dagger] \\ &= \frac{1}{2} \{\mathbb{1}_2 + [AR_{\vec{n}}(\xi) AR_{\vec{n}}(\xi) \vec{r}] \cdot \vec{\sigma}\}, \end{aligned} \quad (16)$$

where $R_{\vec{n}}(\xi)$ is given in Eq. (4) and A is a matrix related to the parameters of the general Pauli channel \mathcal{E}_{GP} , which is given by

$$\begin{aligned} A &= \begin{pmatrix} a_{xx} & 0 & 0 \\ 0 & a_{yy} & 0 \\ 0 & 0 & a_{zz} \end{pmatrix}, \\ a_{xx} &= p_I + p_x - p_y - p_z, \\ a_{yy} &= p_I - p_x + p_y - p_z, \\ a_{zz} &= p_I - p_x - p_y + p_z. \end{aligned} \quad (17)$$

Denoting the Bloch vector of ρ_{out} by $\vec{r}_{\{\xi, \vec{p}\}}$, i.e., $\vec{r}_{\{\xi, \vec{p}\}} = AR_{\vec{n}}(\xi) AR_{\vec{n}}(\xi) \vec{r}$, one has the quantum Fisher information for the conventional cascaded channel [34]

$$\mathcal{F}_{\mathcal{Q}}^c(\xi) = \frac{(\vec{r}_{\{\xi, \vec{p}\}} \partial_\xi \vec{r}_{\{\xi, \vec{p}\}})^2}{1 - |\vec{r}_{\{\xi, \vec{p}\}}|^2} + (\partial_\xi \vec{r}_{\{\xi, \vec{p}\}})^2, \quad (18)$$

which is named after the traditional limit in the following discussion.

For the switch channel illustrated in Fig. 1(b), using (13), we can express $S_{00}(\rho_{\text{in}})$ as

$$\begin{aligned} S_{00}(\rho_{\text{in}}) &= p_I^2 U_\xi^2 \rho_{\text{in}} U_\xi^{\dagger 2} + p_I \sum_{l=x,y,z} p_l U_\xi \sigma_l U_\xi \rho_{\text{in}} U_\xi^\dagger \sigma_l U_\xi^\dagger \\ &\quad + p_I \sum_{l=x,y,z} p_l \sigma_l U_\xi^2 \rho_{\text{in}} U_\xi^{\dagger 2} \sigma_l \\ &\quad + \sum_{l,l'=x,y,z} p_l p_{l'} \sigma_l U_\xi \sigma_{l'} U_\xi \rho_{\text{in}} U_\xi^\dagger \sigma_{l'} U_\xi^\dagger \sigma_l. \end{aligned} \quad (19)$$

Similarly, using (14), we can express $S_{01}(\rho_{\text{in}})$ as

$$\begin{aligned} S_{01}(\rho_{\text{in}}) &= p_I^2 U_\xi^2 \rho_{\text{in}} U_\xi^{\dagger 2} + p_I \sum_{l=x,y,z} p_l U_\xi \sigma_l U_\xi \rho_{\text{in}} U_\xi^\dagger U_\xi^\dagger \sigma_l \\ &\quad + p_I \sum_{l=x,y,z} p_l \sigma_l U_\xi^2 \rho_{\text{in}} U_\xi^\dagger \sigma_l U_\xi^\dagger \\ &\quad + \sum_{l,l'=x,y,z} p_l p_{l'} \sigma_l U_\xi \sigma_{l'} U_\xi \rho_{\text{in}} U_\xi^\dagger \sigma_{l'} U_\xi^\dagger \sigma_l'. \end{aligned} \quad (20)$$

To characterize the switch channel with the input state (1), we can use the equations

$$S_{00}[\tfrac{1}{2}(\mathbb{1}_2 + \vec{r} \cdot \vec{\sigma})] = \tfrac{1}{2}S_{00}(\mathbb{1}_2) + \tfrac{1}{2}S_{00}(\vec{r} \cdot \vec{\sigma}), \quad (21)$$

$$S_{01}[\tfrac{1}{2}(\mathbb{1}_2 + \vec{r} \cdot \vec{\sigma})] = \tfrac{1}{2}S_{01}(\mathbb{1}_2) + \tfrac{1}{2}S_{01}(\vec{r} \cdot \vec{\sigma}), \quad (22)$$

where $S_{00}(\vec{r} \cdot \vec{\sigma})$ and $S_{01}(\vec{r} \cdot \vec{\sigma})$ are functions of $\vec{\sigma}$, and therefore $\text{Tr}[S_{00}(\vec{r} \cdot \vec{\sigma})] = \text{Tr}[S_{01}(\vec{r} \cdot \vec{\sigma})] = 0$.

We prepare the initial control state in $|\psi\rangle_c = \sqrt{p_c}|0\rangle_c + \sqrt{1-p_c}|1\rangle_c$. After interacting with the switch channel and projecting onto the basis $(|\pm\rangle_c = |0\rangle_c \pm |1\rangle_c)/\sqrt{2}$, the probe state becomes

$$\begin{aligned} \rho_\pm &= {}_c \langle \pm | S(\rho_{\text{in}} \otimes \rho_c) | \pm \rangle_c \\ &= \tfrac{1}{2} S_{00}(\rho_{\text{in}}) \pm \sqrt{p_c(1-p_c)} S_{01}(\rho_{\text{in}}). \end{aligned} \quad (23)$$

The probability of the control state being projected onto $|\pm\rangle_c$ is given by $\text{Tr}(\rho_\pm)$, which can be expressed as

$$\begin{aligned} P_\pm &= \text{Tr}(\rho_\pm) \\ &= \tfrac{1}{2} \text{Tr}[\tfrac{1}{2} S_{00}(\mathbb{1}_2) \pm \sqrt{p_c(1-p_c)} S_{01}(\mathbb{1}_2)] \\ &= \tfrac{1}{2} \pm \sqrt{p_c(1-p_c)} Q(\xi, p_I, p_x, p_y, p_z), \end{aligned} \quad (24)$$

with

$$\begin{aligned} Q(\xi, p_I, p_x, p_y, p_z) &= \sum_{i=l,x,y,z} p_i^2 - 4p_I \sin^2 \frac{\xi}{2} \sum_{i=x,y,z} p_i (1 - n_i^2) \\ &\quad + 2p_I(1 - p_I) - 2(p_x p_y + p_y p_z + p_z p_x) \\ &\quad + 4 \sin^2 \frac{\xi}{2} [p_x p_y (1 - n_x^2) \\ &\quad + p_y p_z (1 - n_x^2) + p_z p_x (1 - n_y^2)]. \end{aligned} \quad (25)$$

The classical Fisher information for the switch channel can be calculated from the probability distribution (24) as

$$\mathcal{F}_c^s(\xi) = \frac{4p_c(1-p_c)[\partial_\xi Q(\xi, p_I, p_x, p_y, p_z)]^2}{1 - 4p_c(1-p_c)Q^2(\xi, p_I, p_x, p_y, p_z)}. \quad (26)$$

The classical Fisher information is maximized when $p_c = \frac{1}{2}$.

The above equation is a general result with the general Pauli channels. We have also investigated other scenarios, such as the phase-flip channel, which are discussed in Appendix D. Remarkably, we find from Eq. (24) that the measurement outcomes are independent of the probe state, which means that optimizing probe states is not required in this case. This property is attributed to the noncommutativity of the Kraus operators of the quantum channel. When adopting two identical channels, i.e., $K_i^{(1)} = K_i^{(2)} = K_i$, with $O_{ij} = [K_i^{(1)}, K_j^{(2)}]$, we establish a probe-state-independent criterion. If probe-state independence is required, the condition holds for $\mathcal{O} = \sum_{i,j,i < j} O_{ij}^\dagger O_{ij}$,

$$\mathcal{O} = c \mathbb{1}_2, \quad c \in \mathbb{R}. \quad (27)$$

In the case of a general Pauli channel, this condition is satisfied (the details can be found in Appendix C). In the case of a general amplitude damping channel, this condition is not satisfied, so the measurement outcomes are dependent on the probe state, which has been verified by the previous work in [26]. Furthermore, only a parameter-independent projection measurement is needed on the control qubit. These advantages make the phase-estimation task of the SU(2) system more accessible.

We have also shown that the above results are irrelevant to the specific Kraus representations of the two channels \mathcal{E}_1 and \mathcal{E}_2 (see Appendix B for details). So we can choose the experimentally friendly Kraus operators when implementing quantum channels. Different quantum channels may be related to a unitary operator such as a phase-flip channel and a phase-damping channel, and this situation is very similar to a quantum channel with two kinds of Kraus representations. So the results of phase-flip channels and phase-damping channels are the same if we choose the same noise level in the two channels.

The density matrix of the control qubit after the switch channel is given by

$$\begin{aligned} \rho_{\text{out},c}^s &= \tfrac{1}{2} \text{Tr}[S_{00}(\mathbb{1}_2)] |0\rangle_c \langle 0| + \tfrac{1}{2} \text{Tr}[S_{01}(\mathbb{1}_2)] |0\rangle_c \langle 1| \\ &\quad + \tfrac{1}{2} \text{Tr}[S_{01}^\dagger(\mathbb{1}_2)] |1\rangle_c \langle 0| + \tfrac{1}{2} \text{Tr}[S_{11}(\mathbb{1}_2)] |1\rangle_c \langle 1|. \end{aligned} \quad (28)$$

The Bloch vector of the control qubit is given by

$$\vec{r}_{\text{out},c}^s = \frac{1}{2} \begin{pmatrix} \text{Tr}[S_{01}(\mathbb{1}_2)] \langle 0|\rho_c|1\rangle + \text{Tr}[S_{01}^\dagger(\mathbb{1}_2)] \langle 1|\rho_c|0\rangle \\ i \text{Tr}[S_{01}(\mathbb{1}_2)] \langle 0|\rho_c|1\rangle - i \text{Tr}[S_{01}^\dagger(\mathbb{1}_2)] \langle 1|\rho_c|0\rangle \\ \text{Tr}[S_{00}(\mathbb{1}_2)] \langle 0|\rho_c|0\rangle - \text{Tr}[S_{11}(\mathbb{1}_2)] \langle 1|\rho_c|1\rangle \end{pmatrix}. \quad (29)$$

Because $\text{Tr}[S_{01}(\mathbb{1}_2)] = 2Q(\xi, p_I, p_x, p_y, p_z)$,

$$\vec{r}_{\text{out},c}^s = \begin{pmatrix} 2\sqrt{p_c(1-p_c)}Q(\xi, p_I, p_x, p_y, p_z) \\ 0 \\ 2p_c - 1 \end{pmatrix}. \quad (30)$$

According to Eq. (18), the quantum Fisher information of the control qubit is given by

$$\begin{aligned} \mathcal{F}_{\mathcal{Q}}^s(\xi) &= \frac{16p_c^2(1-p_c)^2Q^2[\partial_\xi Q]^2}{1-4p_c(1-p_c)Q^2} + 4p_c(1-p_c)(\partial_\xi Q)^2 \\ &= \frac{4p_c(1-p_c)[\partial_\xi Q(\xi, p_I, p_x, p_y, p_z)]^2}{1-4p_c(1-p_c)Q^2(\xi, p_I, p_x, p_y, p_z)}. \end{aligned} \quad (31)$$

The quantum Fisher information equals the classical Fisher information in Eq. (26), which means that our projection measurement on the control qubit is optimal.

IV. EXPERIMENTAL DETAILS

A. State preparation

An arbitrary polarization pure state can be generated using a quarter waveplate and a half waveplate such that

$$\hat{H}(\theta)\hat{Q}(\nu)|H\rangle \rightarrow \rho = |\psi\rangle\langle\psi|. \quad (32)$$

The Bloch vector of ρ is

$$\vec{r} = [\sin(4\theta - 2\nu)\cos(2\nu), \sin(2\nu), \cos(4\theta - 2\nu)\cos(2\nu)], \quad (33)$$

where $\hat{H}(\theta)$ and $\hat{Q}(\nu)$ refer to the half waveplate (HWP) and quarter waveplate (QWP), respectively, and θ and ν represent the angles between the fast axis of the waveplate and the horizontal polarization direction. To implement an SU(2) transformation and Pauli operators, we adopt an SU(2) gadget [35,36]. By cascading $\hat{Q}(\tau_3)\hat{H}(\tau_2)\hat{Q}(\tau_1)$ wave plates and adjusting $\{\tau_3, \tau_2, \tau_1\}$, an arbitrary SU(2) transformation can be achieved (refer to Appendix F for details).

B. Experimental setup and calibration

In our experiment, we encode the spatial mode of the photon in each path of the interferometer as the control qubit and the polarization degree of freedom as the probe qubit carrying the estimated parameter. By exploiting the incoherent property of the quantum channel, we implement the general Pauli channel by performing the four Pauli operators separately in each path, resulting in 16 different experimental configurations. We then combine the experimental outcomes with the classical probability vector $\vec{p} = p_I, p_x, p_y, p_z$ and measure the control qubit in the $|\pm\rangle_c$ basis by carefully adjusting the phase of the interferometer using a piezoelectric transducer.

To achieve precision estimation of the unknown parameter and its variance, we carefully count the number of photons N injected into the interferometer, accounting for all loss and detection efficiency. To satisfy the asymptotic reachability of the Cramér bound, a sufficiently large N is required. In our experiment, N is chosen to be 1000 using Monte Carlo simulation and we sample with probability distribution \vec{p} to determine the number of photons n_1, \dots, n_{16} injected into all 16 combinations. We record the time tags of all detected photons for each experimental trial and extract the values

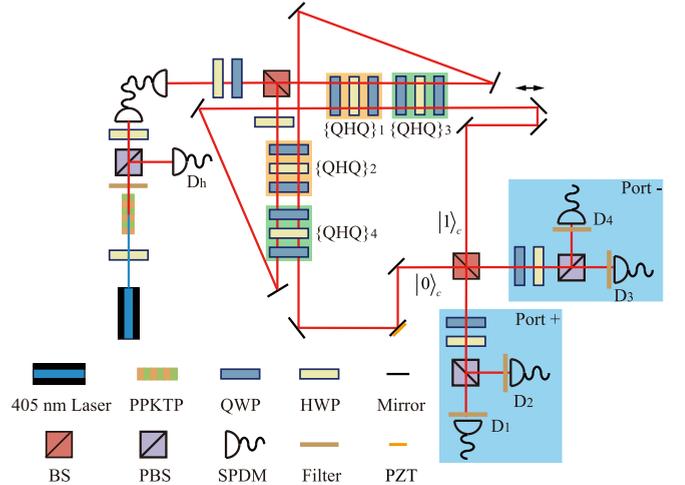


FIG. 2. We utilize a continuous-wave 405-nm laser diode to pump a 3-mm periodically poled potassium titanyl phosphate (PPKTP) crystal for the generation of degenerate photon pairs centered at 810 nm via type-II spontaneous parametric down-conversion. One of the photons acts as a trigger to herald the signal photon, which is fed into a Mach-Zehnder interferometer. The two paths of the interferometer are encoded as control qubits $|0\rangle_c$ and $|1\rangle_c$, respectively. The unknown parameter ξ is initially encoded on the polarization degree of freedom of the signal photon using the first set of QWP, HWP, and QWP (QHQ) in the orange shade in each path. Then the second set of QHQ in the green shade implements Pauli operators. A piezoelectric transducer is used to control the relative phase of the interferometer. Finally, the control qubit is measured on the $|\pm\rangle$ basis by a polarization-maintaining single-mode fiber and a single-photon detection module (SPDM) and detected by DPC-230 TCSPC. Here BS denotes beam splitter and PBS polarization beam splitter.

n_1, \dots, n_{16} for analysis. The experimental trial number is 200. We also provide an analysis of the probability distribution for each single experiment in Appendix E.

First, we calibrate the detection efficiency by blocking the reflection beam of the first beam splitter and setting the four detection probabilities equal to 0.25 theoretically. Since only the output photon counts of the second beam splitter matter, we introduce two virtual detectors D_+ and D_- , which register photons ejected from two ports of the second beam splitter, port + and port -, respectively. By tuning the relative phase of the interferometer with a piezoelectric transducer, D_+ is in the interference-enhanced path with the total counts of D_1 and D_2 , while D_- is in the interference-suppressed path with the total counts of D_3 and D_4 . We denote the five detectors in Fig. 2 by D_h, D_1, D_2, D_3 , and D_4 in counterclockwise order. Adopting a postselection manner, we measure the efficiency of D_+ and D_- as $\eta_+ = 7.95\% \pm 0.18\%$ and $\eta_- = 7.49\% \pm 0.17\%$, respectively.

We verify the validity of utilizing discrete Pauli operators to simulate Pauli channels. This involves setting all waveplates in the $\{\hat{Q}\hat{H}\hat{Q}\}_1, \{\hat{Q}\hat{H}\hat{Q}\}_2$, and $\{\hat{Q}\hat{H}\hat{Q}\}_4$ combination to 0° . The $\{\hat{Q}\hat{H}\hat{Q}\}_3$ combination is used to perform Pauli operators. By postprocessing the coincidence counts using a classical probability distribution (p_I, p_x, p_y, p_z) , we can simulate a general Pauli channel. The channel is characterized

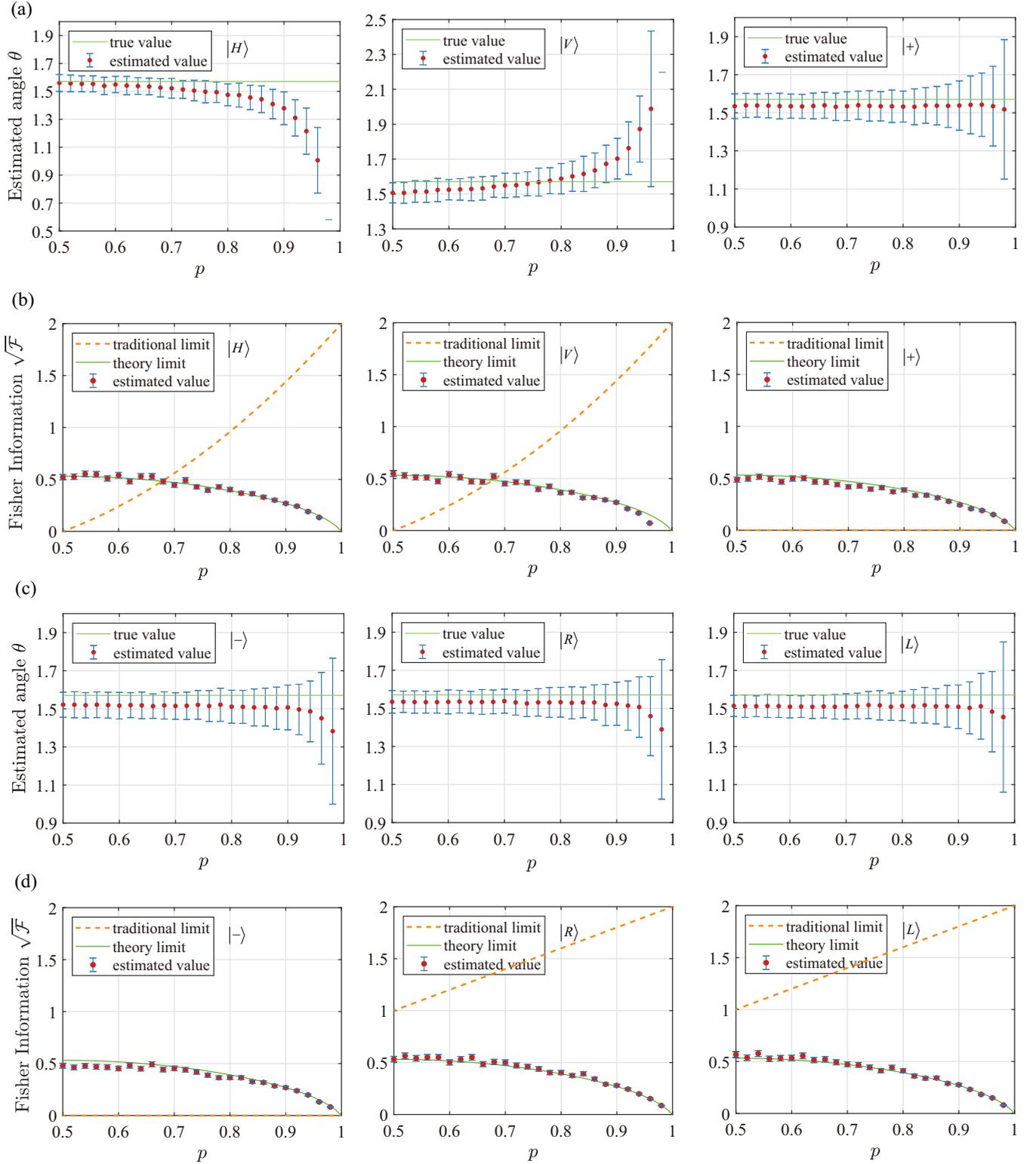


FIG. 3. Estimated value and Fisher information for various probe states in the phase-flip channel. The true value of ξ is $\pi/2$ and the rotation axis $\vec{n} = (1, 0, 0)^T$. The theory limit of quantum Fisher information $\sqrt{\mathcal{F}}$ is calculated with Eq. (26) taking experimental imperfection into consideration, while the estimated value of $\sqrt{\mathcal{F}}$ is calculated with Eq. (G20) using experimental data. The traditional limit is calculated with Eq. (18) under two cascaded phase-flip channels depicted in Fig. 1(a). (a) Estimated value of ξ for the $|H\rangle$, $|V\rangle$, and $|+\rangle$ probe states. (b) Square root of the Fisher information per photon for the $|H\rangle$, $|V\rangle$, and $|+\rangle$ probe states. (c) Estimated value of ξ for the $|-\rangle$, $|R\rangle$, and $|L\rangle$ probe states. (d) Square root of the Fisher information per photon for the $|-\rangle$, $|R\rangle$, and $|L\rangle$ probe states. Here $|+\rangle = 1/\sqrt{2}(|H\rangle + |V\rangle)$, $|-\rangle = 1/\sqrt{2}(|H\rangle - |V\rangle)$, $|R\rangle = 1/\sqrt{2}(|H\rangle + i|V\rangle)$, and $|L\rangle = 1/\sqrt{2}(|H\rangle - i|V\rangle)$.

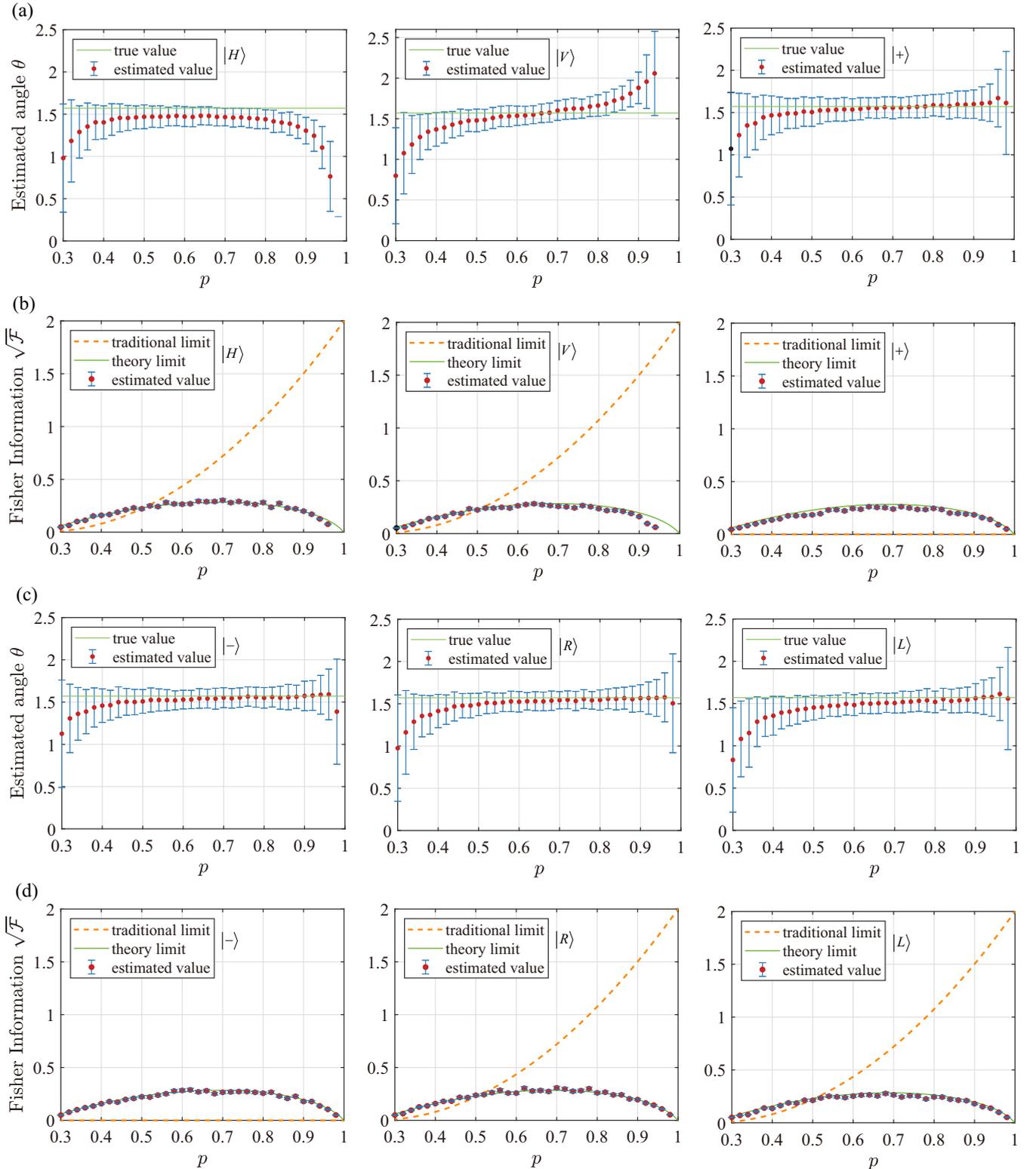


FIG. 4. Estimated value and Fisher information for various probe states in the depolarization channel. The true value of ξ is $\pi/2$ and the rotation axis $\vec{n} = (1, 0, 0)^T$. The theory limit of quantum Fisher information $\sqrt{\mathcal{F}}$ is calculated with Eq. (26) taking experimental imperfection into consideration, while the estimated value of $\sqrt{\mathcal{F}}$ is calculated with Eq. (G20) using experimental data. The traditional limit is calculated with Eq. (18) under the two cascaded depolarization channels depicted in Fig. 1(a). (a) Estimated value of ξ for the $|H\rangle$, $|V\rangle$, and $|+\rangle$ probe states. (b) Square root of the Fisher information per photon for the $|H\rangle$, $|V\rangle$, and $|+\rangle$ probe states. (c) Estimated value of ξ for the $|-\rangle$, $|R\rangle$, and $|L\rangle$ probe states. (d) Square root of the Fisher information per photon for the $|-\rangle$, $|R\rangle$, and $|L\rangle$ probe states. Here $|+\rangle = 1/\sqrt{2}(|H\rangle + |V\rangle)$, $|-\rangle = 1/\sqrt{2}(|H\rangle - |V\rangle)$, $|R\rangle = 1/\sqrt{2}(|H\rangle + i|V\rangle)$, and $|L\rangle = 1/\sqrt{2}(|H\rangle - i|V\rangle)$.

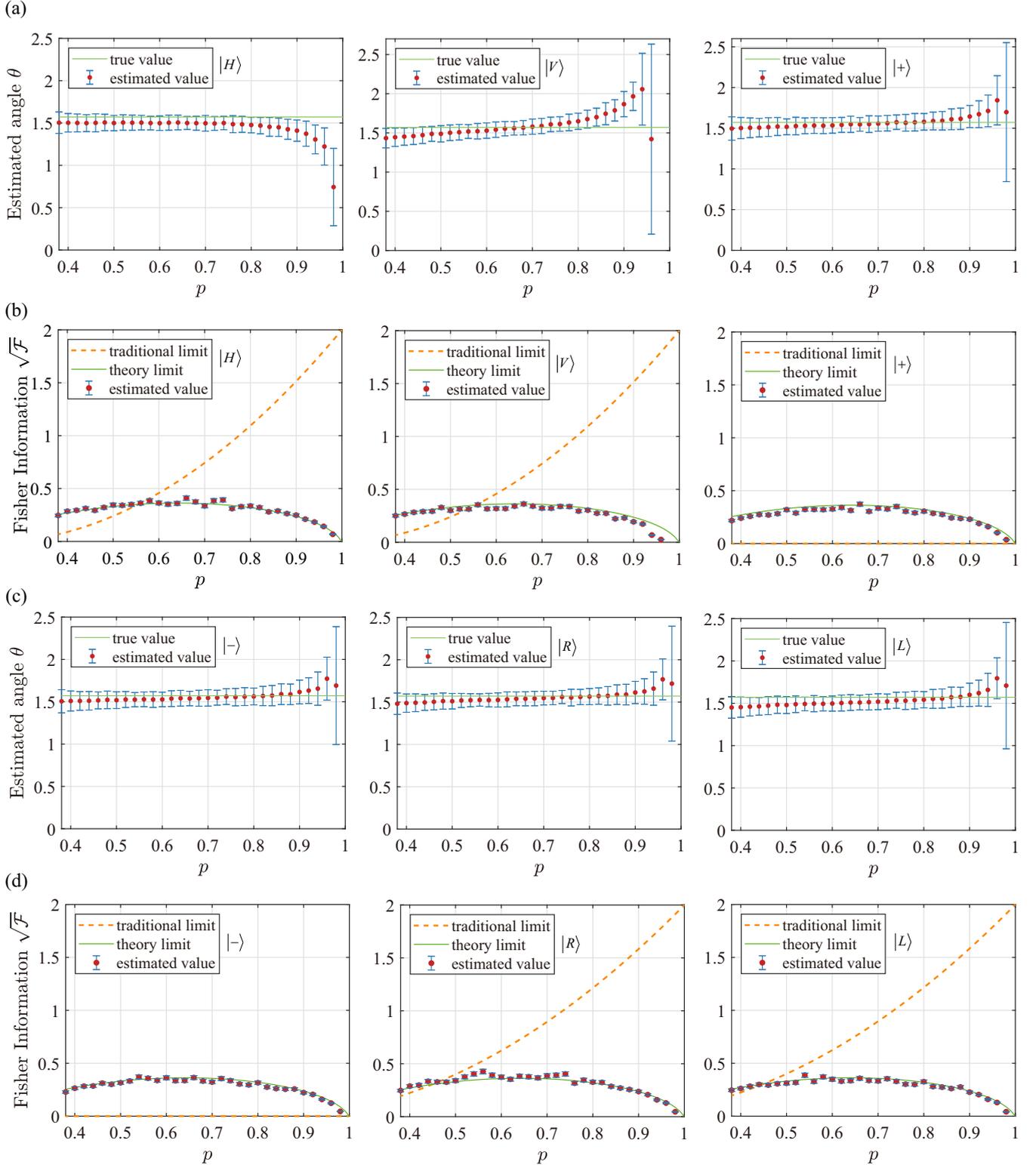


FIG. 5. Estimated value and Fisher information for various probe states in the arbitrary Pauli channel $\mathcal{N}_1(\rho) = p\rho + \frac{1-p}{5}\sigma_x\rho\sigma_x + \frac{3(1-p)}{10}\sigma_y\rho\sigma_y + \frac{1-p}{2}\sigma_z\rho\sigma_z$. The true value of ξ is $\pi/2$ and the rotation axis $\vec{n} = (1, 0, 0)^T$. The theory limit of the quantum Fisher information $\sqrt{\mathcal{F}}$ is calculated with Eq. (26) taking experimental imperfection into consideration, while the estimated value of $\sqrt{\mathcal{F}}$ is calculated with Eq. (G20) using experimental data. The traditional limit is calculated with Eq. (18) under the two cascaded general Pauli channels $\mathcal{N}_1(\rho)$ depicted in Fig. 1(a). (a) Estimated value of ξ for the $|H\rangle$, $|V\rangle$, and $|+\rangle$ probe states. (b) Square root of the Fisher information per photon for the $|H\rangle$, $|V\rangle$, and $|+\rangle$ probe states. (c) Estimated value of ξ for the $|-\rangle$, $|R\rangle$, and $|L\rangle$ probe states. (d) Square root of the Fisher information per photon for the $|-\rangle$, $|R\rangle$, and $|L\rangle$ probe states. Here $|+\rangle = 1/\sqrt{2}(|H\rangle + |V\rangle)$, $|-\rangle = 1/\sqrt{2}(|H\rangle - |V\rangle)$, $|R\rangle = 1/\sqrt{2}(|H\rangle + i|V\rangle)$, and $|L\rangle = 1/\sqrt{2}(|H\rangle - i|V\rangle)$.

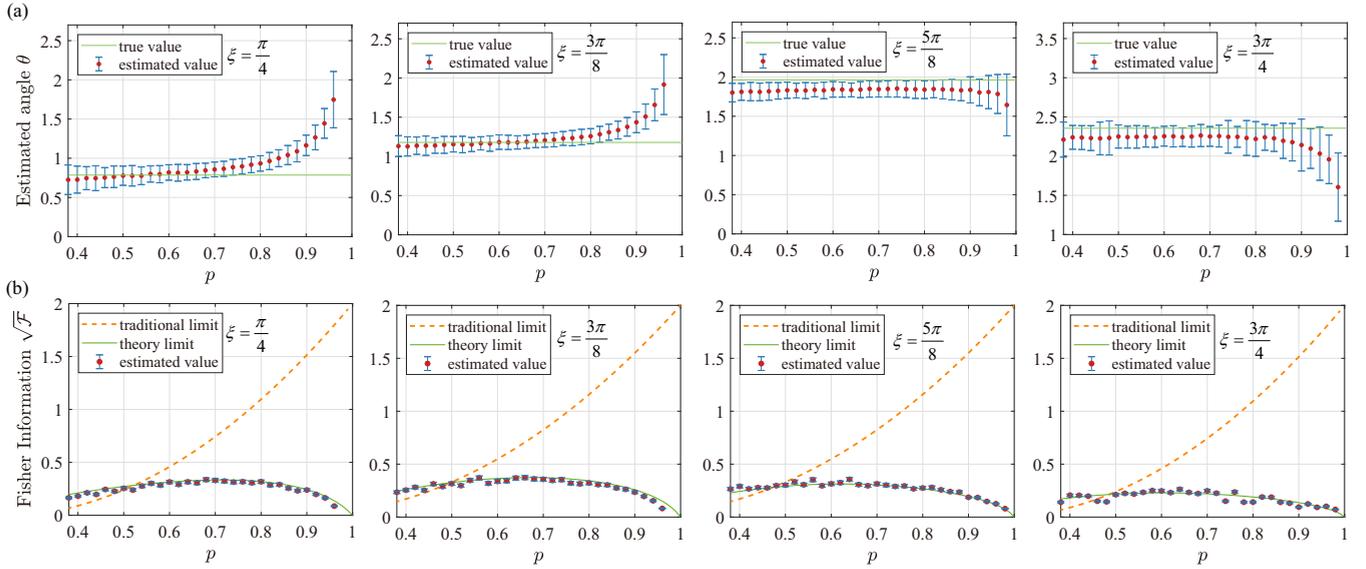


FIG. 6. (a) Estimated value and (b) Fisher information for various values of the parameter ξ in the general Pauli channel $\mathcal{N}_1(\rho) = p\rho + \frac{1-p}{5}\sigma_x\rho\sigma_x + \frac{3(1-p)}{10}\sigma_y\rho\sigma_y + \frac{1-p}{2}\sigma_z\rho\sigma_z$. The probe state is $|R\rangle = 1/\sqrt{2}(|H\rangle + i|V\rangle)$ and the rotation axis $\vec{n} = (1, 0, 0)$.

using quantum process tomography and the average quantum channel fidelity exceeds 95%.

In our experiment, we set $\vec{n} = (1, 0, 0)^T$ and $\xi = \pi/2$, which is achieved using the $\hat{Q}(\frac{\pi}{2})\hat{H}(-\frac{3\pi}{8})\hat{Q}(\frac{\pi}{2})$ combination in the orange shade in each path of Fig. 2. In the beginning, we characterize the interferometer visibility by setting the relative phase of two paths at 0 and π . When the input state is $|H\rangle$, we obtain a lowest visibility around 91.32%. This result is attributed to the polarization dependence of the second beam splitter. The $|H\rangle$ photons in the $|1\rangle_c$ path experience a slight loss upon transmission through the second beam splitter, while the reflection rate exceeds the transmission rate in both input ports. The imperfect device may introduce systematic errors in the estimation task, which can be reduced by introducing calibrated parameters to the estimators. (See Appendix G for more experimental details and the calibrated process.)

C. Experimental result

To show that our protocol is independent of the probe state, we prepared the polarization states $|H\rangle, |V\rangle, |+\rangle = 1/\sqrt{2}(|H\rangle + |V\rangle)$, $|-\rangle = 1/\sqrt{2}(|H\rangle - |V\rangle)$, $|R\rangle = 1/\sqrt{2}(|H\rangle + i|V\rangle)$, and $|L\rangle = 1/\sqrt{2}(|H\rangle - i|V\rangle)$ and utilized the phase-flip channel, depolarization channel, and arbitrary Pauli channel in the quantum switch, respectively. We define $p_I = p$ in the following discussion. The experiment results for the phase-flip channel with $p_x = p_y = 0$ and $p_z = 1 - p$ are shown in Fig. 3. We estimated ξ and extracted the Fisher information per photon at different noise levels. Figures 3(a) and 3(c) display the estimated value of ξ for the probe states $|H\rangle, |V\rangle, |+\rangle, |-\rangle, |R\rangle$, and $|L\rangle$, while Figs. 3(b) and 3(d) show the square root of the Fisher information per photon. We observe that the estimated value is in good agreement with the true value of $\pi/2$, and the Fisher information is consistent with the theoretical value. Furthermore, we find that the estimated value and Fisher information are insensitive

to the probe states. When the noise level is high, the estimated value is close to the true value and the Fisher information get larger, indicating a better performance at high noise levels.

It is important to note that a discrepancy between the estimated value and true value exists as p gradually increases to 1, and the adopted approach becomes invalid when $p = 1$. We emphasize that noise is crucial in our experimental scheme. When $p = 1$, the two quantum noise channels degenerate into a unitary transformation, making it impractical to evaluate ξ . The observed discrepancy between the experimentally estimated value and true value is due to the limited visibility of the interferometer. This can be addressed by revising the estimator, as described in Appendix G 2. There is a little photon loss in the second beam splitter, which is dependent on polarization. After calibrating our setup with the probe state $|R\rangle = 1/\sqrt{2}(|H\rangle + i|V\rangle)$, all the superpositions of $|H\rangle$ and $|V\rangle$ work well, while the behaviors of $|H\rangle$ and $|V\rangle$ are in the opposite direction when p tends to 1.

The standard deviation of the estimated value σ/\sqrt{N} reaches the same classical bound $\sqrt{\mathcal{F}}$ for various probe states, demonstrating that our approach is independent of the probe state. Even when the probe state $|+\rangle$ is parallel to the rotation axis $\hat{n} = (1, 0, 0)^T$, the traditional methods fail; however, our approach is still effective. When the probe state is $|H\rangle$ or $|V\rangle$, our approach outperforms the traditional methods in a high level of noise ($p < 0.6893$).

Our approach only requires collecting photon counts from the two ports of the interferometer, and the information carried by the polarization degree of freedom is discarded, which may decrease the precision of the parameter ξ estimation. However, this simplifies the experimental implementation, as it does not require parameter-dependent optimal measurement setups. The projection measurement $|\pm\rangle\langle\pm|_c$ on the control qubit is deterministic and can achieve the quantum Cramér-Rao bound (QCRB). It is also important to note that the phase stability of the interferometer is crucial for our experiment.

Since the experiment has a combination property, we can adjust the relative phase to 0 before each single experiment. The collection time for each single experiment lasts for 10 s and the phase drift can be negligible.

We also performed experiments with different values of ξ ranging from 0 to π in order to eliminate systematic errors caused by imperfect optical elements and limited visibility. To achieve this, we introduced calibrated parameters α , β , and ξ_e . Using the least-squares method, we determined the parameters for the phase-flip, depolarization, and arbitrary Pauli channels by minimizing the discrepancy between the experimental probability distribution and the theoretical probability distribution at different ξ and p . Details of this process are shown in Appendix G 2. For the depolarization channel with $p_x = p_y = p_z = (1 - p)/3$, the experimental results are shown in Fig. 4. It is notable that when depolarization noise is at a high level ($p < 0.5205$), the quantum switch scenario will outperform the traditional approach, which exhibits robustness to noise. For the arbitrary Pauli channel \mathcal{N}_1 with $p_x = (1 - p)/5$, $p_y = 3(1 - p)/10$, and $p_z = (1 - p)/2$, the experimental results are shown in Fig. 5 for various probe states. The quantum switch outperforms the traditional strategy when $0.375 \leq p \leq 0.5626$ for $|H\rangle, |V\rangle$ and $0.375 \leq p \leq 0.4590$ for $|R\rangle, |L\rangle$. Figure 6 shows the estimation for various values of ξ with the probe state $|R\rangle$. There are some systematic errors in the experiment caused by the imperfect optical elements.

V. CONCLUSION

In summary, we have expanded upon the theoretical proposal presented in Ref. [32] by applying it to the SU(2) general Pauli channels. We have analyzed theoretically the origin of the probe-state-independent property in the estimation precision and established a criterion to avoid precisely preparing the probe state. We have shown that the specific Kraus representations of the channels have no impact on the estimation precision and probe-state-independent property. This provides an experimentally friendly way to implement quantum switch channels. We have verified experimentally the advantages of using an indefinite causal order scheme on a photonics platform. By utilizing three special cases of general Pauli channels, namely, the phase-flip channel, depolarization channel, and arbitrary Pauli channel, we have successfully demonstrated the SU(2) phase estimation in a noisy environment by implementing a switch channel. Our approach achieves estimation precision that can reach the quantum Cramér-Rao bound and exhibits robustness at high noise levels. Notably, our approach allows for arbitrary probe states and deterministic measurement settings, making it more feasible for real-world physical systems compared to conventional SU(2) phase-estimation tasks [33]. It is worth noting that that conventional cascaded channel scheme and switch channel scheme in our paper are special cases of the adaptive feedback and casual superposition scheme,s respectively, in Ref. [30]. Our result is consistent with the general single-parameter-estimation theory in Ref. [30]. Furthermore, our work can be extended to multiparameter noisy metrology [37], which would provide insight into easy-to-implement schemes in quantum metrology.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China (Grants No. 12074307, No. 11804271, No. 91736104, and No. 12204371), Shaanxi Fundamental Science Research Project for Mathematics and Physics (Grant No. 22JSZ004), China Postdoctoral Science Foundation via Project No. 2020M673366, and Shaanxi Natural Science Basic Research Program (Grant No. 2021JQ-008).

APPENDIX A: FISHER INFORMATION

The Fisher information, denoted by $\mathcal{F}(\xi)$, quantifies the amount of information present in an observed data set about an unknown parameter ξ in a statistical model. It serves as an upper bound on the precision of estimating the parameter, as stated by the Cramér-Rao bound,

$$\text{var}(\hat{\xi}) \geq \frac{1}{\mathcal{F}(\xi)}, \quad (\text{A1})$$

where $\hat{\xi}$ represents an unbiased estimator of the unknown parameter ξ and the variance of $\hat{\xi}$ is constrained by the reciprocal of the Fisher information $\mathcal{F}(\xi)$. The classical Fisher information, denoted by \mathcal{F}_c , is defined for a probability density function $f(x|\xi)$ that depends on an unknown parameter ξ . It is given by the integral expression

$$\mathcal{F}_c = \int \frac{[\partial_\xi f(x|\xi)]^2}{f(x|\xi)} dx, \quad (\text{A2})$$

where x represents the classical measurement value and $f(x|\xi)$ denotes the conditional probability of obtaining x given the value of ξ .

In quantum parameter-estimation tasks, the choice of POVMs M_{mi} can affect the conditional probability density functions $f_i(x|\xi)$ and consequently the Fisher information. To obtain as much information as possible, we aim to maximize the Fisher information over all possible POVMs, leading to the QCRB

$$\text{var}(\hat{\xi}) \geq \frac{1}{\mathcal{F}_Q(\xi)}, \quad (\text{A3})$$

where $\hat{\xi}$ represents an unbiased estimator of the unknown parameter ξ and the denominator $\mathcal{F}_Q(\xi)$ denotes the quantum Fisher information (QFI), which is the maximum attainable Fisher information. Achieving the QCRB requires finding an experimentally feasible POVM that can yield the maximum QFI, allowing for optimal precision in estimating the parameter ξ from the quantum data.

APPENDIX B: PROOF OF KRAUS REPRESENTATION INDEPENDENCE IN THE SWITCH CHANNEL

We prove that the phase-estimation precision is independent of the Kraus representation of quantum channels. The Kraus operators of the switch channel are

$$K_{ij} = K_j^{(2)} K_i^{(1)} \otimes |0\rangle\langle 0|_c + K_i^{(1)} K_j^{(2)} \otimes |1\rangle\langle 1|_c. \quad (\text{B1})$$

If we choose two different sets of Kraus operators $\{K_i^{(1)}\}$ and $\{E_i^{(1)}\}$ of channel (1), $\{K_i^{(1)}\}$ and $\{E_i^{(1)}\}$ are related to a unitary transformation R , $K_i^{(1)} = \sum_l R_{il} E_l^{(1)}$.

When adopting two identical channels, because of $S_{00}(\rho) = S_{11}(\rho)$ and $S_{01}(\rho) = S_{01}^\dagger(\rho)$, we only need to prove that $S_{00}(\rho)$ and $S_{01}(\rho)$ are independent of the Kraus representation of channel (1),

$$\begin{aligned}
S_{00}(\rho) &= \sum_{i,j} \left[\left(\sum_l K_j^{(2)} R_{il} E_l^{(1)} \right) \rho \left(\sum_m E_m^{(1)\dagger} R_{im}^* K_j^{(2)\dagger} \right) \right] \\
&= \sum_{i,j} \left[\left(\sum_{l=m} K_j^{(2)} R_{il} E_l^{(1)} \rho E_l^{(1)\dagger} R_{il}^* K_j^{(2)\dagger} \right) \right] \\
&\quad + \sum_{i,j} \left[\left(\sum_{l \neq m} K_j^{(2)} R_{il} E_l^{(1)} \right) \rho \left(\sum_{m \neq l} E_m^{(1)\dagger} R_{im}^* K_j^{(2)\dagger} \right) \right] \\
&= \sum_{j,l} \left(\sum_i R_{il} R_{il}^* K_j^{(2)} E_l^{(1)} \rho E_l^{(1)\dagger} K_j^{(2)\dagger} \right) \\
&\quad + \sum_{i,j} \left(\sum_{l,m,l \neq m} R_{il} R_{im}^* K_j^{(2)} E_l^{(1)} \rho E_m^{(1)\dagger} K_j^{(2)\dagger} \right) \\
&= \sum_{j,l} (K_j^{(2)} E_l^{(1)} \rho E_l^{(1)\dagger} K_j^{(2)\dagger}) \\
&\quad + \sum_j \left[\sum_{l,m,l \neq m} \left(\sum_i R_{il} R_{im}^* \right) K_j^{(2)} E_l^{(1)} \rho E_m^{(1)\dagger} K_j^{(2)\dagger} \right] \\
&= \sum_{j,l} (K_j^{(2)} E_l^{(1)} \rho E_l^{(1)\dagger} K_j^{(2)\dagger}), \tag{B2}
\end{aligned}$$

$$\begin{aligned}
S_{01}(\rho) &= \sum_{i,j} \left[\left(\sum_l K_j^{(2)} R_{il} E_l^{(1)} \right) \rho \left(\sum_m K_j^{(2)\dagger} E_m^{(1)\dagger} R_{im}^* \right) \right] \\
&= \sum_{i,j} \left[\left(\sum_{l=m} K_j^{(2)} R_{il} E_l^{(1)} \rho K_j^{(2)\dagger} E_l^{(1)\dagger} R_{il}^* \right) \right] \\
&\quad + \sum_{i,j} \left[\left(\sum_{l \neq m} K_j^{(2)} R_{il} E_l^{(1)} \right) \rho \left(\sum_{m \neq l} K_j^{(2)\dagger} E_m^{(1)\dagger} R_{im}^* \right) \right] \\
&= \sum_{j,l} \left(\sum_i R_{il} R_{il}^* K_j^{(2)} E_l^{(1)} \rho K_j^{(2)\dagger} E_l^{(1)\dagger} \right) \\
&\quad + \sum_{i,j} \left(\sum_{l,m,l \neq m} R_{il} R_{im}^* K_j^{(2)} E_l^{(1)} \rho K_j^{(2)\dagger} E_m^{(1)\dagger} \right) \\
&= \sum_{j,l} (K_j^{(2)} E_l^{(1)} \rho K_j^{(2)\dagger} E_l^{(1)\dagger}) \\
&\quad + \sum_j \left[\sum_{l,m,l \neq m} \left(\sum_i R_{il} R_{im}^* \right) K_j^{(2)} E_l^{(1)} \rho K_j^{(2)\dagger} E_m^{(1)\dagger} \right] \\
&= \sum_{j,l} (K_j^{(2)} E_l^{(1)} \rho K_j^{(2)\dagger} E_l^{(1)\dagger}), \tag{B3}
\end{aligned}$$

where we used the fact that $\sum_i R_{il} R_{im}^* = \delta_{lm}$. We can find that the Kraus operators of the switch channel are independent of the Kraus representation of quantum channel (1). Since channels (1) and (2) are mathematically equivalent in the above

procedure, we prove that the output of the switch channel is independent of the Kraus representation of the quantum channel, which also implies that the phase-estimation precision is irrelevant to the Kraus representation of quantum channels.

APPENDIX C: PROBE-STATE-INDEPENDENT CRITERION

The probe-state-independent criterion is described as follows. When adopting two identical channels, i.e., $K_i^{(1)} = K_i^{(2)} = K_i$, define $O_{ij} = [K_i^{(1)}, K_j^{(2)}]$ and $\mathcal{O} = \sum_{i,j,i < j} O_{ij}^\dagger O_{ij}$. If the SU(2) phase estimation with the indefinite causal order dynamics is probe-state independent, the condition

$$\mathcal{O} = c\mathbb{I}_2, \quad c \in \mathbb{R}, \tag{C1}$$

must be held.

The probe-state-independent property of the switch channel relates to the noncommutativity of the Kraus operators of the quantum channel. We only need to prove that both $\text{Tr}[S_{00}(\rho_{\text{in}})]$ and $\text{Tr}[S_{01}(\rho_{\text{in}})]$ are independent of the probe state \vec{r} because of $\text{Tr}[S_{11}(\rho_{\text{in}})] = \text{Tr}[S_{00}(\rho_{\text{in}})]$ and $\text{Tr}[S_{01}^\dagger(\rho_{\text{in}})] = \text{Tr}[S_{01}(\rho_{\text{in}})]$. The details are as follows. For $S_{00}(\rho_{\text{in}})$,

$$\begin{aligned}
S_{00}(\rho_{\text{in}}) &= \frac{1}{2} [S_{00}(\mathbb{1}_2)] + \frac{1}{2} [S_{00}(\vec{r} \cdot \vec{\sigma})] \\
&= \frac{1}{2} \sum_{i,j} K_j^{(2)} K_i^{(1)} K_i^{(1)\dagger} K_j^{(2)\dagger} \\
&\quad + \frac{1}{2} \sum_{i,j} K_j^{(2)} K_i^{(1)} (\vec{r} \cdot \vec{\sigma}) K_i^{(1)\dagger} K_j^{(2)\dagger}. \tag{C2}
\end{aligned}$$

Utilizing $\sum K_i^{(1)\dagger} K_i^{(1)} = \sum K_j^{(2)\dagger} K_j^{(2)} = \mathbb{1}_2$, we obtain

$$\begin{aligned}
\text{Tr}[S_{00}(\mathbb{1}_2)] &= \text{Tr} \left(\sum_{i,j} K_j^{(2)} K_i^{(1)} K_i^{(1)\dagger} K_j^{(2)\dagger} \right) \\
&= \text{Tr}(\mathbb{1}_2) \\
&= 2, \tag{C3}
\end{aligned}$$

$$\begin{aligned}
\text{Tr}[S_{00}(\vec{r} \cdot \vec{\sigma})] &= \text{Tr} \left(\sum_{i,j} K_j^{(2)} K_i^{(1)} (\vec{r} \cdot \vec{\sigma}) K_i^{(1)\dagger} K_j^{(2)\dagger} \right) \\
&= \text{Tr}[(\vec{r} \cdot \vec{\sigma})] \\
&= 0. \tag{C4}
\end{aligned}$$

It is obvious that $\text{Tr}[S_{00}(\rho_{\text{in}})]$ is independent of the probe state \vec{r} . For $S_{01}(\rho_{\text{in}})$,

$$\begin{aligned}
S_{01}(\rho_{\text{in}}) &= \frac{1}{2} S_{01}(\mathbb{1}_2) + \frac{1}{2} S_{01}(\vec{r} \cdot \vec{\sigma}) \\
&= \frac{1}{2} \sum_{i,j} K_j^{(2)} K_i^{(1)} K_j^{(2)\dagger} K_i^{(1)\dagger} \\
&\quad + \frac{1}{2} \sum_{i,j} K_j^{(2)} K_i^{(1)} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger}. \tag{C5}
\end{aligned}$$

An arbitrary two-dimensional matrix can be expanded in the Pauli basis, defining $\sum K_i^{(1)} K_i^{(1)\dagger} = a\mathbb{1}_2 + \vec{k}_1 \cdot \vec{\sigma}$ and $\sum K_j^{(2)} K_j^{(2)\dagger} = b\mathbb{1}_2 + \vec{k}_2 \cdot \vec{\sigma}$. If $\sum K_i^{(1)} K_i^{(1)\dagger} = \sum K_j^{(2)} K_j^{(2)\dagger}$

$= \mathbb{1}_2$, channels 1 and 2 are called unital channels:

$$\begin{aligned} S_{01}(\mathbb{1}_2) &= \sum_{i,j} K_j^{(2)} K_i^{(1)} K_j^{(2)\dagger} K_i^{(1)\dagger} \\ &= \sum_{i,j} (K_i^{(1)} K_j^{(2)} - O_{ij}) K_j^{(2)\dagger} K_i^{(1)\dagger} \\ &= \sum_{i,j} K_i^{(1)} K_j^{(2)} K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j} O_{ij} K_j^{(2)\dagger} K_i^{(1)\dagger}. \quad (\text{C6}) \end{aligned}$$

When adopting two identical channels, i.e., $K_i^{(1)} = K_i^{(2)} = K_i$,

$$\begin{aligned} & - \sum_{i,j} O_{ij} K_j^{(2)\dagger} K_i^{(1)\dagger} \\ &= - \sum_{i,j,i<j} O_{ij} K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,i>j} O_{ij} K_j^{(2)\dagger} K_i^{(1)\dagger} \end{aligned}$$

$$\begin{aligned} &= - \sum_{i,j,i<j} O_{ij} K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,j>i} O_{ji} K_i^{(2)\dagger} K_j^{(1)\dagger} \\ &= - \sum_{i,j,i<j} O_{ij} K_j^{(2)\dagger} K_i^{(1)\dagger} + \sum_{i,j,j>i} O_{ij} K_i^{(2)\dagger} K_j^{(1)\dagger} \\ &= - \sum_{i,j,i<j} O_{ij} (K_j^{(2)\dagger} K_i^{(1)\dagger} - K_i^{(2)\dagger} K_j^{(1)\dagger}) \\ &= - \sum_{i,j,i<j} O_{ij} (K_j^{(2)\dagger} K_i^{(1)\dagger} - K_i^{(1)\dagger} K_j^{(2)\dagger}) \\ &= - \sum_{i,j,i<j} O_{ij} [K_j^{(2)\dagger}, K_i^{(1)\dagger}] \\ &= - \sum_{i,j,i<j} O_{ij} O_{ij}^\dagger, \quad (\text{C7}) \end{aligned}$$

$$\begin{aligned} S_{01}(\vec{r} \cdot \vec{\sigma}) &= \sum_{i,j} K_j^{(2)} K_i^{(1)} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} \\ &= \sum_{i,j} (K_i^{(1)} K_j^{(2)} - O_{ij}) (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} \\ &= \sum_{i,j} K_i^{(1)} K_j^{(2)} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j} O_{ij} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} \\ &= \sum_{i,j} K_i^{(1)} K_j^{(2)} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,i<j} O_{ij} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,i>j} O_{ij} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} \\ &= \sum_{i,j} K_i^{(1)} K_j^{(2)} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,i<j} O_{ij} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,j>i} O_{ji} (\vec{r} \cdot \vec{\sigma}) K_i^{(2)\dagger} K_j^{(1)\dagger} \\ &= \sum_{i,j} K_i^{(1)} K_j^{(2)} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,i<j} O_{ij} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} + \sum_{i,j,j>i} O_{ij} (\vec{r} \cdot \vec{\sigma}) K_i^{(2)\dagger} K_j^{(1)\dagger} \\ &= \sum_{i,j} K_i^{(1)} K_j^{(2)} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,i<j} O_{ij} (\vec{r} \cdot \vec{\sigma}) (K_j^{(2)\dagger} K_i^{(1)\dagger} - K_i^{(2)\dagger} K_j^{(1)\dagger}) \\ &= \sum_{i,j} K_i^{(1)} K_j^{(2)} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,i<j} O_{ij} (\vec{r} \cdot \vec{\sigma}) (K_j^{(2)\dagger} K_i^{(1)\dagger} - K_i^{(1)\dagger} K_j^{(2)\dagger}) \\ &= \sum_{i,j} K_i^{(1)} K_j^{(2)} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,i<j} O_{ij} (\vec{r} \cdot \vec{\sigma}) [K_j^{(2)\dagger}, K_i^{(1)\dagger}] \\ &= \sum_{i,j} K_i^{(1)} K_j^{(2)} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,i<j} O_{ij} (\vec{r} \cdot \vec{\sigma}) O_{ij}^\dagger, \quad (\text{C8}) \end{aligned}$$

$$\begin{aligned} \text{Tr}[S_{01}(\rho)] &= \frac{1}{2} \text{Tr}[S_{01}(\mathbb{1}_2) + S_{01}(\vec{r} \cdot \vec{\sigma})] \\ &= \frac{1}{2} \text{Tr}[S_{01}(\mathbb{1}_2)] + \frac{1}{2} \text{Tr}[S_{01}(\vec{r} \cdot \vec{\sigma})] \\ &= \frac{1}{2} \text{Tr} \left(\sum_{i,j} K_i^{(1)} K_j^{(2)} K_j^{(2)\dagger} K_i^{(1)\dagger} \right) - \frac{1}{2} \text{Tr} \left(\sum_{i,j,i<j} O_{ij} O_{ij}^\dagger \right) \\ &\quad + \frac{1}{2} \text{Tr} \left(\sum_{i,j} K_i^{(1)} K_j^{(2)} (\vec{r} \cdot \vec{\sigma}) K_j^{(2)\dagger} K_i^{(1)\dagger} - \sum_{i,j,i<j} O_{ij} (\vec{r} \cdot \vec{\sigma}) O_{ij}^\dagger \right) \end{aligned}$$

$$\begin{aligned}
&= b - \frac{1}{2} \text{Tr} \left(\sum_{i,j,i<j} O_{ij} O_{ij}^\dagger \right) - \frac{1}{2} \text{Tr} \left(\sum_{i,j,i<j} O_{ij} (\vec{r} \cdot \vec{\sigma}) O_{ij}^\dagger \right) \\
&= b - \text{Tr} \left(\sum_{i,j,i<j} O_{ij}^\dagger O_{ij} \left[\frac{1}{2} (\mathbb{1}_2 + \vec{r} \cdot \vec{\sigma}) \right] \right). \tag{C9}
\end{aligned}$$

We can define a two-dimensional Hermitian operator \mathcal{O} , which can be expanded in the Pauli basis,

$$\mathcal{O} \equiv \sum_{i,j,i<j} O_{ij}^\dagger O_{ij} = c \mathbb{1}_2 + \vec{\delta} \cdot \vec{\sigma}, \tag{C10}$$

$$\begin{aligned}
&\mathcal{O} \left[\frac{1}{2} (\mathbb{1}_2 + \vec{r} \cdot \vec{\sigma}) \right] \\
&= (c \mathbb{1}_2 + \vec{\delta} \cdot \vec{\sigma}) \left[\frac{1}{2} (\mathbb{1}_2 + \vec{r} \cdot \vec{\sigma}) \right] \\
&= \frac{1}{2} [c \mathbb{1}_2 + (\vec{\delta} \cdot \vec{r}) \mathbb{1}_2 + (c \vec{r} + \vec{\delta}) \cdot \vec{\sigma} + i(\vec{\delta} \times \vec{r}) \cdot \vec{\sigma}]. \tag{C11}
\end{aligned}$$

It is straightforward to see that $\vec{\delta} = 0$ when probe-state independence is required. When general Pauli channels are used,

$$\begin{aligned}
\frac{1}{2} c &= -n_y^2 [\cos(\xi) - 1] (p + p_z) (p_x - p_y) \\
&\quad - n_z^2 [\cos(\xi) - 1] (p + p_y) (p_x - p_z) \\
&\quad - \cos(\xi) (p - p_x) (p_y + p_z) \\
&\quad + p_x p_y + p_x p_z + 2p_y p_z + p p_y + p p_z, \tag{C12}
\end{aligned}$$

and $\vec{\delta} = 0$, exhibiting the probe-state independence property of the switch channel.

APPENDIX D: SEVERAL CHANNELS

The general Pauli channel defined in Eq. (9) can be reduced to the depolarization channel, phase-damping channel, and phase-flip channel, depending on the value of p_i .

1. Depolarization channel

The general Pauli channel reduces to the depolarization channel D when $p_i = p$ and $p_x = p_y = p_z = (1 - p)/3$. The depolarization channel is isotropic and can be described by the equation

$$\mathcal{E}_D(\rho) = p\rho + \frac{1-p}{3} (\sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y + \sigma_z \rho \sigma_z), \tag{D1}$$

where ρ is the input density matrix; σ_x , σ_y , and σ_z are the Pauli matrices; and p lies in the interval $[\frac{1}{4}, 1]$ in the depolarization channel. The depolarization channel can be visualized as a uniform Bloch ball shrinking process, resulting in a maximally mixed state when $p = \frac{1}{4}$,

$$\mathbb{1}_2 = \frac{\rho + \sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y + \sigma_z \rho \sigma_z}{2}. \tag{D2}$$

2. Phase-flip and phase-damping channel

When $p_x = p_y = 0$, the generalized Pauli channel reduces to the phase-flip channel, as given by

$$\begin{aligned}
\mathcal{E}_{\text{PF}}(\rho) &= p\rho + (1-p)\sigma_z \rho \sigma_z \\
&= \begin{bmatrix} \rho_{11} & (2p-1)\rho_{12} \\ (2p-1)\rho_{21} & \rho_{22} \end{bmatrix}. \tag{D3}
\end{aligned}$$

The Kraus operators for the phase flip (PF) are $K_0^{\text{PF}} = \sqrt{p} \mathbb{1}_2$ and $K_1^{\text{PF}} = \sqrt{1-p} \sigma_z$, where $\mathbb{1}_2$ is the two-dimensional identity matrix. The probability of obtaining the \pm outcomes in the switch channel composed of the PF, denoted by P_\pm , can be expressed as

$$\begin{aligned}
P_\pm &= \text{Tr}(\rho_{p\pm}) = \frac{1}{2} \pm \sqrt{p(1-p)} \mathcal{Q}(\xi, p, 0, 0, 1-p) \\
&= \frac{1}{2} \pm \sqrt{p(1-p)} \left(1 - 4p(1-p) \sin^2 \frac{\xi}{2} \right). \tag{D4}
\end{aligned}$$

Furthermore, the Kraus operators for the phase-damping channel are

$$K_0^{\text{PD}} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \quad K_1^{\text{PD}} = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\gamma} \end{bmatrix}, \tag{D5}$$

where γ is the damping parameter. The Kraus operators for the PF and phase damping (PD) are related by a unitary transformation given by

$$\begin{bmatrix} K_0^{\text{PD}} \\ K_1^{\text{PD}} \end{bmatrix} = \begin{bmatrix} \frac{1+\sqrt{1-\gamma}}{2\sqrt{p}} & \frac{1-\sqrt{1-\gamma}}{2\sqrt{1-p}} \\ \frac{\sqrt{\gamma}}{2\sqrt{p}} & \frac{-\sqrt{\gamma}}{2\sqrt{1-p}} \end{bmatrix} \begin{bmatrix} K_0^{\text{PF}} \\ K_1^{\text{PF}} \end{bmatrix}. \tag{D6}$$

It can be shown that when $\sqrt{1-\gamma} = 2p - 1$, the PD is equivalent to the PF. The minimum value of p corresponds to the case where the PF completely destroys the coherence of the quantum state, resulting in vanishing off-diagonal elements in the density matrix. The maximum value of p is 1, indicating no noise present in the channel. Thus, the valid range for p is $[\frac{1}{2}, 1]$.

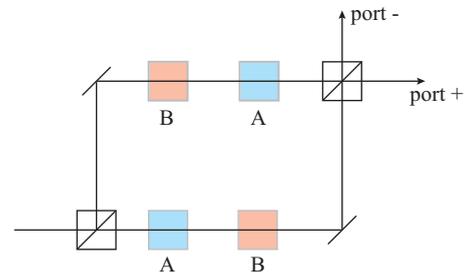


FIG. 7. Schematic diagram illustrating the experimental setup with \hat{A} and \hat{B} represented as $\sigma_i U_\xi$ and $\sigma_j U_\xi$, respectively. Photons emitted from port + and port - are collected by virtual detectors D_+ and D_- , respectively.

APPENDIX E: SINGLE-EXPERIMENT ANALYSIS AND COMBINATION STRATEGY

We will now derive the explicit expression (24) from a combination perspective of the single experiment (see Fig. 7). Denote the states of the single photon emitted from port + and port - by $|\Psi_1\rangle$ and $|\Psi_2\rangle$, respectively. When the operators \hat{A} and \hat{B} are exchanged with each other, $|\Psi_2\rangle$ becomes $-|\Psi_2\rangle$ while leaving $|\Psi_1\rangle$ unchanged.

Denote the ratios of the first and second beam splitters by \mathcal{D}_1 and \mathcal{D}_2 , respectively. The action of the beam splitter on the

input state $|\psi_{in}\rangle$ is given by

$$|\psi_{in}\rangle \xrightarrow{\text{beam splitter}} \sqrt{\mathcal{D}_i}|\psi_t\rangle + i\sqrt{1-\mathcal{D}_i}|\psi_r\rangle. \quad (\text{E1})$$

The expressions for $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are then given by

$$|\Psi_1\rangle = i[\sqrt{\mathcal{D}_1(1-\mathcal{D}_2)}\hat{B}\hat{A} + \sqrt{\mathcal{D}_2(1-\mathcal{D}_1)}\hat{A}\hat{B}]|\psi_{in}\rangle,$$

$$|\Psi_2\rangle = [\sqrt{\mathcal{D}_1\mathcal{D}_2}\hat{B}\hat{A} - \sqrt{(1-\mathcal{D}_1)(1-\mathcal{D}_2)}\hat{A}\hat{B}]|\psi_{in}\rangle. \quad (\text{E2})$$

Using these expressions, we can calculate the inner products as follows:

$$\langle\Psi_1|\Psi_1\rangle = \mathcal{D}_1(1-\mathcal{D}_2) + \mathcal{D}_2(1-\mathcal{D}_1) + \sqrt{\mathcal{D}_1\mathcal{D}_2(1-\mathcal{D}_1)(1-\mathcal{D}_2)}\langle\psi_{in}|(\hat{A}^\dagger\hat{B}^\dagger\hat{A}\hat{B} + \hat{B}^\dagger\hat{A}^\dagger\hat{B}\hat{A})|\psi_{in}\rangle,$$

$$\langle\Psi_2|\Psi_2\rangle = \mathcal{D}_1\mathcal{D}_2 + (1-\mathcal{D}_1)(1-\mathcal{D}_2) - \sqrt{\mathcal{D}_1\mathcal{D}_2(1-\mathcal{D}_1)(1-\mathcal{D}_2)}\langle\psi_{in}|(\hat{A}^\dagger\hat{B}^\dagger\hat{A}\hat{B} + \hat{B}^\dagger\hat{A}^\dagger\hat{B}\hat{A})|\psi_{in}\rangle. \quad (\text{E3})$$

We conducted an experiment with three categories of \hat{A} and \hat{B} combinations and observed that for all situations, the matrix $\hat{O} = \hat{A}^\dagger\hat{B}^\dagger\hat{A}\hat{B} + \hat{B}^\dagger\hat{A}^\dagger\hat{B}\hat{A}$ is diagonal with identical elements.

(1) When $\hat{A} = \hat{B} = \sigma_i U_\xi$, where $\sigma_i \in \mathbb{1}_2, \sigma_x, \sigma_y, \sigma_z$, one can obtain $\hat{O} = 2\mathbb{1}_2$.

(2) When $\hat{A} = U_\xi$ and $\hat{B} = \sigma_i U_\xi$, where $\sigma_i \in \sigma_x, \sigma_y, \sigma_z$,

$$\hat{O} = [2n_i^2(1 - \cos \xi) + 2 \cos \xi] \mathbb{1}_2. \quad (\text{E4})$$

(3) When $\hat{A} = \sigma_i U_\xi$ and $\hat{B} = \sigma_j U_\xi$, where $\sigma_i, \sigma_j \in \sigma_x, \sigma_y, \sigma_z$ and $i \neq j$,

$$\hat{O} = [2n_k^2(\cos \xi - 1) - 2 \cos \xi] \mathbb{1}_2. \quad (\text{E5})$$

If we choose a general Pauli channel (9), we can simulate the probability distribution of a multipath interferometer with that of a two-path interferometer, due to the incoherent property of quantum channels. The probability distributions P_+ and P_- are given by Eq. (E6), where

$$P_+ = \mathcal{D}_1(1-\mathcal{D}_2) + \mathcal{D}_2(1-\mathcal{D}_1) + 2\sqrt{\mathcal{D}_1\mathcal{D}_2(1-\mathcal{D}_1)(1-\mathcal{D}_2)}\mathcal{Q}(n_x, n_y, n_z, p_l, p_x, p_y, p_z), \quad (\text{E6})$$

$$P_- = \mathcal{D}_1\mathcal{D}_2 + (1-\mathcal{D}_1)(1-\mathcal{D}_2) - 2\sqrt{\mathcal{D}_1\mathcal{D}_2(1-\mathcal{D}_1)(1-\mathcal{D}_2)}\mathcal{Q}(n_x, n_y, n_z, p_l, p_x, p_y, p_z),$$

$$\begin{aligned} \mathcal{Q}(n_x, n_y, n_z, p_l, p_x, p_y, p_z) = & 2[n_y^2(p_l + p_z)(p_x - p_y) + n_z^2(p_l + p_y)(p_x - p_z) + (p_l - p_x)(p_y + p_z)] \cos \xi \\ & - 2(n_x^2 p_y p_z + n_y^2 p_x p_z + n_z^2 p_x p_y) + 2p_l(n_x^2 p_x + n_y^2 p_y + n_z^2 p_z) + p_l^2 + p_x^2 + p_y^2 + p_z^2. \end{aligned} \quad (\text{E7})$$

If \mathcal{D}_2 is exactly 0.5 and we replace \mathcal{D}_1 with p in Eq. (E6), the resulting expression will be identical to Eq. (24). This observation confirms the validity and soundness of our combination strategy.

APPENDIX F: UNITARY TRANSFORMATION IMPLEMENTATION

The expression for an SU(2) operator, denoted by U_2 ,

$$U_2 = \exp\left(-i\frac{\xi}{2}\vec{n} \cdot \vec{\sigma}\right), \quad (\text{F1})$$

can be written as using the Z-Y decomposition as

$$\begin{aligned} U_2 &= \hat{R}_y(\alpha)\hat{R}_z(\beta)\hat{R}_y(\gamma) \\ &= \exp\left(\frac{-i\alpha\sigma_y}{2}\right)\exp\left(\frac{-i\beta\sigma_z}{2}\right)\exp\left(\frac{-i\gamma\sigma_y}{2}\right) \\ &= \exp\left(\frac{i\pi}{2}\right)\hat{Q}(\tau_3)\hat{H}(\tau_2)\hat{Q}(\tau_1), \end{aligned} \quad (\text{F2})$$

where $\hat{R}_i(j)$ represents a rotation around the i axis by an angle j , and α, β , and γ are the Euler angles. The parameters τ_1, τ_2 ,

and τ_3 are given by (see Fig. 8)

$$\begin{aligned} \tau_1 &= -\frac{\gamma}{2} + \frac{\pi}{4}, \\ \tau_2 &= \frac{\alpha + \beta - \gamma}{4} + \frac{\pi}{4}, \\ \tau_3 &= \frac{\alpha}{2} + \frac{\pi}{4}. \end{aligned} \quad (\text{F3})$$

By equating Eqs. (F1) and (F2), we obtain the following set of equations that establish the relationship between the SU(2)

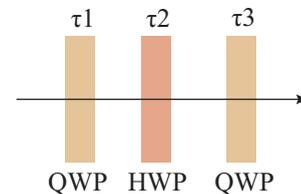


FIG. 8. Arbitrary SU(2) transformation implementation according to $\hat{Q}(\tau_3)\hat{H}(\tau_2)\hat{Q}(\tau_1)$ cascaded waveplates. Here τ_1, τ_2 , and τ_3 represent the angles between the fast axis of the waveplate and the horizontal polarization direction.

TABLE I. Implementation of Pauli operators according to $\hat{Q}(\tau_1)$, $\hat{H}(\tau_2)$, and $\hat{Q}(\tau_3)$.

SU(2) operator	τ_1	τ_2	τ_3
$\mathbb{1}$	0	0	0
$i\sigma_x$	0	45°	0
$i\sigma_y$	90°	45°	0
$i\sigma_z$	90°	0	0

parameters and the waveplate set $\hat{Q}(\tau_3)\hat{H}(\tau_2)\hat{Q}(\tau_1)$:

$$\begin{aligned}\cos \frac{\xi}{2} &= \cos(2\tau_2 - \tau_1 - \tau_3) \cos(\tau_1 - \tau_3), \\ n_x \sin \frac{\xi}{2} &= -\sin(2\tau_2 - \tau_1 - \tau_3) \cos(\tau_1 + \tau_3), \\ n_y \sin \frac{\xi}{2} &= -\cos(2\tau_2 - \tau_1 - \tau_3) \sin(\tau_1 - \tau_3), \\ n_z \sin \frac{\xi}{2} &= \sin(2\tau_2 - \tau_1 - \tau_3) \sin(\tau_1 + \tau_3).\end{aligned}\quad (\text{F4})$$

The angles of the waveplate set to implement Pauli operators and SU(2) transformation $\exp(-i\xi\hat{n}\cdot\hat{\sigma}/2)$ can be solved using Eqs. (F4). We list the solutions in Tables I and II.

APPENDIX G: EXPERIMENTAL DETAILS

The piezoelectric linear stage (Newport AG-LS25-27) with a minimum incremental motion of 0.2 μm is used to search for the optimal interference point. In Fig. 9(a) the maximal interference visibility is 95.00% when the piezoelectric linear stage is at 6.9066 mm. The phase of the interferometer is controlled by a piezoelectric transducer (PZT) (Thorlabs PA4FKW). We acquire the interference fringe by scanning the voltage exerted on the PZT with a step of 5 V as shown in Fig. 9(b). We characterize the simulated general Pauli channels by quantum process tomography. The results of the bit-flip channel x , bit-phase-flip channel y , phase-flip channel z , and depolarization channel d are shown in Fig. 10(a). The average quantum channel fidelity exceeds 95%. The visibility of the Mach-Zehnder interferometer for six polarization states is shown in Fig. 10(b). When the input state is $|H\rangle$, we obtain a lowest visibility around 91.32%.

TABLE II. Implementation of the SU(2) transformation $\exp(-i\xi\hat{n}\cdot\hat{\sigma}/2)$ using the waveplate operators $\hat{Q}(\tau_1)$, $\hat{H}(\tau_2)$, and $\hat{Q}(\tau_3)$, where $\vec{n} = (1, 0, 0)$.

ξ	τ_1	τ_2	τ_3
0°	90°	-90°	90°
22.5°	90°	-84.375°	90°
45°	90°	-78.75°	90°
67.5°	90°	-73.125°	90°
90°	90°	-67.5°	90°
112.5°	90°	-61.875°	90°
135°	90°	-56.25°	90°
157.5°	90°	-50.625°	90°
180°	90°	-45°	90°

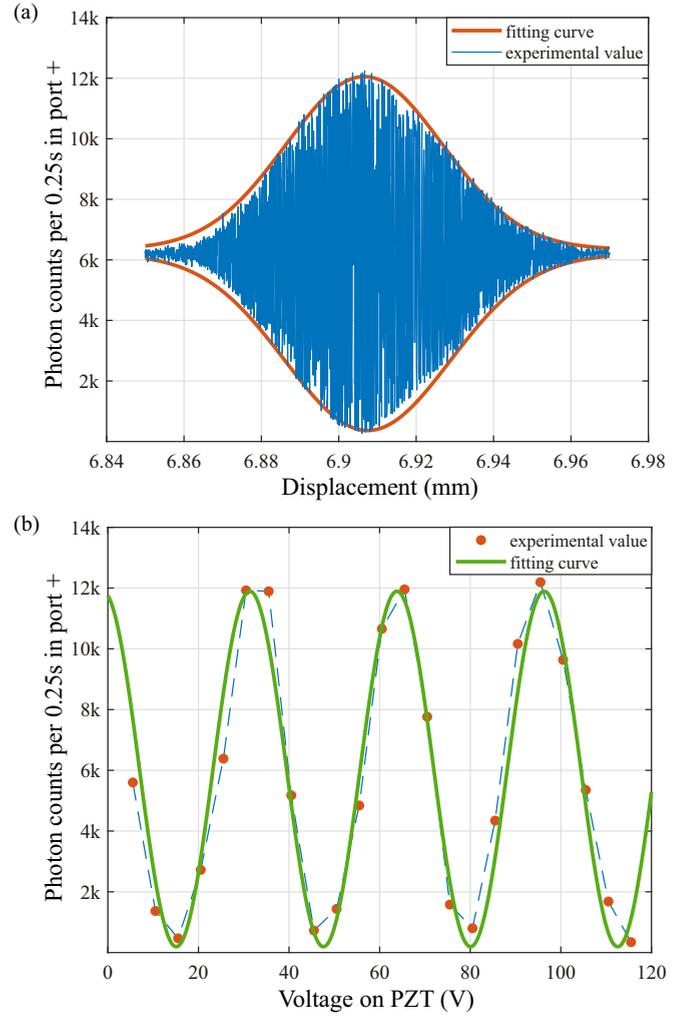


FIG. 9. Degree of first-order coherence of the heralded single-photon source and the interference fringe of the Mach-Zehnder interferometer. (a) Signal photons measured in port + with a minimum incremental motion of 0.2 μm of the piezoelectric linear stage. (b) Interference fringe at the optimal interference point of 6.9066 mm.

1. Quantum process tomography

A quantum channel can be expressed as

$$\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger, \quad (\text{G1})$$

where $\{E_i\}$ is the Kraus operator and satisfies $\sum_i E_i^\dagger E_i = \mathbb{1}_2$. In the experiment, we aim to characterize an unknown quantum channel using a special group of operators \tilde{E}_m to express the quantum channel as follows:

$$\mathcal{E}(\rho) = \sum_i \tilde{E}_m \rho \tilde{E}_n^\dagger \chi_{mn}. \quad (\text{G2})$$

Quantum process tomography is used to determine the χ matrix by measuring the output states of the quantum channel. For a qubit system, the operators \tilde{E}_m are chosen as

$$\begin{aligned}\tilde{E}_0 &= \mathbb{1}_2, & \tilde{E}_1 &= \sigma_x, \\ \tilde{E}_2 &= i\sigma_y, & \tilde{E}_3 &= \sigma_z.\end{aligned}\quad (\text{G3})$$

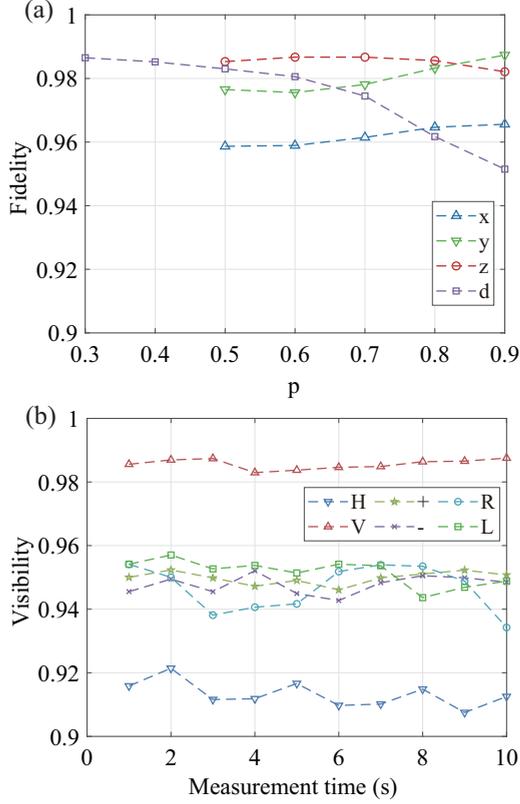


FIG. 10. Channel fidelity for different kinds of quantum channel and visibility of the interferometer for different probe states. (a) Channel fidelity for four simulated quantum channels at different noise levels: x , bit-flip channel; y , bit-phase-flip channel; z , phase-flip channel; and d , depolarization channel. (b) Visibility of the interferometer for six polarization states in 10 s. Here $|+\rangle = 1/\sqrt{2}(|H\rangle + |V\rangle)$, $|-\rangle = 1/\sqrt{2}(|H\rangle - |V\rangle)$, $|R\rangle = 1/\sqrt{2}(|H\rangle + i|V\rangle)$, and $|L\rangle = 1/\sqrt{2}(|H\rangle - i|V\rangle)$.

Then we need to determine $\rho'_1 = \mathcal{E}(|0\rangle\langle 0|)$, $\rho'_4 = \mathcal{E}(|1\rangle\langle 1|)$, $\mathcal{E}(|+\rangle\langle +|)$, and $\mathcal{E}(|R\rangle\langle R|)$ according to quantum state tomography. The χ matrix can be determined as follows:

$$\chi = \frac{1}{2} \begin{pmatrix} \mathbb{1}_2 & \sigma_x \\ \sigma_x & -\mathbb{1}_2 \end{pmatrix} \begin{pmatrix} \rho'_1 & \rho'_2 \\ \rho'_3 & \rho'_4 \end{pmatrix} \begin{pmatrix} \mathbb{1}_2 & \sigma_x \\ \sigma_x & -\mathbb{1}_2 \end{pmatrix}, \quad (\text{G4})$$

$$\begin{aligned} \rho'_2 &= \mathcal{E}(|+\rangle\langle +|) - i\mathcal{E}(|R\rangle\langle R|) \\ &\quad - (1 - i)(\mathcal{E}(|0\rangle\langle 0|) + \mathcal{E}(|1\rangle\langle 1|))/2, \\ \rho'_3 &= \mathcal{E}(|+\rangle\langle +|) + i\mathcal{E}(|R\rangle\langle R|) \\ &\quad - (1 + i)(\mathcal{E}(|0\rangle\langle 0|) + \mathcal{E}(|1\rangle\langle 1|))/2. \end{aligned} \quad (\text{G5})$$

The real and imaginary parts of the χ matrix for the phase-flip channel at different values of p are illustrated in Figs. 11 and 12, respectively.

2. Estimator

We observe that the probability of projecting onto $|\pm\rangle_c = |0\rangle_c + |1\rangle_c$ has the form

$$P_{\pm} = f_{\pm}(p) \cos \xi + g_{\pm}(p), \quad (\text{G6})$$

where f and g are functions that depend on p . Taking into account the effects of detector efficiency, nonunit visibility, and imperfect beam splitters, it is expected that f and g may deviate from the true values. To eliminate systematic errors in our experimental setup, we choose $\vec{n} = (1, 0, 0)$ and perform the experiment with different ξ values. By fitting the curve of P_{\pm} with respect to p , we can obtain the calibrated functions f_{\pm}^{cal} and g_{\pm}^{cal} , which are given by

$$P_{\pm}^{\text{cal}} = f_{\pm}^{\text{cal}}(p) \cos(\xi - \xi_e) + g_{\pm}^{\text{cal}}(p), \quad (\text{G7})$$

where ξ_e represents the system error. Denote the detection efficiencies by η_+ and η_- for detectors D_+ and D_- , respectively.

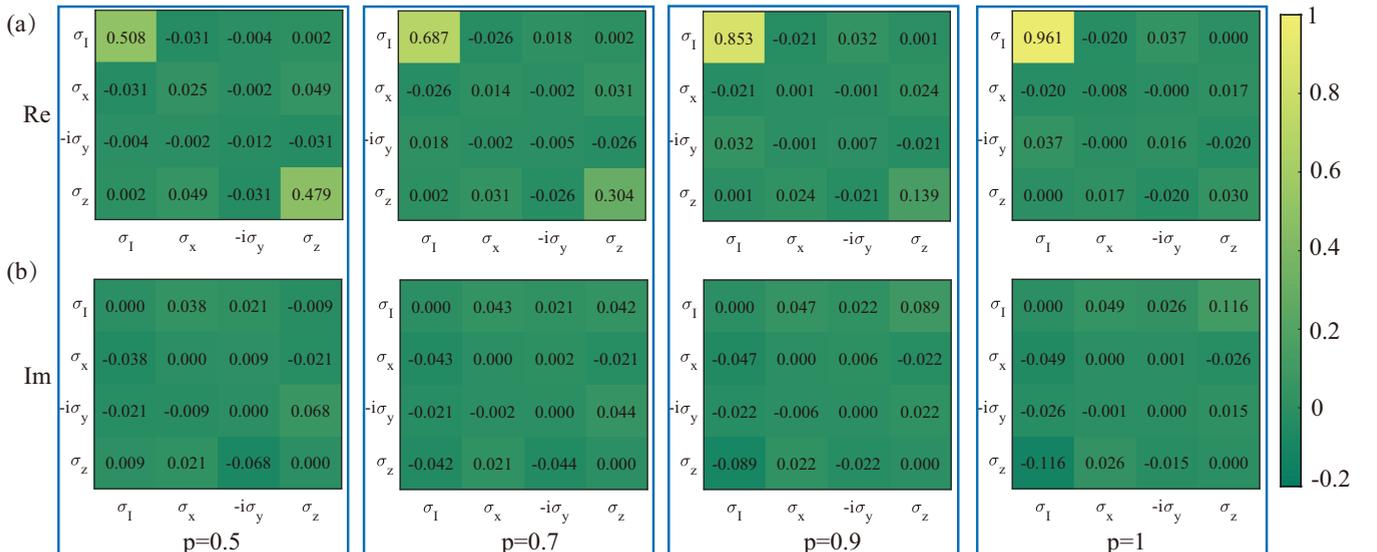


FIG. 11. Quantum process tomography result for the quantum-flip channel under different noise levels p : (a) real part of the χ matrix and (b) imaginary part of the χ matrix.

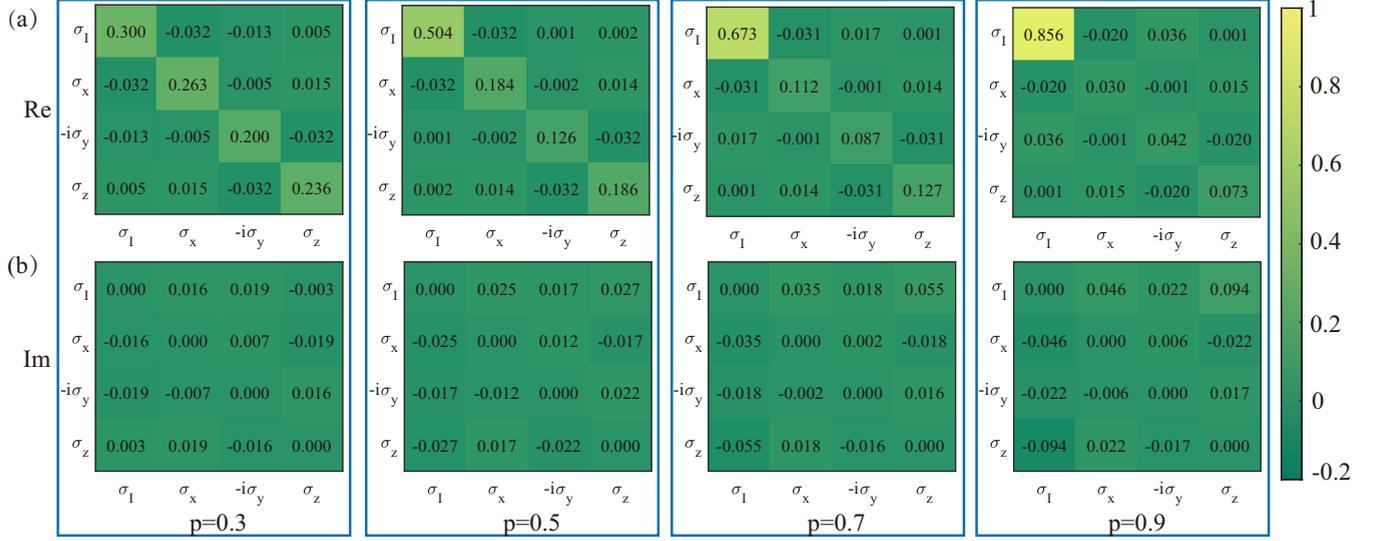


FIG. 12. Quantum process tomography result for the quantum depolarization channel under different noise levels p : (a) imaginary part of the χ matrix and (b) imaginary part of the χ matrix.

We can obtain the renormalized probabilities as

$$P_+^{\text{re}} = \frac{\eta_+ P_+^{\text{cal}}}{\eta_+ P_+^{\text{cal}} + \eta_- P_-^{\text{cal}}}, \quad P_-^{\text{re}} = \frac{\eta_- P_-^{\text{cal}}}{\eta_+ P_+^{\text{cal}} + \eta_- P_-^{\text{cal}}}. \quad (\text{G8})$$

We can determine the functions $f_{\pm}^{\text{cal}}(p)$, $g_{\pm}^{\text{cal}}(p)$, and ξ_e by fitting the experimental data. For instance, for the phase-flip switched channel with $\vec{n} = (1, 0, 0)$, the probabilities P_{\pm} are given by

$$P_{\pm} = \pm 2\sqrt{p(1-p)}p(1-p)\cos\xi \pm \sqrt{p(1-p)}(2p^2 - 2p + 1) + \frac{1}{2}, \quad (\text{G9})$$

where

$$f_{\pm} = \pm 2\sqrt{p(1-p)}p(1-p), \quad g_{\pm} = \pm \sqrt{p(1-p)}(2p^2 - 2p + 1) + \frac{1}{2}. \quad (\text{G10})$$

To calibrate our setup, we introduce the parameters α and β ,

$$f_{\pm}^{\text{cal}} = \alpha f_{\pm}, \quad g_{\pm}^{\text{cal}} = g_{\pm} \pm \beta. \quad (\text{G11})$$

Using maximum-likelihood estimation (MLE), we can obtain

$$\mathcal{L}(\xi) = (\eta_+ P_+^{\text{cal}})^{N_+^c} (\eta_- P_-^{\text{cal}})^{N_-^c} (1 - \eta_+ P_+^{\text{cal}} - \eta_- P_-^{\text{cal}})^{N_h - N_+^c - N_-^c},$$

$$\frac{\partial \ln \mathcal{L}(\xi)}{\partial \xi} = N_+^c \frac{\partial_{\xi} \eta_+ P_+^{\text{cal}}}{\eta_+ P_+^{\text{cal}}} + N_-^c \frac{\partial_{\xi} \eta_- P_-^{\text{cal}}}{\eta_- P_-^{\text{cal}}} + (N_h - N_+^c - N_-^c) \frac{\partial_{\xi} (1 - \eta_+ P_+^{\text{cal}} - \eta_- P_-^{\text{cal}})}{1 - \eta_+ P_+^{\text{cal}} - \eta_- P_-^{\text{cal}}} = 0. \quad (\text{G14})$$

Taking into account the fact that $P_+^{\text{cal}} + P_-^{\text{cal}} = 0$, we can simplify Eq. (G14) as

$$\frac{N_+^c}{P_+^{\text{cal}}} - \frac{N_-^c}{P_-^{\text{cal}}} + (N_h - N_+^c - N_-^c) \frac{\eta_- - \eta_+}{1 - \eta_+ P_+^{\text{cal}} - \eta_- P_-^{\text{cal}}} = 0. \quad (\text{G15})$$

So the probabilities P_{\pm}^{cal} for the phase-flip switched channel become

$$P_{\pm}^{\text{cal}} = \alpha f_{\pm} \cos(\xi + \xi_e) + g_{\pm} \pm \beta. \quad (\text{G12})$$

We can use the least-squares method to fit the experimental data and obtain the values of α , β , and ξ_e as

$$\arg \min_{\alpha, \beta, \xi_e} \mathcal{L} = \sum_{p, \xi} (P_{\pm}^{\text{expt}} - P_{\pm}^{\text{re}})^2, \quad (\text{G13})$$

where P_{\pm}^{expt} is the experimental probability calculated from photon counts directly. For the phase-flip switched channel, we have $\alpha = 0.9427$, $\beta = -0.02407$, and $\xi_e = -0.02155$. For the depolarization switched channel, we have $\alpha = 0.9436$, $\beta = -0.02111$, and $\xi_e = -0.01937$. For the arbitrary Pauli channel \mathcal{N}_1 , we have $\alpha = 0.9758$, $\beta = -0.01452$, and $\xi_e = -0.02398$.

We consider the scenario where we send N_h heralded photon pairs into the interferometer. Out of these, some are detected by D_+ , some are detected by D_- , and some are not detected by any detector. We can denote the number of heralded photon pairs detected by D_+ by N_+^c , the number of heralded photon pairs detected by D_- by N_-^c , and the number of loss photons due to transmit loss, coupling efficiency, and detector efficiency by $N_h - N_+^c - N_-^c$.

In order to obtain the MLE estimator, we need to determine the values of η_+ and η_- first. Based on the given equations,

$$\begin{aligned} N_i &= \eta_i P_i N_h, \\ N_i^c &= \eta_h \eta_i P_i N_h, \end{aligned} \quad (\text{G16})$$

where N_i represents the total counts of channel i ($i = +, -, h$), P_i is the theoretical probability of detecting a single photon in channel i , and η_h is the efficiency of the heralded channel. Here we assume that the events where D_+ or D_- detects a photon while D_h does not detect a photon are omitted. We can immediately deduce

$$\frac{\eta_i P_i}{P_h} = \frac{N_i^c}{N_h}. \quad (\text{G17})$$

The calibrated Fisher information is

$$\begin{aligned} \mathcal{F}(\xi) &= \frac{(\eta_+ \partial_\xi P_+^{\text{cal}})^2}{\eta_+ P_+^{\text{cal}}} + \frac{(\eta_- \partial_\xi P_-^{\text{cal}})^2}{\eta_- P_-^{\text{cal}}} + \frac{(\eta_+ \partial_\xi P_+^{\text{cal}} + \eta_- \partial_\xi P_-^{\text{cal}})^2}{1 - \eta_+ P_+^{\text{cal}} - \eta_- P_-^{\text{cal}}} \\ &= (\partial_\xi P_+^{\text{cal}})^2 \left(\frac{\eta_+}{P_+^{\text{cal}}} + \frac{\eta_-}{P_-^{\text{cal}}} + \frac{(\eta_+ - \eta_-)^2}{1 - \eta_+ P_+^{\text{cal}} - \eta_- P_-^{\text{cal}}} \right) \\ &= \frac{(\partial_\xi P_+^{\text{cal}})^2}{\eta_+} \left(\frac{1}{P_+^{\text{cal}}} + \frac{\eta_- / \eta_+}{P_-^{\text{cal}}} + \frac{(1 - \eta_- / \eta_+)^2}{1 / \eta_+ - P_+^{\text{cal}} - \eta_- / \eta_+ P_-^{\text{cal}}} \right). \end{aligned}$$

The Cramér-Rao lower bound on the variance of the estimator $\hat{\xi}$ is given by

$$\text{var}(\hat{\xi}) \geq \frac{1}{N_h \mathcal{F}(\xi)}. \quad (\text{G19})$$

Equality holds when using MLE for large N_h asymptotically.

In our experiment, the total resource usage is determined by the coincidence counts $N_+^c + N_-^c$, which are considered as the consumed resource due to the postselection approach. Since only the relative efficiency of detectors D_+ and D_- is important, we can set $\eta_+^r = 1$ and consequently $\eta_-^r = \eta_- / \eta_+ = 94.21\%$. Taking into account the detector efficiency, the effective number of photons used in the experiment is given by $N_+^c + N_-^c / \eta_-^r$. The Fisher information extracted from the experiment is expressed as

$$\mathcal{F}(\xi) = \frac{\eta_-^r}{(\eta_-^r N_+^c + N_-^c) \text{var}(\hat{\xi})}, \quad (\text{G20})$$

where $\text{var}(\hat{\xi})$ is the variance of the estimator $\hat{\xi}$.

By adopting the postselection approach, where $P_h = 1$ and $\eta_h = 1$ always hold, we can simplify the equation as

$$\eta_i = \frac{N_i^c}{N_h P_i}. \quad (\text{G18})$$

The detector efficiencies η_+ and η_- can be measured in a preexperiment where one path of the interferometer is blocked. By removing all waveplates in the interferometer, we can obtain the coincidence counts of ports $+$ and $-$, respectively. Using Eq. (G18) with $P_\pm = \frac{1}{2}$, we can calculate $\eta_+ = 7.95\% \pm 0.18\%$ and $\eta_- = 7.49\% \pm 0.17\%$. The heralding efficiency of our setup can be calculated as $(N_+^c + N_-^c) / \sqrt{N_h(N_+ + N_-)} = 14.67\% \pm 0.13\%$.

3. Multiple photon pairs in spontaneous parametric down-conversion

We utilized a laser diode centered at 405 nm to pump a 3-mm PPKTP crystal, generating correlated photon pairs through type-II spontaneous parametric down-conversion. However, in this process, there may also be undesired emissions of multiple photon pairs, which need to be taken into account for accurate resource calculation in the context of quantum metrology. In particular, we focus on the probability of emitting more than one pair of photons, denoted by $N^{(m)}$. This probability can be expressed as

$$\begin{aligned} P^{(m)} &= \frac{N^{(m)}}{N} = \frac{N_{ijh}^c}{\eta_i \eta_j P_i P_j N_h} \\ &\approx 0.08373\% \pm 0.01510\%. \end{aligned} \quad (\text{G21})$$

When the total coincidence count is set to 20 000, the number of events with multiple photons is estimated to be around 17, which is considered negligible in the experiment.

- [1] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology, *Nat. Phys.* **7**, 406 (2011).
- [2] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, The elusive Heisenberg limit in quantum-enhanced metrology, *Nat. Commun.* **3**, 1063 (2012).
- [3] R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, P. Hyllus, L. Pezzé, and A. Smerzi, Useful multiparticle entanglement and sub-shot-noise sensitivity in experimental phase estimation, *Phys. Rev. Lett.* **107**, 080504 (2011).

- [4] R. Demkowicz-Dobrzański and L. Maccone, Using entanglement against noise in quantum metrology, *Phys. Rev. Lett.* **113**, 250801 (2014).
- [5] K. Wang, X. Wang, X. Zhan, Z. Bian, J. Li, B. C. Sanders, and P. Xue, Entanglement-enhanced quantum metrology in a noisy environment, *Phys. Rev. A* **97**, 042112 (2018).
- [6] Z. Huang, C. Macchiavello, and L. Maccone, Usefulness of entanglement-assisted quantum metrology, *Phys. Rev. A* **94**, 012101 (2016).

- [7] M. Kacprowicz, R. Demkowicz-Dobrzański, W. Wasilewski, K. Banaszek, and I. A. Walmsley, Experimental quantum-enhanced estimation of a lossy phase shift, *Nat. Photon.* **4**, 357 (2010).
- [8] H. Yuan, Sequential feedback scheme outperforms the parallel scheme for Hamiltonian parameter estimation, *Phys. Rev. Lett.* **117**, 160801 (2016).
- [9] Z. Hou, R.-J. Wang, J.-F. Tang, H. Yuan, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, Control-enhanced sequential scheme for general quantum parameter estimation at the Heisenberg limit, *Phys. Rev. Lett.* **123**, 040501 (2019).
- [10] S. Pang and A. N. Jordan, Optimal adaptive control for quantum metrology with time-dependent Hamiltonians, *Nat. Commun.* **8**, 14695 (2017).
- [11] R. Chaves, J. B. Brask, M. Markiewicz, J. Kołodyński, and A. Acín, Noisy metrology beyond the standard quantum limit, *Phys. Rev. Lett.* **111**, 120401 (2013).
- [12] Y. Yang, S. Ru, M. An, Y. Wang, F. Wang, P. Zhang, and F. Li, Multiparameter simultaneous optimal estimation with an SU(2) coding unitary evolution, *Phys. Rev. A* **105**, 022406 (2022).
- [13] J. B. Brask, R. Chaves, and J. Kołodyński, Improved quantum magnetometry beyond the standard quantum limit, *Phys. Rev. X* **5**, 031010 (2015).
- [14] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, *Nat. Commun.* **3**, 1092 (2012).
- [15] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, Experimental superposition of orders of quantum gates, *Nat. Commun.* **6**, 7913 (2015).
- [16] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, Experimental verification of an indefinite causal order, *Sci. Adv.* **3**, e1602589 (2017).
- [17] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, Indefinite causal order in a quantum switch, *Phys. Rev. Lett.* **121**, 090503 (2018).
- [18] D. Ebler, S. Salek, and G. Chiribella, Enhanced communication with the assistance of indefinite causal order, *Phys. Rev. Lett.* **120**, 120502 (2018).
- [19] Y. Guo, X.-M. Hu, Z.-B. Hou, H. Cao, J.-M. Cui, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, and G. Chiribella, Experimental transmission of quantum information using a superposition of causal orders, *Phys. Rev. Lett.* **124**, 030502 (2020).
- [20] P. A. Guérin, G. Rubino, and Č. Brukner, Communication through quantum-controlled noise, *Phys. Rev. A* **99**, 062317 (2019).
- [21] G. Rubino, L. A. Rozema, D. Ebler, H. Kristjánsson, S. Salek, P. A. Guérin, A. A. Abbott, C. Branciard, Č. Brukner, G. Chiribella, and P. Walther, Experimental quantum communication enhancement by superposing trajectories, *Phys. Rev. Res.* **3**, 013093 (2021).
- [22] D. Felce and V. Vedral, Quantum refrigeration with indefinite causal order, *Phys. Rev. Lett.* **125**, 070603 (2020).
- [23] H. Cao, N.-N. Wang, Z. Jia, C. Zhang, Y. Guo, B.-H. Liu, Y.-F. Huang, C.-F. Li, and G.-C. Guo, Quantum simulation of indefinite causal order induced quantum refrigeration, *Phys. Rev. Res.* **4**, L032029 (2022).
- [24] X. Nie, X. Zhu, K. Huang, K. Tang, X. Long, Z. Lin, Y. Tian, C. Qiu, C. Xi, X. Yang, J. Li, Y. Dong, T. Xin, and D. Lu, Experimental realization of a quantum refrigerator driven by indefinite causal orders, *Phys. Rev. Lett.* **129**, 100603 (2022).
- [25] X. Zhao, Y. Yang, and G. Chiribella, Quantum metrology with indefinite causal order, *Phys. Rev. Lett.* **124**, 190503 (2020).
- [26] F. Chapeau-Blondeau, Indefinite causal order for quantum metrology with quantum thermal noise, *Phys. Lett. A* **447**, 128300 (2022).
- [27] M. Ban, Quantum Fisher information of phase estimation in the presence of indefinite causal order, *Phys. Lett. A* **468**, 128749 (2023).
- [28] F. Chapeau-Blondeau, Indefinite causal order for quantum phase estimation with Pauli noise, *Fluct. Noise Lett.* **22**, 2350036 (2023).
- [29] Q. Liu, Z. Hu, H. Yuan, and Y. Yang, Optimal strategies of quantum metrology with a strict hierarchy, *Phys. Rev. Lett.* **130**, 070803 (2023).
- [30] S. Kurdziałek, W. Górecki, F. Albarelli, and R. Demkowicz-Dobrzański, Using adaptiveness and causal superpositions against noise in quantum metrology, *Phys. Rev. Lett.* **131**, 090801 (2023).
- [31] P. Yin, X. Zhao, Y. Yang, Y. Guo, W.-H. Zhang, G.-C. Li, Y.-J. Han, B.-H. Liu, J.-S. Xu, G. Chiribella, G. Chen, C.-F. Li, and G.-C. Guo, Experimental super-Heisenberg quantum metrology with indefinite gate order, *Nat. Phys.* **19**, 1122 (2023).
- [32] F. Chapeau-Blondeau, Noisy quantum metrology with the assistance of indefinite causal order, *Phys. Rev. A* **103**, 032615 (2021).
- [33] F. Chapeau-Blondeau, Optimizing qubit phase estimation, *Phys. Rev. A* **94**, 022334 (2016).
- [34] F. Chapeau-Blondeau, Optimized probing states for qubit phase estimation with general quantum noise, *Phys. Rev. A* **91**, 052310 (2015).
- [35] R. Simon and N. Mukunda, Universal SU(2) gadget for polarization optics, *Phys. Lett. A* **138**, 474 (1989).
- [36] V. Bagini, R. Borghi, F. Gori, M. Santarsiero, F. Frezza, G. Schettini, and G. S. Spagnolo, The Simon-Mukunda polarization gadget, *Eur. J. Phys.* **17**, 279 (1996).
- [37] A. Z. Goldberg, K. Heshami, and L. L. Sánchez-Soto, Evading noise in multiparameter quantum metrology with indefinite causal order, *Phys. Rev. Res.* **5**, 033198 (2023).