Indefinite Time Directed Quantum Metrology

Gaurang Agrawal[1,2], Pritam Halder[2], and Aditi Sen(De)[2]

[1]Indian Institute of Science Education and Research, Homi Bhabha Rd, Pashan, Pune 411 008, India
[2]Harish-Chandra Research Institute, A CI of Homi Bhabha National Institute, Chhatnag Road, Jhunsi, Allahabad - 211019, India

We explore the performance of the metrology scheme by employing a quantum time flip during encoding, a specific case of processes with indefinite time direction, which we refer to as indefinite time directed metrology (ITDM). In the case of single parameter estimation of a unitary, we demonstrate that our protocol can achieve Heisenberg scaling $(1/N)$ with product probe states, surpassing the standard quantum limit $(1/\sqrt{N})$, where N is the number of particles in the probe. We establish this by computing the quantum Fisher information (QFI) which is a lower bound on the root mean square error occurred during parameter estimation. Although we analytically prove the optimality of the symmetric product probe state in ITDM, entangled probe states produce a higher QFI than optimal product probes without enhancing scaling, highlighting the non-essentiality of entanglement. For phase estimation, we propose a single-qubit measurement on the control qubit that accomplishes near-optimal Fisher information and eventually reaches Heisenberg scaling. Our findings reveal the best orientation of product probe states in every pertinent situation, emphasizing its independence from the parameter to be estimated in the limiting case. Furthermore, we illustrate the benefits of ITDM in noisy metrology, outperforming existing techniques in some situations.

1 Introduction

The goal of quantum metrology [1–7] is to make extremely accurate estimates of a system's properties which is crucial in diverse scenarios such as magnetometry [8, 9] quantum imaging [10–12], atomic clock [13–16], biological sensing [17, 18], dark matter search [19, 20], gravitational wave detection [21–23]. In particular, the metrological procedure involves measuring a quantum state that encodes the parameter to be estimated, and its accuracy is typically quantified by the root mean square error [3]. Under plausible assumptions, the quantity is constrained below by Fisher information (FI), also known as the Cramer-Rao bound [24], while optimizing overall measurement procedures provides quantum Fisher information (QFI), which specifies the ultimate precision limit [25]. It was demonstrated that, while uncorrelated probes yield QFI that scales only linearly with the number of subsystems, achieving the standard quantum limit (SQL), nonclassical resources such as entanglement or spin-squeezing can improve sensitivity by exceeding linear growth to quadratic or beyond, referred to as the Heisenberg limit (HL) or super-Heisenberg precision [3, 4, 26–29].

In conventional metrological protocols, the encoding operation is performed within a framework that adheres to a well-defined causal order and a fixed temporal direction. However, recent studies [30–38] have explored metrological enhancements of HL and beyond with the aid of indefinite causal order (ICO) [39–45] in which the quantum switch [46] serves as an example of a quantum-controlled superposition of several causally ordered transformations [47]. Importantly, such protocols have already been implemented in laboratories [36, 37]. Further, beyond enhancing sensitivity, ICO has exhibited benefits across various domains including state and channel discrimination [48–51], communication capacity and complexity [52–62], along with thermodynamical tasks including charging, refrigeration, work extraction and thermal state activation [63–69].

Similar to the concept of ICO, a new broader

class of higher-order operations [70] has recently been discovered, termed quantum operations with indefinite time direction or indefinite input-output direction [71]. These operations lack a fixed temporal direction and involve a coherent superposition of forward and backward time directions. In the backward time direction, the roles of input and output port are interchanged, defining quantum processes which receive the input in the future and output in the past timeline. Indefinite time direction can be realized in bidirectional quantum devices [71, 72] which include crystals that rotate the polarization of single photons such as half-wave plates, and quarter-wave plates. Analogous to the quantum switch in the context of ICO, the quantum "time flip" (TF) [71, 72], a paradigmatic example of indefinite time directed processes, has shown promising advantages in certain operator discrimination tasks [71], communication capacity [72], emergence of dynamical memory effect in memoryless phase-covariant processes [73], and also answering foundational queries such as the thermodynamic "arrow of time" and work-extraction [74, 75]. Furthermore, its feasibility has been established through experimental implementations [76, 77]. Despite these advancements, the potential applications of the quantum TF in quantum technologies remain largely unexplored (although see a recent work which uses quantum time flip [78] in continuous variable systems for estimating optical phase and light-beam rotation with photon number and orbital angular momentum being resources respectively).

In this paper, we bridge the gap by introducing a novel framework for quantum metrology that consists of encoding operations with indefinite input-output direction. By exploiting the properties of quantum TF, we devise a protocol that enables Heisenberg-limited precision in the estimation of an arbitrary single parameter of unitary operator, which we refer to as indefinite time directed metrology (ITDM). We prove that initializing the probe as pure product states in which all subsystems align in a specific direction, known as a symmetric product probe state, yields the maximum QFI using a time-flipped encoding approach. Further, we demonstrate that entanglement can yield a higher QFI than the product probes, but it cannot offer a scaling benefit beyond the HL in this encoding procedure. We also extract requirements for encoded state, in which ITDM cannot outperform SQL – we call the no-advantage condition (NAC).

We demonstrate the benefit of TF encoding for estimating phase-shift and axis of a unitary. Specifically, in the case of phase estimation, we devise a measurement strategy that operates exclusively on the control qubit yielding an FI that exhibits Heisenberg scaling. Interestingly, we exhibit that the parameter dependency of the optimal input probe state and FI can be eliminated when the parameter to be evaluated is small. Furthermore, when noise influences the encoding strategy and input probes, our protocol showcases the advantage of the time-flipped strategy over the existing quantum "switched" strategy [32] and regular strategy within specific parameter values and noise regimes. We also briefly analyze the effect of noise in multi qubit noisy ITDM and present a possible way to get a higher FI than the standard method, emphasizing the need for further investigation in this area.

The paper is organized as follows. In Sec. 2, we introduce the strategy based on time-flip encoding and derive the Heisenberg scaling for the general case. We present the optimality of symmetric product probe states among all product states and NAC condition with a measure of advantage in Sec. 2.2. In Sec. 3, we provide the applications of our procedure to phase estimation. We evaluate the use of entangled probes in Sec. 4 and demonstrate that entangled states cannot provide a higher QFI scaling than the product probes in this set-up. In Sec. 5, we investigate the advantages of our scheme under noise and compare them with existing protocols. Finally, the results are summarized in Sec. 6.

2 Framework for Indefinite Time Directed Metrology

Quantum time flip is defined on the set of bidirectional quantum processes which can operate either in forward or backward time directions. A quantum channel $\mathcal{C}$ is referred to as *bidirectional* if $\Omega(\mathcal{C})$ is a completely positive and trace-preserving (CPTP) map where Ω denotes the *input-output inversion map*. As demanded in Ref. [71], this map Ω should have certain properties including "order reversing", "identity preserving", $\Omega(\mathcal{A}) \neq \Omega(\mathcal{B})$ when channels $\mathcal{A}, \mathcal{B}$ are un-

equal, and $\Omega(q\mathcal{A}+(1-q)\mathcal{B}) = q\Omega(\mathcal{A})+(1-q)\Omega(\mathcal{B})$ with $1 \geq q \geq 0$. Moreover, the Kraus operators $\{C_i\}$ of $\mathcal{C}(*) = \sum_i C_i(*)C_i^\dagger$ satisfy bistochasticity conditions, i.e., $\sum_i C_i^\dagger C_i = \sum_i C_i C_i^\dagger = I$, with I being the identity matrix on the Hilbert space of the system. The only two possibilities of Ω are $\Omega(\mathcal{C})(*) = \mathcal{C}^\dagger(*) = \sum_i C_i^\dagger(*)C_i$ or $\Omega(\mathcal{C})(*) = \mathcal{C}^T(*) = \sum_i C_i^T(*)C_i^{T\dagger}$, up to unitary equivalence. $\mathcal{C}$ as the input of the quantum time flip supermap generates the time flipped channel, denoted as $T_\mathcal{C}$. The operation of $T_\mathcal{C}$ is described by $T_\mathcal{C}(\rho_c \otimes \rho) = \sum_i T_i^\mathcal{C}(\rho_c \otimes \rho)T_i^{\mathcal{C}\dagger}$ where $T_i^\mathcal{C} = |0\rangle\langle 0| \otimes C_i + |1\rangle\langle 1| \otimes \Omega(C_i)$ are the Kraus operators and the first register functions as the control for the time direction. We perform encoding in parameter estimation protocol with the aid of indefinite time ordering, defined as ITDM.

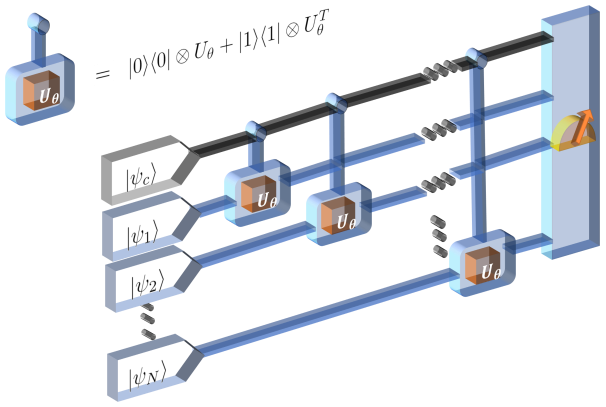

Figure 1: **ITDM scheme.** Taking N-qubit pure product probe states, $\{|\psi_i\rangle\}_{i=1}^N$ along with a control qubit, $|\psi_c\rangle$ we encode the parameter θ of U_θ on the probe state using time flipped gate, $T_{U_\theta} = |0\rangle\langle 0|\otimes U_\theta + |1\rangle\langle 1|\otimes U_\theta^T$. Therefore, the whole encoding procedure is performed in a single time step with N queries of T_{U_θ}, given by $\mathbb{T}_{U_\theta} = \prod_{i=1}^N (T_{U_\theta}^i \otimes \mathbb{I}_{N-1}^{\neq i}) = |0\rangle\langle 0|\otimes U_\theta^{\otimes N} + |1\rangle\langle 1|\otimes U_\theta^{T\otimes N}$ with $\mathbb{I}_{N-1}^{\neq i} = \bigotimes_{\substack{j=1 \\ j\neq i}}^N I_j$. This encoded state is then measured either locally on the control qubit or globally on the basis of symmetric logarithmic derivative operator to obtain the information about θ.

2.1 Time-flipped encoding strategy

In quantum estimation theory, given a unitary gate U which is a function of θ, i.e., $U \equiv U_\theta$, our task is to estimate θ. Note that the parameter θ can be directly related to a physical quantity of interest, such as the strength of a magnetic field, phase shift acquired by a photon's polarization state in a photonic setup. Instead of conventional unitary encoding schemes, we employ TF-encoding strategy comprising of N number of bipartite gates, $T_{U_\theta} = |0\rangle\langle 0|\otimes U_\theta + |1\rangle\langle 1|\otimes U_\theta^T$, acquired by applying the TF-supermap on U_θ, on N-qubit input probe, $\left|\psi^N\right\rangle$, assisted by a control qubit, $|\psi_c\rangle = \sqrt{p_c}\,|0\rangle + e^{i\theta_c}\sqrt{1-p_c}\,|1\rangle$ with $0 \leq p_c \leq 1$ and $0 \leq \theta_c \leq 2\pi$ (see Fig. 1). Mathematically, the state after encoding operation can be written as $|\Phi\rangle = \mathbb{T}_{U_\theta}|\Psi\rangle = \mathbb{T}_{U_\theta}|\psi_c\rangle \otimes \left|\psi^N\right\rangle$, where

$$\begin{aligned}\mathbb{T}_{U_\theta} &= \prod_{i=1}^N (T_{U_\theta}^i \otimes \mathbb{I}_{N-1}^{\neq i}),\\ &= \prod_{i=1}^N \Big(|0\rangle\langle 0| \otimes I^{\otimes i-1} \otimes U_\theta \otimes I^{\otimes N-i} \\ &\qquad + |1\rangle\langle 1| \otimes I^{\otimes i-1} \otimes U_\theta^T \otimes I^{\otimes N-i} \Big),\\ &= |0\rangle\langle 0| \otimes U_\theta^{\otimes N} + |1\rangle\langle 1| \otimes U_\theta^{T\otimes N}, \end{aligned} \tag{1}$$

with $\mathbb{I}_{N-1}^{\neq i} = \bigotimes_{\substack{j=1 \\ j\neq i}}^N I_j$. Specifically, when $\left|\psi^N\right\rangle$ is fully separable, i.e., $\left|\psi^N\right\rangle = \bigotimes_{i=1}^N |\psi_i\rangle$, the encoded state can be written as

$$|\Phi\rangle = \sqrt{p_c}\,|0\rangle \bigotimes_{i=1}^N |\chi_i\rangle + e^{i\theta_c}\sqrt{1-p_c}\,|1\rangle \bigotimes_{i=1}^N |\omega_i\rangle\,, \tag{2}$$

where $U_\theta|\psi_i\rangle = |\chi_i\rangle$, $U_\theta^T|\psi_i\rangle = |\omega_i\rangle$. Subsequently, we analyze the performance of the estimation protocol with product probe states and compare its performance with the entangled ones.

2.2 Quantum Fisher information in ITDM

The goal here is to estimate the parameter θ involved in U_θ with a high precision which can be fulfilled by performing measurements on the encoded state. Importantly, for any kind of measurements, the precision in estimating θ with an unbiased estimator, quantified by the variance of θ, is bounded below by the Cramer-Rao bound via Fisher information, $\mathcal{F}_\theta$ as $(\Delta\theta)^2 \geq (\lambda\mathcal{F}_\theta)^{-1}$, where λ denotes the number of measurements available [2, 25]. In particular, when the decoding measurement is performed in the basis of symmetric logarithmic derivative (SLD), Cramer-Rao bound is saturated and FI is termed as quantum Fisher information. The QFI for a pure state $|\Xi\rangle$ can be written as

$$\mathcal{Q}_\theta = 4\left(\langle\dot{\Xi}|\dot{\Xi}\rangle - \left|\langle\dot{\Xi}|\Xi\rangle\right|^2\right), \tag{3}$$

where $|\dot{\Xi}\rangle = \frac{\mathrm{d}\,|\Xi\rangle}{\mathrm{d}\theta}$. For N number of uncorrelated particles, $(\Delta\theta)^2$ is known to scale as $1/N$, referred to as the standard quantum limit (SQL) while with the help of quantum entanglement, it is possible to diminish the error as $(\Delta\theta)^2 \sim 1/N^2$ [2–5], known as the Heisenberg limit (HL). It will be interesting to find that when encoding is performed with TF operator, how quantum Fisher information scales with N. In our case of product probes in Eq. (2), QFI takes the form as (see Appendix. A)

$$\begin{aligned}\mathcal{Q} &= 4\Big[p_c\left(\sum_{k=1}^{N}|\dot{\chi}_k| + \sum_{\substack{j,k=1,\\ j\neq k}}^{N}\alpha_k\alpha_j^*\right)\\ &\quad +(1-p_c)\left(\sum_{k=1}^{N}|\dot{\omega}_k| + \sum_{\substack{j,k=1,\\ j\neq k}}^{N}\beta_k\beta_j^*\right)\\ &\quad -\Big|p_c\sum_{k=1}^{N}\alpha_k + (1-p_c)\sum_{k=1}^{N}\beta_k\Big|^2\Big], \end{aligned} \tag{4}$$

where

$$\begin{aligned}\langle\dot{\chi}_k|\chi_k\rangle = \alpha_k, \quad \langle\dot{\chi}_k|\dot{\chi}_k\rangle = |\dot{\chi}_k|,\\ \langle\dot{\omega}_k|\omega_k\rangle = \beta_k, \quad \langle\dot{\omega}_k|\dot{\omega}_k\rangle = |\dot{\omega}_k|.\end{aligned} \tag{5}$$

To analyze the scaling of QFI, let us concentrate on the simplest scenario where all the qubits in the probe state are equal, i.e., $|\psi_i\rangle = |\psi\rangle\ \forall i$, leading to a N-qubit symmetric probe state while for the asymmetric one, $|\psi_i\rangle \neq |\psi\rangle\ \ \forall i$. Therefore, in the symmetric scenario, the QFI reduces to

$$\begin{aligned}&\mathcal{Q}^S\\ &= 4N^2\Big[p_c|\alpha|^2 + (1-p_c)|\beta|^2 - |p_c\alpha + (1-p_c)\beta|^2\Big]\\ &\quad +4N\Big[p_c\Big(|\dot{\chi}| - |\alpha|^2\Big) + (1-p_c)\Big(|\dot{\omega}| - |\beta|^2\Big)\Big],\\ &= 4N^2p_c(1-p_c)|\alpha-\beta|^2\\ &\quad +4N\Big[p_c\Big(|\dot{\chi}| - |\alpha|^2\Big) + (1-p_c)\Big(|\dot{\omega}| - |\beta|^2\Big)\Big],\end{aligned} \tag{6}$$

with $\alpha_k = \alpha,\ \beta_k = \beta,\ \ |\dot{\chi}_k| = |\dot{\chi}|$, and $|\dot{\omega}_k| = |\dot{\omega}|$ $\forall k$. From the above expression, it is clear that $\mathcal{Q}^S \sim N^2$ as $N \to \infty$, which implies that by using symmetric product state $\left|\Psi^S\right\rangle = |\psi_c\rangle\otimes|\psi\rangle^{\otimes N}$ and encoding θ by the time-flipped operation which comprises multiple two-qubit controlled gates (see Fig. 1), the precision limit of θ can go beyond SQL and can achieve HL.

Let us now inquire whether we can beat SQL for all kinds of parameterized unitaries. We find that this is not the case, specifically, ITDM can not surpass the standard quantum limit when $\alpha = \beta$. This condition, referred to as the "no advantage condition" (NAC), simplifies to $(\partial_\theta U_\theta)^\dagger U_\theta = (\partial_\theta U_\theta^T)^\dagger U_\theta^T$. If U_θ satisfies NAC, it implies that there do not exist any probe and control state for which one can go beyond the SQL. Notably, the NAC is satisfied when $U^T = \pm U$, although, additional solutions for U may also exist.

Motivated by no-advantage condition, we propose a measure of advantage (M_A) in our protocol as the value of coefficient of N^2 in Eq. (6). It turns out to be $M_A = 4p_c(1-p_c)\max_{|\psi\rangle}\left|\langle(\partial_\theta U_\theta)^\dagger U_\theta - (\partial_\theta U_\theta^T)^\dagger U_\theta^T\rangle\right|^2 = 4p_c(1-p_c)|\lambda_{\max}|^2$ where $\lambda_{\max}$ is the eigenvalue of $(\partial_\theta U_\theta)^\dagger U_\theta - (\partial_\theta U_\theta^T)^\dagger U_\theta^T$ with the biggest modulus. There is no advantage when either $p_c = 0$ and 1, i.e., when the time flip does not work or when NAC is satisfied where $|\lambda_{\max}|^2 = 0$ thereby, resulting in a vanishing measure. On the other hand, when the measure is non-zero, we can, in principle, set our probe to the eigenvector corresponding to $\lambda_{\max}$ and gain QFI with an $N^2 4p_c(1-p_c)|\lambda_{\max}|^2$.

It is now natural to ask – *Is the symmetric product probe state optimal for ITDM*? In other words, if we can exhibit that one cannot achieve higher scaling with asymmetric product probes than the one obtained with symmetric states, the query above is resolved. Indeed, we answer this question positively in the following Theorem.

Theorem 1. *Among all pure product states, given an asymmetric state, there exists a symmetric state which provides a higher value of QFI in ITDM.*

Proof. Given an asymmetric product probe state $|\Psi\rangle = |\psi_c\rangle\bigotimes_{i=1}^N|\psi_i\rangle$ with QFI $\mathcal{Q}^A$, we denote the set of N symmetric product states as $\mathcal{S}_\Psi := \{\left|\Psi_l^S\right\rangle\}_{l=1}^N$. Here, the state $\left|\Psi_l^S\right\rangle = |\psi_c\rangle\,|\psi_l\rangle^{\otimes N}$ in this set have QFI $\mathcal{Q}_l^S$. Without loss of generality, we assume that the optimal probe state among all symmetric product states in $\mathcal{S}_\Psi$ is $\left|\Psi_1^S\right\rangle$. Therefore, mathematically, we can write $\mathcal{Q}_1^S \geq \mathcal{Q}_l^S \quad \forall l \in [N]$, implying $\mathcal{Q}_1^S - \mathcal{Q}_{avg} \geq 0$, where $\mathcal{Q}_{avg} = \frac{1}{N}\sum_{l=1}^N \mathcal{Q}_l^S$. Let us now consider the quantity, $\mathcal{Q}_{avg} - \mathcal{Q}^A$. On the other hand, it

can be shown that (see Appendix. B)

$$\begin{aligned}&\mathcal{Q}_{avg} - \mathcal{Q}^A \\ &= 4p_c(1-p_c)\left(N\sum_{k=1}^{N}|\alpha_k-\beta_k|^2 - \left|\sum_{k=1}^{N}\alpha_k-\beta_k\right|^2\right), \\ &\geq 0. \end{aligned} \tag{7}$$

Therefore, we get $\mathcal{Q}_1^S \geq \mathcal{Q}^A$. Hence, the theorem is proved. □

This result establishes the fact that in the space of N-qubit product states, there always exist a symmetric product probe state that maximizes the QFI. Consequently, from Theorem 1, we can conclude the following.

Corollary 2. *Maximal QFI in the case of ITDM with product probe states is achieved by symmetric states of the form* $\left|\Psi^S\right\rangle = |\psi_c\rangle\,|\psi\rangle^{\otimes N}$ *where we have to optimize over only the doubly-parametrized set of states,* $|\psi\rangle = \sqrt{p_s}\,|0\rangle + e^{i\theta_s}\sqrt{1-p_s}\,|1\rangle$ *with* $\theta_s \in [0, 2\pi]$ *and* $p_s \in [0,1]$ *along with optimal control qubit and encoding axis.*

3 Estimating phases with ITDM

After establishing the advantage of TF operation in the context of quantum metrology, let us concentrate on an explicit phase estimation scheme and exhibit how one can achieve Heisenberg limit with TF operation. In particular, given a unitary $U_\theta = e^{-i\frac{\theta}{2}\vec{\sigma}.\hat{n}}$, our goal is to estimate the value of the parameter θ, where $\hat{n} = \hat{x}n_1 + \hat{y}n_2 + \hat{z}n_3$ is referred as the encoding axis. As realized in the previous section, by initializing the input state as $\left|\Psi^S\right\rangle = |\psi_c\rangle\,|\psi\rangle^{\otimes N}$, we find QFI in terms of

$$\begin{aligned}\alpha &= in_3(p_s-1/2) \\ &\quad + i\sqrt{p_s(1-p_s)}(n_1\cos\theta_s + n_2\sin\theta_s), \\ \beta &= in_3(p_s-1/2) \\ &\quad + i\sqrt{p_s(1-p_s)}(n_1\cos\theta_s - n_2\sin\theta_s), \\ |\dot{\chi}| &= 1/4, \text{ and } |\dot{\omega}| = 1/4, \end{aligned} \tag{8}$$

in Eq. (6). In particular, $\mathcal{Q}_\theta(A,B) = AN^2 + BN$, where

$$A = 4p_c(1-p_c) * 4p_s(1-p_s)n_2^2\sin^2\theta_s, \tag{9}$$

and

$$\begin{aligned}B &= 4p_c\Big[\frac{1}{4} - \Big\{n_3(p_s-1/2) \\ &\quad + \sqrt{p_s(1-p_s)}(n_1\cos\theta_s + n_2\sin\theta_s)\Big\}^2\Big] \\ &\quad + 4(1-p_c)\Big[\frac{1}{4} - \Big\{n_3(p_s-1/2) \\ &\quad + \sqrt{p_s(1-p_s)}(n_1\cos\theta_s - n_2\sin\theta_s)\Big\}^2\Big], \end{aligned} \tag{10}$$

is independent of the parameter θ itself which we want to estimate, and also on θ_c. Notice that NAC, in this case reduces to $(\vec{\sigma}.\hat{n})^T = \vec{\sigma}.\hat{n}$ which implies $\hat{n}_2 = 0$.

Let us now find the maximum value of QFI, $\bar{\mathcal{Q}}_\theta^{\hat{n}}$ for a given encoding direction $\hat{n}$, which requires optimization over the control and the probe state. To address this, let us group the parameters θ_s, p_s, p_c in a set $\boldsymbol{\mu}$, i.e., $\boldsymbol{\mu} = \{\theta_s, p_s, p_c\}$. Mathematically, the optimization can be cast into the form $\mathcal{Q}_\theta^{\hat{n}} = \max_{\boldsymbol{\mu}} \mathcal{Q}_\theta(A,B)$. Since optimization over multiple (three) parameters is not straightforward, we maximize the coefficient A to obtain M_A. Looking at the form of A in Eq. (9), we can easily find that the maximum value of $A = n_2^2$ with $\boldsymbol{\mu}_A = \{\frac{\pi}{2}, \frac{1}{2}, \frac{1}{2}\}$ which leads to QFI $\mathcal{Q}_\theta|_{\boldsymbol{\mu}_A} = N^2n_2^2 + N(1-n_2^2)$. Let us now prove in the following Theorem that calculating maximum advantage of time flip, i.e., M_A, maximizes the QFI.

Theorem 3. *The maximum value of QFI, achieved by a N-qubit symmetric product probe state in the ITDM protocol for phase estimation is* $\bar{\mathcal{Q}}_\theta^{\hat{n}} = N^2n_2^2 + N(1-n_2^2)$.

Proof. To demonstrate that $\mathcal{Q}_\theta|_{\boldsymbol{\mu}_A}$ represents the optimal value of the QFI, we have to prove $\mathcal{Q}_\theta|_{\boldsymbol{\mu}_A} - \mathcal{Q}_\theta(A,B) \geq 0$, for all other values of θ_s, p_s, p_c, i.e.,

$$\begin{aligned}&N^2n_2^2 + N(1-n_2^2) - (AN^2 + BN) \geq 0, \\ \implies\ &N(n_2^2 - A) \geq B + n_2^2 - 1. \end{aligned} \tag{11}$$

As n_2^2 is the optimal value of A, $n_2^2 - A \geq 0$. Therefore, if Eq. (11) is true for $N = 1$, then it remains valid for all values of $N > 1$. For $N = 1$, it simplifies to $A + B \leq 1$, which can be proven (see Appendix. C). Hence, $\bar{\mathcal{Q}}_\theta^{\hat{n}} = N^2n_2^2 + N(1-n_2^2)$. □

Let us now determine the optimal encoding direction $\hat{n}$ for which the maximum of QFI is attained. To do so, we define

$$\bar{\mathcal{Q}}_\theta = \max_{\hat{n}} \bar{\mathcal{Q}}_\theta^{\hat{n}} = \max_{\hat{n},\boldsymbol{\mu}} \mathcal{Q}_\theta(A, B), \tag{12}$$

which includes optimization over $\hat{n}$, as well as input and control qubits. As a direct consequence of Theorem 3, the following corollary can be established by choosing $n_2 = 1$.

Corollary 4. *By fixing the optimal encoding axis $\hat{n} = \hat{y}$ and $\boldsymbol{\mu}_A = \{\frac{\pi}{2}, \frac{1}{2}, \frac{1}{2}\}$, the maximum QFI turns out to be N^2, achieving Heisenberg scaling.*

Remark. The maximum QFI for the estimation of the axis parameter of $U = \exp\left(-i\frac{\theta}{2}\vec{\sigma}\cdot\hat{n}\right)$, where either ϕ or ξ of the axis $\hat{n} = \hat{x}\sin\phi\cos\xi + \hat{y}\sin\phi\sin\xi + \hat{z}\cos\phi$ is to be estimated, is determined to be $4N^2$ following Eq. (6) (for details, see Appendix. D). This explicitly demonstrates that our protocol based on TF is capable of achieving the Heisenberg limit for arbitrary parameters in unitaries, as expected from our derivations in Section 2. For example, it can be used to estimate the errors in the parameters of a gate-set [79], required for universal quantum computation.

Measurement scheme to attain maximum QFI. Let us exhibit that the maximum QFI, i.e., $\mathcal{Q}_\theta = N^2$ for optimal encoding is, in fact, achievable by performing local measurement during decoding. The scheme consists of the following steps: (1) Prepare the control state as $|\psi_c\rangle = |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and each qubit of the probe state as $|\psi\rangle = |{+i}\rangle = (|0\rangle + i\,|1\rangle)/\sqrt{2}$, i.e., the input state is $|\Psi^s\rangle = |+\rangle\,|{+i}\rangle^{\otimes N}$. (2) The encoded state after the application of $\mathbb{T}_{U_\theta}$ with $U_\theta = e^{-i\theta\sigma_y/2}$ (i.e., $\hat{n} = \hat{y}$) on $|\Psi_s\rangle$ is given by

$$|\Phi\rangle = \frac{1}{\sqrt{2}} e^{-iN\theta/2}\left(|0\rangle + e^{iN\theta}\,|1\rangle\right) \otimes |{+i}\rangle^{\otimes N}. \tag{13}$$

(3) Perform projective measurement on the control qubit in the basis of σ_x, i.e., $\{|+\rangle, |-\rangle\}$. (4) Repeat the above procedure λ number of times and record the measurement outcomes.

The probability statistics for the outcome $|\pm\rangle$ is given by $p(\pm|\theta) = \frac{1}{2}[1 \pm \cos(N\theta)]$. Therefore, the FI [25] coincides with $\bar{\mathcal{Q}}_\theta$, i.e.,

$$\mathcal{F}_\theta = \sum_{b\in\{+,-\}} p(b|\theta)\left[\frac{\partial \ln p(b|\theta)}{\partial\theta}\right]^2, \tag{14}$$

$$= N^2 = \bar{\mathcal{Q}}_\theta. \tag{15}$$

Notably, in this scenario, the probe and control states (Eq. (13)) remain as a pure product state even after bidriectional encoding along y-axis, which is a global operation, while still providing HL. This demonstrates that entanglement is not necessary for surpassing SQL in ITDM protocol. However, the two qubit unitaries generated by time flip may generate entanglement for arbitrary input product probe states.

The above-discussed protocol is optimal when $\hat{n} = \hat{y}$ and the root mean square error scales as $1/N$. However, it is interesting to ask – *Given an arbitrary direction of the encoding axis, how does FI, $\mathcal{F}_\theta$, scale when the decoding measurement basis is fixed in $\{|+\rangle, |-\rangle\}$?* We answer this question in the following subsection.

3.1 Fisher information by restricted measurements

By initializing the control and probe states as a product state (Eq. (2)) and encoding the axis as $\hat{n}$, let us consider the decoding measurement only on the control qubit in $\{|+\rangle, |-\rangle\}$ basis. In this scenario, the outcome statistics read

$$p(\pm|\theta) = \mathrm{Tr}\left[(|\pm\rangle\langle\pm|_c \otimes I_p)\,|\Phi\rangle\langle\Phi|\right]$$
$$= \frac{1}{2} \pm \frac{\sqrt{p_c(1-p_c)}}{2}\left(e^{i\theta_c}\prod_{i=1}^{N}\langle\chi_i|\omega_i\rangle + e^{-i\theta_c}\prod_{i=1}^{N}\langle\omega_i|\chi_i\rangle\right). \tag{16}$$

Since $\langle\chi_i|\omega_i\rangle$ is a complex number, we parametrize it as $\langle\chi_i|\omega_i\rangle = r_i(\boldsymbol{\nu})e^{if_i(\boldsymbol{\nu})}$ where $\boldsymbol{\nu} = \{\theta, \theta_s, p_s, \hat{n}\}$. According to this parametrization,

$$p(\pm|\theta)$$
$$= \frac{1}{2} \pm \sqrt{p_c(1-p_c)}\cos\left(\sum_{i=1}^{N} f_i(\boldsymbol{\nu}) + \theta_c\right)\prod_{i=1}^{N} r_i(\boldsymbol{\nu}), \tag{17}$$

and the FI can be calculated as

$$\mathcal{F}_\theta$$
$$= 4p_c(1-p_c)\Big[\sum_{j=1}^{N}\Big(\prod_{i\neq j}^{N} r_i(\boldsymbol{\nu})\Big)\Big\{\dot{r}_j(\boldsymbol{\nu})\cos\Big(\sum_{i=1}^{N} f_i(\boldsymbol{\nu})+\theta_c\Big)-\Big(\frac{r_j(\boldsymbol{\nu})}{N}\Big)\sin\Big(\sum_{i=1}^{N} f_i(\boldsymbol{\nu})+\theta_c\Big)\Big(\sum_{i=1}^{N}\dot{f}_i(\boldsymbol{\nu})\Big)\Big\}\Big]^2$$
$$\times\left[1-4p_c(1-p_c)\Big(\prod_{i=1}^{N} r_i(\boldsymbol{\nu})\Big)^2\cos^2\Big(\sum_{i=1}^{N} f_i(\boldsymbol{\nu})+\theta_c\Big)\right]^{-1}, \tag{18}$$

which can be shown to be optimal at $p_c = 1/2$. We consider $|\psi_i\rangle = |\psi\rangle$ for all i, which is determined to be optimal based on rigorous numerical investigations. This implies $r_i(\boldsymbol{\nu}) = r(\boldsymbol{\nu})$, $f_i(\boldsymbol{\nu}) = f(\boldsymbol{\nu})\ \forall i$ and along with $p_c = 1/2$, we have

$$\mathcal{F}_\theta = \frac{N^2 r^{2N-2}(\boldsymbol{\nu})\left\{\dot{r}(\boldsymbol{\nu})\cos\left(Nf(\boldsymbol{\nu})+\theta_c\right) - r(\boldsymbol{\nu})\dot{f}(\boldsymbol{\nu})\sin\left(Nf(\boldsymbol{\nu})+\theta_c\right)\right\}^2}{1-r^{2N}(\boldsymbol{\nu})\cos^2\left(Nf(\boldsymbol{\nu})+\theta_c\right)}. \tag{19}$$

To approach the question of the optimal probe state, similar to QFI, we define

$$\bar{\mathcal{F}}_\theta = \max_{\hat{n}} \bar{\mathcal{F}}_\theta^{\hat{n}}, \tag{20}$$

where $\bar{\mathcal{F}}_\theta^{\hat{n}}$ is the maximum FI for a given axis $\hat{n}$ by choosing optimal probe and performing decoding measurement in $\{|+\rangle, |-\rangle\}$ basis on optimal control qubit, i.e., $\bar{\mathcal{F}}_\theta^{\hat{n}} = \max_{\boldsymbol{\mu},\theta_c} \mathcal{F}_\theta$. Note that $U_\theta^\dagger U_\theta^T$ being unitary, we have $0 \leq r(\boldsymbol{\nu}) \leq 1$. Therefore, from Eq. (19), it is evident that if $r(\boldsymbol{\nu}) < 1$, $\mathcal{F}_\theta$ decreases exponentially with increasing value of N. This means that it is desirable to set $r(\boldsymbol{\nu}) = 1$ by taking initial probe state $|\psi\rangle = |u(\hat{n},\theta)^\pm\rangle$ where $|u(\hat{n},\theta)^\pm\rangle = (\mathrm{g}^\mp(\hat{n},\theta)|0\rangle + |1\rangle)/\sqrt{1+|\mathrm{g}^\mp(\hat{n},\theta)|^2}$ is the eigenstate of $U_\theta^\dagger U_\theta^T$, i.e., $U_\theta^\dagger U_\theta^T |u(\hat{n},\theta)^\pm\rangle = \exp\left(\pm i\mathrm{f}(n_2,\theta)\right)|u(\hat{n},\theta)^\pm\rangle$. In this scenario, $p_s = \mathrm{g}^\mp(\hat{n},\theta)^2/\sqrt{1+|\mathrm{g}^\mp(\hat{n},\theta)|^2}$ and $\theta_s = i\ln(1+|\mathrm{g}^\mp(\hat{n},\theta)|^2 - \mathrm{g}^\mp(\hat{n},\theta)^2)/2$. Moreover, we have

$$\mathrm{f}(n_2,\theta) = \tan^{-1}\left(\frac{\sqrt{4n_2^2\sin^2\frac{\theta}{2}\left(1-n_2^2\sin^2\frac{\theta}{2}\right)}}{1-2n_2^2\sin^2\frac{\theta}{2}}\right), \tag{21}$$

and the explicit forms of $g^\mp(\hat{n},\theta)$ are given in Appendix. E, which shows that the eigenvectors depends on the encoding axis and the parameter θ itself. This implies $\dot{r}(\boldsymbol{\nu}) = 0$ and $f(\boldsymbol{\nu}) = \mathrm{f}(n_2,\theta)$. Therefore, preparing the probe $|\psi\rangle^{\otimes N}$ in either $|u(\hat{n},\theta)^+\rangle^{\otimes N}$ or $|u(\hat{n},\theta)^-\rangle^{\otimes N}$ and $|\psi_c\rangle = (|0\rangle + e^{i\theta_c}|1\rangle)/\sqrt{2}$ and measuring the control qubit in the $\{|+\rangle, |-\rangle\}$ basis lead to FI (following Eq. (19))

$$\bar{\mathcal{F}}_\theta^{\hat{n}} = N^2\dot{\mathrm{f}}(n_2,\theta)^2. \tag{22}$$

However, the complexity is two-folded – (1) the input state parameters, θ_s, p_s, is θ-dependent, (2) $\bar{\mathcal{F}}_\theta^{\hat{n}}$ is θ-dependent which means that the estimation accuracy may fluctuate or degrade depending on the true value of θ. Notice that choosing $\hat{n} = \pm\hat{y}$ gets rid of such complexities and gives $\mathrm{f} = -\theta \pm \pi$ which achieves $\bar{\mathcal{F}}_\theta = N^2 = \bar{\mathcal{Q}}_\theta$ which is consistent with Eq. (15). Next, we show that the above-mentioned subtlety can be removed when the parameter θ is taken to be small.

3.1.1 Optimal input probes for small θ

Let us write the input probe as $\rho = |\psi\rangle\langle\psi| = (I + \hat{s}.\vec{\sigma})/2$ where $\hat{s} = \hat{x}s_x + \hat{y}s_y + \hat{z}s_z$ and $s_x = 2\sqrt{p_s(1-p_s)}\cos\theta_s$, $s_y = 2\sqrt{p_s(1-p_s)}\sin\theta_s$, and $s_z = (2p_s - 1)$. The overlap $\langle\chi|\omega\rangle$ can be written as

$$\begin{aligned}\langle\chi|\omega\rangle &= \langle\psi|U_\theta^\dagger U_\theta^T|\psi\rangle = \mathrm{Tr}\left(U_\theta^T\rho U_\theta^\dagger\right)\\ &= 1 - n_2^2(1-\cos\theta) + in_2n_3s_x(1-\cos\theta)\\ &\quad -in_1n_2s_z(1-\cos\theta) + in_2s_y\sin\theta. \end{aligned}\tag{23}$$

Therefore, the modulus and phase of the complex number in Eq. (23) is given by

$$r(\boldsymbol{\nu}) = \Big[\Big\{1 - n_2^2(1-\cos\theta)\Big\}^2 + \{n_2 s_y \sin\theta + (n_2 n_3 s_x - n_1 n_2 s_z)(1-\cos\theta)\}^2\Big]^{1/2}, \tag{24}$$

and

$$f(\boldsymbol{\nu}) = \tan^{-1}\frac{n_2 s_y \sin\theta + (n_2 n_3 s_x - n_1 n_2 s_z)(1-\cos\theta)}{1 - n_2^2(1-\cos\theta)}, \tag{25}$$

respectively. When θ is very small ($\theta = \tilde{\theta}$), expanding Eqs. (24) and (25) upto second order in θ we obtain (see details in Appendix. F)

$$\begin{aligned}\mathcal{F}_{\tilde{\theta}} &= s_y^2 n_2^2 N^2 + (1-s_y^2) n_2^2 N \delta_{\theta_c,0}, \\ &= s_y^2 n_2^2 N(N - \delta_{\theta_c,0}) + n_2^2 N \delta_{\theta_c,0}, \end{aligned} \tag{26}$$

where $\delta_{\theta_c,0} = 1$ if $\theta_c = 0$ and $\delta_{\theta_c,0} = 0$ elsewhere. Interestingly, in the limit $N \to \infty$, Fisher information $\mathcal{F}_{\tilde{\theta}}|_{N\to\infty} = s_y^2 n_2^2 N^2$ is $\tilde{\theta}$-independent and reaches HL irrespective of any probe state. From Eq. (26), it immediately follows that the optimal probe for $\tilde{\theta}$ estimation is given by $s_y = \pm 1$, i.e., $\boldsymbol{\mu} = \boldsymbol{\mu}_A = \{\pi/2 + 2m\pi, 1/2, 1/2\} \forall m \in \mathbb{Z}$ which implies

$$\bar{\mathcal{F}}_{\tilde{\theta}}^{\hat{n}} = n_2^2 N^2 = \bar{\mathcal{Q}}_{\tilde{\theta}}^{\hat{n}}|_{N\to\infty}. \tag{27}$$

Hence, we arrive at the following result, which shows that the two-fold complexity previously described can be eradicated under the given scenario:

Theorem 5. *The FI, $\bar{\mathcal{F}}_{\tilde{\theta}}^{\hat{n}} = n_2^2 N^2$, independent of $\tilde{\theta}$ to be estimated which is small, saturates the QFI, $\bar{\mathcal{Q}}_{\tilde{\theta}}^{\hat{n}}|_{N\to\infty}$ asymptotically, using the control qubit $|\psi_c\rangle = (|0\rangle + e^{i\theta_c}|1\rangle)/\sqrt{2}$ along with the probe state as $|\pm i\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$ and the decoding measurement in the σ_x basis.*

4 Entanglement can not beat N^2-scaling

Our analysis upto now has focused on utilizing product states as the initial probe state. Let us now examine whether the use of an entangled probe can improve scaling, surpassing the Heisenberg limit or not. To address the same, we write the general N-qubit probe state as

$$|\Psi\rangle = \sum_{i=1}^{2^N} \Lambda_i |\lambda_i\rangle, \tag{28}$$

where $\sum_{i=1}^{2^N} |\lambda_i\rangle\langle\lambda_i| = I_{2^N}$ and $\langle\lambda_i|\lambda_j\rangle = \delta_{ij}$. Defining $|\lambda_i\rangle = \bigotimes_{j=1}^N |\psi_j^i\rangle$, we rewrite $|\Psi\rangle = \sum_{i=1}^{2^N} \Lambda_i \bigotimes_{j=1}^N |\psi_j^i\rangle$, where $|\psi_j^i\rangle$ is the ith basis state for the jth qubit with $\langle\psi_j^i|\psi_j^k\rangle = \delta_{ik}$ and $\sum_{i=1}^2 |\psi_j^i\rangle\langle\psi_j^i| = I_2$. Extending beyond the use of product states as probes, our theorem below indicates that entanglement does not provide any additional advantage in achieving scaling of the QFI with the number of qubits beyond the HL.

Theorem 6. *In the ITDM protocol, the maximum QFI extractible from any probe state, including all entangled states, is upper bounded by $\gamma N^2 + \zeta N$, where γ and ζ are constants, independent of N.*

Proof. Let us present a brief outline of our comprehensive proof, (for details, see Appendix. G). Using the general N-qubit probe state in Eq. (28), we derive an upper bound on QFI as

$$\begin{aligned}\mathcal{Q}_\theta \le \tilde{\mathcal{Q}}_\theta &= 4\langle\dot{\Phi}|\dot{\Phi}\rangle, \\ &= 4\sum_{i=1}^{2^N}\sum_{k=1}^{2^N} \Lambda_i \Lambda_k^* \Bigg[p_c \sum_{m=1}^N \Big(\dot{\chi}_m^{ki} \prod_{j\neq m} a_j^{ki} + \sum_{l\neq m} \alpha_m^{ki}\alpha_l^{ik*} \prod_{j\neq m,l} a_j^{ki}\Big) \\ &\quad + (1-p_c)\sum_{m=1}^N \Big(\dot{\omega}_m^{ki} \prod_{j\neq m} b_j^{ki} + \sum_{l\neq m} \beta_m^{ki}\beta_l^{ik*} \prod_{j\neq m,l} b_j^{ki}\Big)\Bigg], \end{aligned} \tag{29}$$

where

$$\begin{aligned}\langle\dot{\chi}_m^k|\dot{\chi}_m^i\rangle &= \dot{\chi}_m^{ki}, \quad \langle\chi_j^k|\chi_j^i\rangle = a_j^{ki}, \\ \langle\dot{\omega}_m^k|\dot{\omega}_m^i\rangle &= \dot{\omega}_m^{ki}, \quad \langle\omega_j^k|\omega_j^i\rangle = b_j^{ki}, \\ \langle\dot{\chi}_m^k|\chi_m^i\rangle &= \alpha_m^{ki}, \quad \langle\dot{\omega}_j^k|\omega_j^i\rangle = \beta_j^{ki}, \end{aligned} \tag{30}$$

and $|\chi_j^i\rangle = U_\theta |\psi_j^i\rangle$ and $|\omega_j^i\rangle = U_\theta^T |\psi_j^i\rangle$. In Eq. (29), the terms $\prod_{j\neq m} a_j^{ki}$, and $\prod_{j\neq m} b_j^{ki}$ are the products of the inner products of $(N-1)$

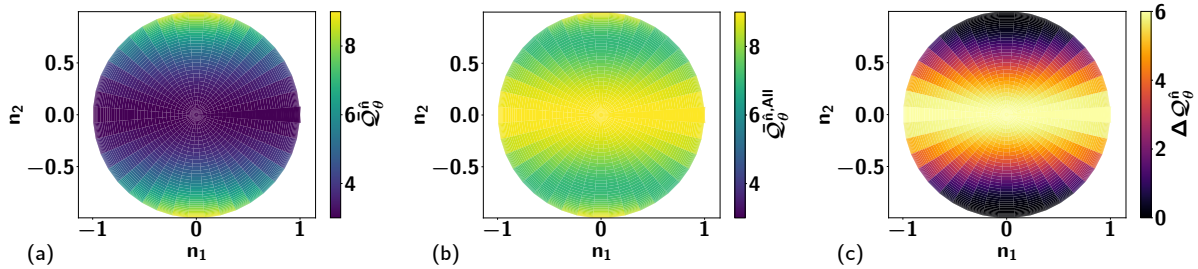

Figure 2: **QFI using three-qubit probe states for estimating θ in U_θ.** Map plot of (a) $\bar{\mathcal{Q}}_\theta^{\hat{n}}$, optimized over all product states, (b) $\bar{\mathcal{Q}}_\theta^{\hat{n},All}$ optimized over all states, (c) $\Delta\mathcal{Q}_\theta^{\hat{n}} = \bar{\mathcal{Q}}_\theta^{\hat{n},All} - \bar{\mathcal{Q}}_\theta^{\hat{n}}$ against encoding axis, n_1 (horizontal axis) and n_2 (vertical axis). We see that the difference vanishes near $n_2 = \pm 1$ precisely where both (a) and (b) attain their maximal QFI of $N^2 = 9$. We conclude that product states can give as much QFI as entangled states when optimized over the axis for a three-qubit probe system. All the axis are dimensionless.

qubits. Hence, if the qubits in $|\lambda_i\rangle$ and $\langle\lambda_k|$ differ in more than one position, $\prod_{j\neq m} a_j^{ki} = \prod_{j\neq m} b_j^{ki} = 0$. Similarly, if the qubits in $|\lambda_i\rangle$ and $\langle\lambda_k|$ differ in more than two positions, $\prod_{j\neq m,l} b_j^{ki} = \prod_{j\neq m,l} a_j^{ki} = 0$. Defining, $\tilde{\mathcal{Q}}_\theta^X$ as the term which involves contribution only from N-qubit basis states where qubits differ in X number of positions, we can write $\tilde{\mathcal{Q}}_\theta = \sum_{X=0}^{N} \tilde{\mathcal{Q}}_\theta^X = \sum_{X=0}^{2} \tilde{\mathcal{Q}}_\theta^X$ since, $\tilde{\mathcal{Q}}_\theta^X = 0\ \forall X \geq 3$. In Appendix. G, we show that each of these three terms is bounded by an expression of the form $\gamma_i N^2 + \zeta_i N$. Hence, $\tilde{\mathcal{Q}}_\theta \leq \gamma N^2 + \zeta N$ where γ, ζ are constants, which are independent of N. □

Although entangled states do not offer superior scaling compared to product probe states, we investigate whether they can yield higher QFI. To this aim, given an encoding axis $\hat{n}$, we perform numerical maximization of QFI over the whole state space (see Eq. (28)) for $N = 3$ denoted as $\bar{\mathcal{Q}}_\theta^{\hat{n},All}$ (see Fig. 2(b)). Interestingly, for $n_2 = 1$, $\bar{\mathcal{Q}}_\theta^{\hat{y},All}$ is the same as that with QFI for optimal product probe state, i.e., $\bar{\mathcal{Q}}_\theta^{\hat{y},All} = \bar{\mathcal{Q}}_\theta^{\hat{y}} = 9 = \bar{\mathcal{Q}}_\theta$. Notice that $\hat{n} = \hat{y}$ is indeed the universal optimal encoding axis which means maximal attainable value ($N^=9$) of QFI by optimizing over the whole state space and the encoding axis. However, as value of n_2 deviates from unity, the difference between the maximum value of QFI with optimal product probe and $\bar{\mathcal{Q}}_\theta^{\hat{n},All}$, i.e., $\Delta\mathcal{Q}_\theta^{\hat{n}} = \bar{\mathcal{Q}}_\theta^{\hat{n},All} - \bar{\mathcal{Q}}_\theta^{\hat{n}}$ increases significantly as demonstrated in Fig. 2(c). For $n_2 = 0$, $\bar{\mathcal{Q}}_\theta^{\hat{n},All} = 9$ and $\bar{\mathcal{Q}}_\theta^{\hat{n}} = 3$ where the difference $\Delta\mathcal{Q}_\theta^{\hat{n}}$ reaches a maximum value of 6. This observation ensures that the optimal probe states which give higher QFI than fully separable probe states must be entangled in at least some bipartition although no scaling benefit is provided by the entangled probes. Notice that $\Delta\mathcal{Q}_\theta^{\hat{n}} \approx \mathcal{O}(10^{-2})$ in the region $\pm 0.95 \lesssim n_2 \leq \pm 1$ irrespective of n_1 and n_3 which shows the robustness of optimal product probe states compared to the entangled ones against nonoptimal encoding.

5 Advantage in noisy unitary encoding

Both theoretical and experimental perspectives of quantum metrology demand the investigation on how the protocol gets affected in the presence of noisy environment. Moreover, it was shown that the precision in estimating parameters in a noisy scenario can also be improved by using indefinite causal order [32]. Therefore, it is interesting to analyze the ITDM scheme under noise. For inspection, we consider a noisy input probe, $\rho = (I + \vec{s}.\vec{\sigma})/2$ with $|\vec{s}| \leq 1$, which is influenced by depolarizing noise, $\mathcal{N}(\rho) = q\rho + (1-q)I/2$ [80] after unitary encoding U_θ. Consequently, the effective channel, $\mathcal{N}_\theta(\rho) = qU_\theta\rho U_\theta^\dagger + (1-q)I/2$ is termed as noisy unitary channel with Kraus operators $\{K_0 = \sqrt{1-p}U_\theta, K_1 = \sqrt{p/3}\sigma_x U_\theta, K_2 = \sqrt{p/3}\sigma_y U_\theta, K_3 = \sqrt{p/3}\sigma_z U_\theta\}$ where $q = 1 - 4p/3 \in [0,1]$. Application of time-flip to such a bistochastic channel [71] leads to a quantum operation $T_{\mathcal{N}_\theta}(\rho_c \otimes \rho) = \sum_{j=0}^{3} F_j(\rho_c \otimes \rho)F_j^\dagger$ with Kraus operators $F_j = |0\rangle\langle 0| \otimes K_j + |1\rangle\langle 1| \otimes K_j^T$ where $\rho_c = |\psi_c\rangle\langle\psi_c|$. Therefore, after the encoding procedure in ITDM scheme with noisy encoding operation (NITDM), the output state is given by $\Theta = \mathbb{T}_{\mathcal{N}_\theta}(\rho_c \bigotimes_{i=1}^{N} \rho_i) = \sum_{j\in\Pi} \mathbb{F}_j(\rho_c \bigotimes_{i=1}^{N} \rho_i)\mathbb{F}_j^\dagger$ with $\mathbb{F}_j = |0\rangle\langle 0| \bigotimes_{i=1}^{N} K_{j[i]} + |1\rangle\langle 1| \bigotimes_{i=1}^{N} K_{j[i]}^T$, where Π is the set of 4^N integer strings j, with each string specifying the Kraus operator $K_{j[i]}$

acted on the qubit i, and $\sum_{j\in\Pi}\mathbb{F}_j^\dagger\mathbb{F}_j = I$ and $\sum_{j\in\Pi}\left(\bigotimes_{i=1}^N K_{j[i]}\right)^\dagger\left(\bigotimes_{i=1}^N K_{j[i]}\right) = I$. This can further be simplified as

$$\begin{aligned}\Theta = p_c|0\rangle\langle 0|\bigotimes_{i=1}^N\left(\sum_{j=0}^3 K_j\rho_i K_j^\dagger\right)\\ + \sqrt{p_c(1-p_c)}e^{-i\theta_c}|0\rangle\langle 1|\bigotimes_{i=1}^N\left(\sum_{j=0}^3 K_j\rho_i K_j^{T\dagger}\right)\\ + \sqrt{p_c(1-p_c)}e^{i\theta_c}|1\rangle\langle 0|\bigotimes_{i=1}^N\left(\sum_{j=0}^3 K_j^T\rho_i K_j^\dagger\right)\\ + (1-p_c)|1\rangle\langle 1|\bigotimes_{i=1}^N\left(\sum_{j=0}^3 K_j^T\rho_i K_j^{T\dagger}\right). \quad (31)\end{aligned}$$

Defining $\mathrm{Tr}\left(\sum_j K_j^T\rho_i K_j^\dagger\right) = r_i(\boldsymbol{\nu})e^{if_i(\boldsymbol{\nu})}$ with $\boldsymbol{\nu} = \{\theta, \vec{s}, \hat{n}, q\}$ as the updated set $\boldsymbol{\nu}$, outcome probabilities of measurement in the σ_x-basis on the control qubit can be written as

$$\begin{aligned}&p(\pm|\theta)\\ &= \frac{1}{2} \pm \sqrt{p_c(1-p_c)}\cos\left(\sum_{i=1}^N f_i(\boldsymbol{\nu}) + \theta_c\right)\prod_{i=1}^N r_i(\boldsymbol{\nu}).\end{aligned} \quad (32)$$

Following Eq. (14), we calculate FI, $\mathcal{F}_\theta^q$ in NITDM which have exactly same mathematical form as Eq. (19), i.e., $\mathcal{F}_\theta^q \equiv \mathcal{F}_\theta$ with updated $r_i(\boldsymbol{\nu}) = \left|\mathrm{Tr}\left(\sum_j K_j^T\rho_i K_j^\dagger\right)\right|$ and $f_i(\boldsymbol{\nu}) = \arg\left[\mathrm{Tr}\left(\sum_j K_j^T\rho_i K_j^\dagger\right)\right]$. With the tools in hand, we are now ready to analyze the effect of noisy encoding.

5.1 Comparison of NITDM with "switched" and regular strategy in single qubit $N = 1$

In the case of $N = 1$, the value of FI corresponding to NITDM is given by (see Appendix. H)

$$\mathcal{F}_\theta^q = \frac{n_2^4 q^2\sin^2\theta}{1 - \left\{\frac{1+q}{2} - n_2^2 q(1-\cos\theta)\right\}^2}, \quad (33)$$

which is independent of the input probe configuration. Being periodic in θ, we can safely choose $\theta \in [0, 2\pi)$ for our analysis. We compare NITDM with two other metrological schemes – (1) "switched" strategy, where encoding is done using the quantum switch as described in Ref. [32], (2) regular strategy [32] where conventional unitary encoding is used. The "switched" FI, $\mathcal{F}_\theta^{q,S}$ and QFI, $\mathcal{Q}_\theta^{q,R}$ in a regular strategy are given by $\mathcal{F}_\theta^{q,S} = \tilde{q}^2\sin^4\theta\left[1 - \left\{\tilde{q}\cos\theta + \frac{1}{4}(1+q)^2\right\}^2\right]^{-1}$ with $\tilde{q} = q(1-q)$, and $\mathcal{Q}_\theta^{q,R} = q^2(\vec{n}\times\vec{s})^2$ respectively. Let us now compare these three strategies in terms of their metrological power.

5.1.1 Fixing optimally encoded direction

Notice that $\mathcal{F}_\theta^q$ is optimal for encoding axis $\hat{n} = \hat{y}$ (see Appendix. H) where optimal QFI for regular probe $\mathcal{Q}_\theta^{q,R} = q^2$ is attained by choosing $\hat{s} \perp \hat{n}$. Due to the θ-dependency of the FI in the NITDM, there exist specific regions of θ where this protocol demonstrates superior performance compared to the regular one, as well as regions where it lags behind. Specifically, $\mathcal{F}_\theta^q|_{\hat{n}=\hat{y}}$ provides TF advantage compared to $\mathcal{Q}_\theta^{q,R}|_{\hat{s}\perp\hat{n}} = q^2$ in the parameter region $\pi/3 < \theta < 4\tan^{-1}\left(\sqrt{q_1 - 4\sqrt{q_2}}\right)$ and $4\tan^{-1}\left(\sqrt{q_1 + 4\sqrt{q_2}}\right) < \theta < 5\pi/3$ where $q_1 = (5+7q)/(3+q)$ and $q_2 = (1+4q+3q^2)/(3+q)^2$. In contrast, in the regime $\cos^{-1}\left[\frac{-1+q}{2(1+q)}\right] < \theta < 2\pi - \cos^{-1}\left[\frac{-1+q}{2(1+q)}\right]$; $0 \le \theta < \pi/3$ and $5\pi/3 < \theta < 2\pi/3$, we have $\mathcal{F}_\theta^q|_{\hat{n}=\hat{y}} < \mathcal{Q}_\theta^{q,R}|_{\hat{s}\perp\hat{n}}$, i.e., regular strategy outperforms NITDM.

Note that these ranges of θ are reliant on the strength of the noise parameter q. Interestingly, there exist intervals of θ where NITDM surpasses regular strategy irrespective of the strength of the noise. Our analysis reveals that, $\mathcal{F}_\theta^q|_{\hat{n}=\hat{y}} > \mathcal{Q}_\theta^{q,R}|_{\hat{s}\perp\hat{n}}\,\forall q$ when $\{\pi/3 < \theta < \pi/2; 3\pi/2 < \theta < 5\pi/3\}$ while $\mathcal{F}_\theta^q|_{\hat{n}=\hat{y}} < \mathcal{Q}_\theta^{q,R}|_{\hat{s}\perp\hat{n}}\,\forall q$ for $\{2\pi/3 < \theta < 4\pi/3; 0 \le \theta < \pi/3; 5\pi/3 < \theta < 2\pi/3\}$. Moreover, setting $\hat{n} = \hat{y}$ as the encoding axis, we always obtain the benefit over the switched scheme, i.e., $\mathcal{F}_\theta^q|_{\hat{n}=\hat{y}} \ge \mathcal{F}_\theta^{q,S}$ (Appendix. H) for any value of θ and arbitrary noise strength q.

Arbitrary probe state. Since $\mathcal{Q}_\theta^{q,R}$ is dependent on probe states, choosing nonoptimal probe states can reduce the performance of regular strategy whereas NITDM bypasses the probe dependency and illustrates supremacy as shown in Fig. 3. In particular, we illustrate the advantage of time-flip by choosing probe state ρ with $s_x = 0.8, s_y = 0.6$, $\mathcal{F}_\theta^q$ can be shown to outperform both $\mathcal{F}_\theta^{q,S}$ and $\mathcal{Q}_\theta^{q,R}$ with $\theta = \pi/4$. Interestingly, even when the probe is in a maximally

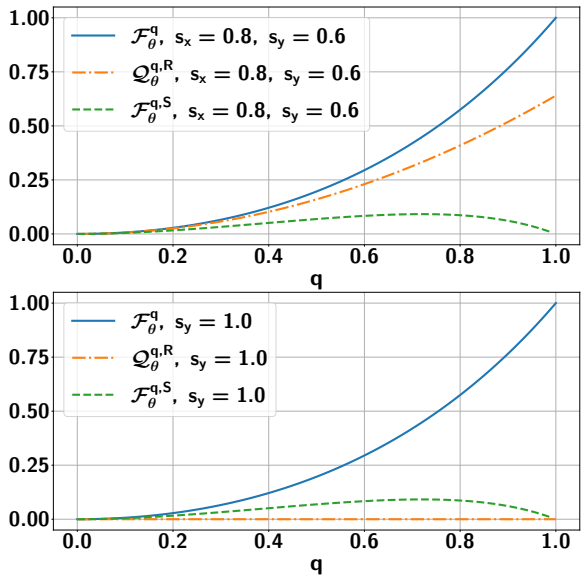

Figure 3: **Fisher information (vertical axis) in the case of time-flipped, F_θ^q, "switched", $\mathcal{F}_\theta^{q,S}$ and regular protocols $\mathcal{Q}_\theta^{q,R}$. with respect to noise parameter, q (horizontal axis).** Choosing $\hat{n} = \hat{y}$ with $\theta = \pi/4$, FI for NITDM, $\mathcal{F}_\theta^q$ (blue lines) always outperforms switched FI, $\mathcal{F}_\theta^{q,S}$ (green lines) which are independent of probe state. $\mathcal{F}_\theta^q$ is higher than $\mathcal{Q}_\theta^{q,R}$ (orange lines) with both $\hat{s} = [0.8, 0.6, 0]$ and $\hat{s} = \hat{y}$ probe state which are not optimal for regular strategy. All axes are dimensionless.

mixed state or orthogonal to $\hat{n}$, $\mathcal{F}_\theta^q$ is always positive (except for $q = 0$ and $\theta \neq 2m\pi \forall \pi \in \mathbb{Z}$) where $\mathcal{Q}_\theta^{q,R} = 0$. Moreover, the dominance of NITDM over switched and regular strategies exists in the low noise region even with nonoptimal encoding as illustrated in Appendix. H.

5.1.2 θ-averaged performance

Due to the θ-dependence of $\mathcal{F}_\theta^q$ and $\mathcal{F}_\theta^{q,S}$, we perform averaging over θ, i.e.,

$$\begin{aligned}\langle \mathcal{F}_\theta^q \rangle &= \int_0^{2\pi} \frac{1}{2\pi} \mathcal{F}_\theta^q d\theta \\ &= 1 - \frac{1}{4}\sqrt{(3+q)(3+q-4qn_2^2)} \\ &\quad - \frac{1}{4}\sqrt{(1-q)(1-q+4qn_2^2)}, \quad (34)\end{aligned}$$

and $\langle \mathcal{F}_\theta^{q,S} \rangle = 1 - \frac{\sqrt{3}}{8}(1-q)\sqrt{(1-q)(3+5q)} - \frac{1}{8}\sqrt{(5+6q-3q^2)(5-2q+5q^2)}$ [32]. Given the noise strength q, although $\langle \mathcal{F}_\theta^q \rangle$ and $\langle \mathcal{F}_\theta^{q,S} \rangle$ are both less than the optimal regular QFI, $\mathcal{Q}_\theta^{q,R}|_{\hat{n}=\hat{y}} = q^2 = \langle \mathcal{Q}_\theta^{q,R}|_{\hat{n}=\hat{y}} \rangle$, in the range $\pm n_2^{\min} \leq n_2 \leq \pm 1$, the "flipped" strategy is better than the "switched" one, i.e., $\langle \mathcal{F}_\theta^q \rangle > \langle \mathcal{F}_\theta^{q,S} \rangle$. Here, $n_2^{\min} = n_2^0[1-q^g]^{h^{-1}}$ represents the χ^2-fitted curve, where $g = 1.14147 \pm (0.48\%)$ and $h = 2.38455 \pm (0.25\%)$ are fitting parameters (see Appendix. I), and $n_2^0 = 0.945742$ denotes the value of $n_2^{\min}$ in the complete depolarizing scenario. This signifies another distinct advantage of the quantum time flip strategy over the switched one. In particular, in the presence of high noises, flip always outperforms switch when $\pm 0.945742 \leq n_2 \leq \pm 1$, and when noise starts to decrease, this range gets increased with the noise parameter, making flip highly favorable in low noise regimes. Moreover, since the FI with switch method does not depend on n_2, for $\pm n_2^{\min}(q) \leq n_2 \leq \pm 1$, the FI obtained with flip is strictly higher than the highest possible FI with the switch protocol at various levels of noise q. Because n_2 can be controlled, we can obtain higher average FI values with time-flip for any value of q compared to the switch-based scheme.

5.2 Advantage in multiparty noisy case

Going beyond the single qubit scenario, we examine how multiparty scenario gets affected by noise. While the optimal strategy for the multi-qubit NITDM falls outside the scope of this study, we provide a brief overview of its key aspects. Note that while $\mathcal{F}_\theta^q$ is independent of the input probe for $N = 1$, this dependence emerges when $N \geq 2$. By assuming $\theta = \pi/4, n_2 = 1$, we initialize the probe state to be $|+i\rangle^{\otimes N}$ and evaluate $\mathcal{F}_\theta^q$ (given by Eq. (19)) for different values of noise content q. Let us compare NITDM (by fixing $\theta = \pi/4$) with the regular strategy when all the N-qubit state is initialized along $\hat{s} = \hat{y}$, which is the optimal product probes in the noiseless case. In this case, $\mathcal{Q}_\theta^{q,R} = Nq^2$ while the behavior of $\mathcal{F}_{\theta=\frac{\pi}{4}}^q$ is depicted in Fig. 4. Due to N^2 term in Eq. (19), $\mathcal{F}_\theta$ increases, while the exponential term is responsible for loss of Fisher information with N as shown in Fig. 4. Specifically, depending on the system configuration and noise content, $\mathcal{F}_\theta^q$ reaches maximum at a certain value of N. However, FI per qubit, $\mathcal{F}_\theta^q/N$, is maximized, where the first slope-change of the envelope of $\mathcal{F}_\theta^q$ occurs. According to this, for a given $\theta, \vec{s}$, and $\hat{n}$, we define the set $\{\tilde{\mathcal{F}}_\theta^q, \tilde{N}\}$ where $\tilde{\mathcal{F}}_\theta^q$ is the value of FI of $\tilde{N}$-qubit NITDM protocol corresponding to the maximized value of $\mathcal{F}_\theta^q/N$ at $\tilde{N}$. Given a high value of $N \gg \tilde{N}$, we perform $\tilde{N}$-qubit NITDM and repeat the procedure for $\lfloor N/\tilde{N} \rfloor$ times to

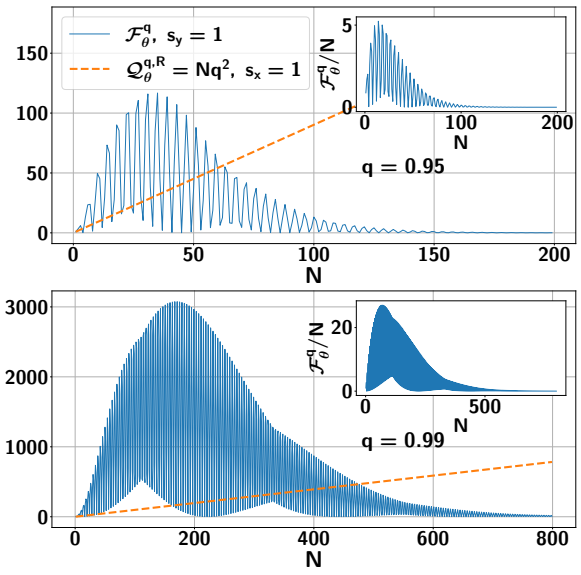

Figure 4: **Fisher information (ordinate) with respect to number or qubits N (abscissa) for different values of noise parameter q.** Fisher information for time-flip, $\mathcal{F}_\theta^q$ (blue lines) shows oscillatory behavior and after reaching the maximum, it decays. Here, we fix the parameter values $\theta = \pi/4$ and $\hat{n}_2 = 1$. Contrastingly, QFI for the regular case, $\mathcal{Q}_\theta^{(q,R)}$ (orange dotted lines) increases linearly. Inset: Fisher information per qubit $\mathcal{F}_\theta^q/N$ (ordinate, blue lines) against number or qubits N (abscissa). All axes are dimensionless.

get an advantage over regular additive strategy with N-qubit product probe states. Note that if $N < \tilde{N}$, we have to perform N-qubit NITDM to get higher value of FI over the regular strategy.

To demonstrate the above claim, we find that maximum $\mathcal{F}_{\theta=\frac{\pi}{4}}^{q=0.99} = 3074.72$ and $\mathcal{F}_{\theta=\frac{\pi}{4}}^{q=0.95} = 116.77$, which are achieved at $N = 170$ and $N = 34$ in the NITDM scheme respectively. On the other hand, the maximum values of $\mathcal{F}_\theta^q/N$ are 27.46 and 5.51, achieved at $\{\tilde{\mathcal{F}}_\theta^q, \tilde{N}\} = \{1894.5, 69\}$ and $\{49.68, 9\}$ respectively. Let us elaborate on the benefit of NITDM by discussing two cases – **Case 1.** For a fixed number of qubits, say $N = 1000$ and $q = 0.95$, employing the NITDM with $\tilde{N} = 9$ qubits for 111 repetitions yields an effective Fisher information of $111\tilde{\mathcal{Q}}_\theta^{q=0.95} = 111 \times 49.68 = 5514.48$, surpassing the QFI achievable with a standard 1000-qubit system under the same noise conditions, $\mathcal{Q}_\theta^{(q=0.95,R)} = Nq^2 = 1000 \times 0.95^2 = 902.5$. **Case 2.** For $q = 0.99$, we can perform NITDM with $\tilde{N} = 69$ repeatedly for 14 times to get the best possible result $14\mathcal{F}_\theta^{q=0.99} = 26523 \gg \mathcal{Q}_\theta^{(q=0.99,R)} = 980.1$.

This shows a notable performance gain by NITDM compared to the regular strategy in a low-noise regime, although it is for a specific parameter θ. However, our study provides evidence that this advantage persists when averaged over the parameter space of θ as previously defined, which we discuss in Appendix. J.

6 Conclusion

We incorporated the concept of an indefinite time directed encoding process in the framework of quantum metrology to enhance the precision limit of parameter estimation beyond the standard quantum limit (SQL). To accomplish the higher precision than the SQL, referred to as the Heisenberg limit (HL), the conventional parameter estimation strategy relies on an entangled input probe, along with unitary encoding and multi qubit measurement on the encoded state for decoding. In contrast, we demonstrated that achieving the HL is possible by employing unentangled input probes with discrete energy levels provided indefinite time directed encoding (ITDM) is performed. Specifically, we proved analytically that the symmetric product probe state is sufficient to reach Heisenberg scaling among all possible multiparty input states although entangled input probe states can increase the precision of parameter estimation compared to the separable probe ones. Further, in certain cases involving product probes, the bidirectionally encoded state remains a product state, indicating that entanglement is not generated throughout the protocol even after the global time-flip operation and, therefore, is unnecessary. Fixing the measurement on the control qubit, we illustrated that in the limit of small parameter values, both the input probe and quantum Fisher information (QFI) which provides a lower bound on the precision are independent of the parameters involved. We also examined the impact of noise on a single-qubit scenario and assessed its advantages in certain cases over the "switched" (based on indefinite causal order) and conventional metrology protocols. Additionally, we discussed strategies for achieving higher Fisher information in noisy multiqubit scenarios, in comparison to the regular scheme.

It is important to note that a quantum time flip operation cannot be realized within the stan-

dard quantum circuit model since no such circuit can transform an unknown unitary gate to its transpose [71, 81]. However, it was demonstrated [76, 77] that the implementation of time-flip operations on photonic platforms can be achieved by leveraging device-dependent symmetries inherent to their experimental apparatus. Specifically, the target qubit, subject to the time flip operation, is encoded in the polarization degree of freedom of light, while its spatial modes serve the purpose of the control qubit. Forward and backward temporal directions are achieved by transmitting photons through appropriately configured waveplates, following the proper Stokes-parameter convention [82]. Consequently, given N queries of the black boxes, U_θ (where θ is the parameter to be estimated), the ITDM scheme should be within experimental reach. Further, this can be used to validate the Heisenberg limit of root mean square error utilizing a bidirectional encoding process without entangled probes, which are typically fragile against noise.

Our results pave the way for the metrological utilization of quantum time flips in the discrete-variable regime. Going forward, it will be fascinating to explore the behavior of QFI in the presence of noise using other estimation procedures such as parallel, adaptive, and indefinite-causal-order approaches [31, 35, 38] in conjunction with indefinite time direction. To determine whether time flip or more generic indefinite time directed [71] techniques outperform typical quantum comb or indefinite-causal-order systems, their hierarchies must be rigorously classified and evaluated possibly using semidefinite optimization techniques [31, 35, 38, 83, 84].

Acknowledgements

We acknowledge the use of QIClib – a modern C++ library for general purpose quantum information processing and quantum computing and cluster computing facility at Harish-Chandra Research Institute. PH acknowledges "INFOSYS scholarship for senior students".

References

[1] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. "Quantum Metrology". Phys. Rev. Lett. **96**, 010401 (2006).

[2] Matteo G. A. Paris. "Quantum estimation for quanutm technology". Int. J. Quantum Inform. **07**, 125–137 (2009).

[3] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. "Advances in quantum metrology". Nat. Photonics **5**, 222–229 (2011).

[4] Géza Tóth and Iagoba Apellaniz. "Quantum metrology from a quantum information science perspective". J. Phys. A: Math. Theor. **47**, 424006 (2014).

[5] C. L. Degen, F. Reinhard, and P. Cappellaro. "Quantum sensing". Rev. Mod. Phys. **89**, 035002 (2017).

[6] Luca Pezzè, Augusto Smerzi, Markus K. Oberthaler, Roman Schmied, and Philipp Treutlein. "Quantum metrology with nonclassical states of atomic ensembles". Rev. Mod. Phys. **90**, 035005 (2018).

[7] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd. "Advances in photonic quantum sensing". Nat. Photonics **12**, 724–733 (2018).

[8] Jonathan A. Jones, Steven D. Karlen, Joseph Fitzsimons, Arzhang Ardavan, Simon C. Benjamin, G. Andrew D. Briggs, and John J. L. Morton. "Magnetic Field Sensing Beyond the Standard Quantum Limit Using 10-Spin NOON States". Science **324**, 1166–1168 (2009).

[9] W. Wasilewski, K. Jensen, H. Krauter, J. J. Renema, M. V. Balabas, and E. S. Polzik. "Quantum Noise Limited and Entanglement-Assisted Magnetometry". Phys. Rev. Lett. **104**, 133601 (2010).

[10] Carlos A. Pérez-Delgado, Mark E. Pearce, and Pieter Kok. "Fundamental Limits of Classical and Quantum Imaging". Phys. Rev. Lett. **109**, 123601 (2012).

[11] Marco Genovese. "Real applications of quantum imaging". J. Opt. **18**, 073002 (2016).

[12] Paul-Antoine Moreau, Ermes Toninelli, Thomas Gregory, and Miles J. Padgett. "Imaging with quantum states of light". Nat. Rev. Phys. **1**, 367–380 (2019).

[13] J. Appel, P. J. Windpassinger, D. Oblak, U. B. Hoff, N. Kjærgaard, and E. S. Polzik. "Mesoscopic atomic entanglement for precision measurements beyond the standard

quantum limit". Proc. Natl. Acad. Sci. U.S.A. **106**, 10960–10965 (2009).

[14] Hidetoshi Katori. "Optical lattice clocks and quantum metrology". Nat. Photonics **5**, 203–210 (2011).

[15] Andrew D. Ludlow, Martin M. Boyd, Jun Ye, E. Peik, and P. O. Schmidt. "Optical atomic clocks". Rev. Mod. Phys. **87**, 637–701 (2015).

[16] B. C. Nichol, R. Srinivas, D. P. Nadlinger, P. Drmota, D. Main, G. Araneda, C. J. Ballance, and D. M. Lucas. "An elementary quantum network of entangled optical atomic clocks". Nature **609**, 689–694 (2022).

[17] Michael A. Taylor and Warwick P. Bowen. "Quantum metrology and its application in biology". Phys. Rep. **615**, 1–59 (2016).

[18] N. P. Mauranyapin, L. S. Madsen, M. A. Taylor, M. Waleed, and W. P. Bowen. "Evanescent single-molecule biosensing with quantum-limited precision". Nat. Photonics **11**, 477–481 (2017).

[19] M. Malnou, D. A. Palken, B. M. Brubaker, Leila R. Vale, Gene C. Hilton, and K. W. Lehnert. "Squeezed Vacuum Used to Accelerate the Search for a Weak Classical Signal". Phys. Rev. X **9**, 021023 (2019).

[20] Steven D. Bass and Michael Doser. "Quantum sensing for particle physics". Nat. Rev. Phys. **6**, 329–339 (2024).

[21] Carlton M. Caves. "Quantum-mechanical noise in an interferometer". Phys. Rev. D **23**, 1693–1708 (1981).

[22] Roman Schnabel, Nergis Mavalvala, David E. McClelland, and Ping K. Lam. "Quantum metrology for gravitational wave astronomy". Nat. Commun. **1**, 1–10 (2010).

[23] "A gravitational wave observatory operating beyond the quantum shot-noise limit" (2011). [Online; accessed 1. Feb. 2025].

[24] Harald Cramér. "Mathematical methods of statistics (pms-9)". Princeton University Press. (1946).

[25] Samuel L. Braunstein and Carlton M. Caves. "Statistical distance and the geometry of quantum states". Phys. Rev. Lett. **72**, 3439–3443 (1994).

[26] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen. "Spin squeezing and reduced quantum noise in spectroscopy". Phys. Rev. A **46**, R6797–R6800 (1992).

[27] D. Leibfried, M. D. Barrett, T. Schaetz, J. Britton, J. Chiaverini, W. M. Itano, J. D. Jost, C. Langer, and D. J. Wineland. "Toward Heisenberg-Limited Spectroscopy with Multiparticle Entangled States". Science **304**, 1476–1478 (2004).

[28] Luca Pezzé and Augusto Smerzi. "Entanglement, Nonlinear Dynamics, and the Heisenberg Limit". Phys. Rev. Lett. **102**, 100401 (2009).

[29] R. Augusiak, J. Kołodyński, A. Streltsov, M. N. Bera, A. Acín, and M. Lewenstein. "Asymptotic role of entanglement in quantum metrology". Phys. Rev. A **94**, 012339 (2016).

[30] Xiaobin Zhao, Yuxiang Yang, and Giulio Chiribella. "Quantum Metrology with Indefinite Causal Order". Phys. Rev. Lett. **124**, 190503 (2020).

[31] Qiushi Liu, Zihao Hu, Haidong Yuan, and Yuxiang Yang. "Optimal Strategies of Quantum Metrology with a Strict Hierarchy". Phys. Rev. Lett. **130**, 070803 (2023).

[32] François Chapeau-Blondeau. "Noisy quantum metrology with the assistance of indefinite causal order". Physical Review A **103**, 032615 (2021).

[33] Michael Frey. "Indefinite causal order aids quantum depolarizing channel identification". Quantum Inf. Process. **18**, 1–20 (2019).

[34] Masashi Ban. "Quantum Fisher information of phase estimation in the presence of indefinite causal order". Phys. Lett. A **468**, 128749 (2023).

[35] Raphaël Mothe, Cyril Branciard, and Alastair A. Abbott. "Reassessing the advantage of indefinite causal orders for quantum metrology". Phys. Rev. A **109**, 062435 (2024).

[36] Min An, Shihao Ru, Yunlong Wang, Yu Yang, Feiran Wang, Pei Zhang, and Fuli Li. "Noisy quantum parameter estimation with indefinite causal order". Phys. Rev. A **109**, 012603 (2024).

[37] Peng Yin, Xiaobin Zhao, Yuxiang Yang, Yu Guo, Wen-Hao Zhang, Gong-Chu Li, Yong-Jian Han, Bi-Heng Liu, Jin-Shi Xu, Giulio Chiribella, Geng Chen, Chuan-Feng

Li, and Guang-Can Guo. "Experimental super-Heisenberg quantum metrology with indefinite gate order". Nat. Phys. **19**, 1122–1127 (2023).

[38] Qiushi Liu, Zihao Hu, Haidong Yuan, and Yuxiang Yang. "Fully-Optimized Quantum Metrology: Framework, Tools, and Applications". Adv. Quantum Technol. **7**, 2400094 (2024).

[39] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner. "Quantum correlations with no causal order". Nat. Commun. **3**, 1–8 (2012).

[40] Mateus Araújo, Cyril Branciard, Fabio Costa, Adrien Feix, Christina Giarmatzi, and Časlav Brukner. "Witnessing causal nonseparability". New J. Phys. **17**, 102001 (2015).

[41] Lorenzo M. Procopio, Amir Moqanaki, Mateus Araújo, Fabio Costa, Irati Alonso Calafell, Emma G. Dowd, Deny R. Hamel, Lee A. Rozema, Časlav Brukner, and Philip Walther. "Experimental superposition of orders of quantum gates". Nat. Commun. **6**, 1–6 (2015).

[42] Giulia Rubino, Lee A. Rozema, Adrien Feix, Mateus Araújo, Jonas M. Zeuner, Lorenzo M. Procopio, Časlav Brukner, and Philip Walther. "Experimental verification of an indefinite causal order". Sci. Adv.**3** (2017).

[43] 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).

[44] Teodor Strömberg, Peter Schiansky, Robert W. Peterson, Marco Túlio Quintino, and Philip Walther. "Demonstration of a Quantum Switch in a Sagnac Configuration". Phys. Rev. Lett. **131**, 060803 (2023).

[45] Lee A. Rozema, Teodor Strömberg, Huan Cao, Yu Guo, Bi-Heng Liu, and Philip Walther. "Experimental aspects of indefinite causal order in quantum mechanics". Nat. Rev. Phys. **6**, 483–499 (2024).

[46] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, and Benoit Valiron. "Quantum computations without definite causal structure". Physical Review A **88**, 022318 (2013).

[47] Julian Wechs, Hippolyte Dourdent, Alastair A. Abbott, and Cyril Branciard. "Quantum Circuits with Classical Versus Quantum Control of Causal Order". PRX Quantum **2**, 030335 (2021).

[48] Giulio Chiribella. "Perfect discrimination of no-signalling channels via quantum superposition of causal structures". Phys. Rev. A **86**, 040301 (2012).

[49] Jessica Bavaresco, Mio Murao, and Marco Túlio Quintino. "Strict Hierarchy between Parallel, Sequential, and Indefinite-Causal-Order Strategies for Channel Discrimination". Phys. Rev. Lett. **127**, 200504 (2021).

[50] Jessica Bavaresco, Mio Murao, and Marco Túlio Quintino. "Unitary channel discrimination beyond group structures: Advantages of sequential and indefinite-causal-order strategies". J. Math. Phys. **63**, 042203 (2022).

[51] Spiros Kechrimparis, James Moran, Athena Karsa, Changhyoup Lee, and Hyukjoon Kwon. "Enhancing quantum state discrimination with indefinite causal order". New J. Phys. (2024).

[52] Daniel Ebler, Sina Salek, and Giulio Chiribella. "Enhanced communication with the assistance of indefinite causal order". Physical Review Letters **120**, 120502 (2018).

[53] Kejin Wei, Nora Tischler, Si-Ran Zhao, Yu-Huai Li, Juan Miguel Arrazola, Yang Liu, Weijun Zhang, Hao Li, Lixing You, Zhen Wang, Yu-Ao Chen, Barry C. Sanders, Qiang Zhang, Geoff J. Pryde, Feihu Xu, and Jian-Wei Pan. "Experimental Quantum Switching for Exponentially Superior Quantum Communication Complexity". Phys. Rev. Lett. **122**, 120504 (2019).

[54] Sina Salek, Daniel Ebler, and Giulio Chiribella. "Quantum communication in a superposition of causal orders" (2018). arXiv:1809.06655.

[55] Zhen Wu, James Fullwood, Zhihao Ma, Siqi Zhou, Qi Zhao, and Giulio Chiribella. "General Communication Enhancement via the Quantum Switch" (2024). arXiv:2407.02726.

[56] Hlér Kristjánsson, Giulio Chiribella, Sina Salek, Daniel Ebler, and Matthew Wilson. "Resource theories of communication". New J. Phys. **22**, 073014 (2020).

[57] Matt Wilson and Giulio Chiribella. "A Dia-

grammatic Approach to Information Transmission in Generalised Switches" (2020). arXiv:2003.08224.

[58] Nicolas Loizeau and Alexei Grinbaum. "Channel capacity enhancement with indefinite causal order". Phys. Rev. A **101**, 012340 (2020).

[59] Yu Guo, Xiao-Min Hu, Zhi-Bo Hou, Huan Cao, Jin-Ming Cui, Bi-Heng Liu, Yun-Feng Huang, Chuan-Feng Li, Guang-Can Guo, and Giulio Chiribella. "Experimental Transmission of Quantum Information Using a Superposition of Causal Orders". Phys. Rev. Lett. **124**, 030502 (2020).

[60] K. Goswami, Y. Cao, G. A. Paz-Silva, J. Romero, and A. G. White. "Increasing communication capacity via superposition of order". Phys. Rev. Res. **2**, 033292 (2020).

[61] Giulio Chiribella, Matt Wilson, and H. F. Chau. "Quantum and Classical Data Transmission through Completely Depolarizing Channels in a Superposition of Cyclic Orders". Phys. Rev. Lett. **127**, 190502 (2021).

[62] Sk Sazim, Michal Sedlak, Kratveer Singh, and Arun Kumar Pati. "Classical communication with indefinite causal order for N completely depolarizing channels". Phys. Rev. A **103**, 062610 (2021).

[63] David Felce and Vlatko Vedral. "Quantum Refrigeration with Indefinite Causal Order". Phys. Rev. Lett. **125**, 070603 (2020).

[64] Tamal Guha, Mir Alimuddin, and Preeti Parashar. "Thermodynamic advancement in the causally inseparable occurrence of thermal maps". Phys. Rev. A **102**, 032215 (2020).

[65] Kyrylo Simonov, Saptarshi Roy, Tamal Guha, Zoltán Zimborás, and Giulio Chiribella. "Activation of thermal states by coherently controlled thermalization processes". New Journal of Physics (2025).

[66] Kyrylo Simonov, Gianluca Francica, Giacomo Guarnieri, and Mauro Paternostro. "Work extraction from coherently activated maps via quantum switch". Phys. Rev. A **105**, 032217 (2022).

[67] Gianluca Francica. "Causal games of work extraction with indefinite causal order". Phys. Rev. A **106**, 042214 (2022).

[68] Xinfang Nie, Xuanran Zhu, Keyi Huang, Kai Tang, Xinyue Long, Zidong Lin, Yu Tian, Chudan Qiu, Cheng Xi, Xiaodong Yang, Jun Li, Ying Dong, Tao Xin, and Dawei Lu. "Experimental Realization of a Quantum Refrigerator Driven by Indefinite Causal Orders". Phys. Rev. Lett. **129**, 100603 (2022).

[69] Gaoyan Zhu, Yuanbo Chen, Yoshihiko Hasegawa, and Peng Xue. "Charging Quantum Batteries via Indefinite Causal Order: Theory and Experiment". Phys. Rev. Lett. **131**, 240401 (2023).

[70] Alessandro Bisio and Paolo Perinotti. "Theoretical framework for higher-order quantum theory". Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **475**, 20180706 (2019).

[71] Giulio Chiribella and Zixuan Liu. "Quantum operations with indefinite time direction". Commun. Phys. **5**, 1–8 (2022).

[72] Zixuan Liu, Ming Yang, and Giulio Chiribella. "Quantum communication through devices with indefinite input-output direction". New J. Phys. **25**, 043017 (2023).

[73] Göktuğ Karpat and Barış Çakmak. "Memory in quantum processes with indefinite time direction and causal order". Phys. Rev. A **110**, 012446 (2024).

[74] Giulia Rubino, Gonzalo Manzano, and Časlav Brukner. "Quantum superposition of thermodynamic evolutions with opposing time's arrows". Commun. Phys. **4**, 1–10 (2021).

[75] Giulia Rubino, Gonzalo Manzano, Lee A. Rozema, Philip Walther, Juan M. R. Parrondo, and Časlav Brukner. "Inferring work by quantum superposing forward and time-reversal evolutions". Phys. Rev. Res. **4**, 013208 (2022).

[76] Teodor Strömberg, Peter Schiansky, Marco Túlio Quintino, Michael Antesberger, Lee A. Rozema, Iris Agresti, Časlav Brukner, and Philip Walther. "Experimental superposition of a quantum evolution with its time reverse". Phys. Rev. Res. **6**, 023071 (2024).

[77] Yu Guo, Zixuan Liu, Hao Tang, Xiao-Min Hu, Bi-Heng Liu, Yun-Feng Huang, Chuan-Feng Li, Guang-Can Guo, and Giulio Chiribella. "Experimental Demonstration of Input-Output Indefiniteness in a Single Quantum Device". Phys. Rev. Lett. **132**, 160201 (2024).

[78] Binke Xia, Jingzheng Huang, Hongjing Li, Zhongyuan Luo, and Guihua Zeng. "Nanoradian-scale precision in light rotation measurement via indefinite quantum dynamics". Sci. Adv.**10** (2024).

[79] Shelby Kimmel, Guang Hao Low, and Theodore J. Yoder. "Robust calibration of a universal single-qubit gate set via robust phase estimation". Phys. Rev. A **92**, 062315 (2015).

[80] Michael A. Nielsen and Isaac L. Chuang. "Quantum computation and quantum information". Cambridge University Press. (2012).

[81] Marco Túlio Quintino, Qingxiuxiong Dong, Atsushi Shimbo, Akihito Soeda, and Mio Murao. "Probabilistic exact universal quantum circuits for transforming unitary operations". Phys. Rev. A **100**, 062339 (2019).

[82] Nicholas J. Frigo, Frank Bucholtz, Geoffrey A. Cranch, and Joseph M. Singley. "On the choice of conventions in polarization evolution calculations and representations". Journal of Lightwave Technology **40**, 179–190 (2022).

[83] Stanisław Kurdziałek, Wojciech Górecki, Francesco Albarelli, and Rafał Demkowicz-Dobrzański. "Using adaptiveness and causal superpositions against noise in quantum metrology". Phys. Rev. Lett. **131**, 090801 (2023).

[84] Stanisław Kurdziałek, Piotr Dulian, Joanna Majsak, Sagnik Chakraborty, and Rafał Demkowicz-Dobrzański. "Quantum metrology using quantum combs and tensor network formalism". New Journal of Physics **27**, 013019 (2025).

[85] Karl Pearson. "X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling". The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **50**, 157–175 (1900).

[86] William G. Cochran. "The χ^2 test of goodness of fit". The Annals of Mathematical Statistics **23**, 315–345 (1952).

A Calculation of QFI for a arbitrary fully separable pure state

By using the form of pure product probe states, given in Eq. (2) of the main text, we obtain

$$|\dot{\Phi}\rangle = \frac{\mathrm{d}\,|\Phi\rangle}{\mathrm{d}\theta} = \sqrt{p_c}\,|0\rangle \otimes \sum_{i=1}^{N} \left(|\dot{\chi}_i\rangle \bigotimes_{\substack{j=1 \\ j\neq i}}^{N} |\chi_j\rangle \right) + e^{i\theta_c}\sqrt{1-p_c}\,|1\rangle \otimes \sum_{i=1}^{N} \left(|\dot{\omega}_i\rangle \bigotimes_{\substack{j=1 \\ j\neq i}}^{N} |\omega_j\rangle \right).$$

Therefore, we can write the inner product of $|\dot{\Phi}\rangle$ with itself as

$$\langle\dot{\Phi}|\dot{\Phi}\rangle = p_c \left(\sum_{k=1}^{N} \langle\dot{\chi}_k|\dot{\chi}_k\rangle + \sum_{\substack{j,k=1 \\ j\neq k}}^{N} \langle\dot{\chi}_k|\chi_k\rangle\langle\chi_j|\dot{\chi}_j\rangle \right) + (1-p_c) \left(\sum_{k=1}^{N} \langle\dot{\omega}_k|\dot{\omega}_k\rangle + \sum_{\substack{j,k=1 \\ j\neq k}}^{N} \langle\dot{\omega}_k|\omega_k\rangle\langle\omega_j|\dot{\omega}_j\rangle \right). \tag{35}$$

On the other hand, we have

$$\langle\dot{\Phi}|\Phi\rangle = p_c \sum_{k=1}^{N} \langle\dot{\chi}_k|\chi_k\rangle + (1-p_c) \sum_{k=1}^{N} \langle\dot{\omega}_k|\omega_k\rangle. \tag{36}$$

Defining

$$\begin{aligned} \langle\dot{\chi}_k|\chi_k\rangle &= \alpha_k, \quad \langle\dot{\chi}_k|\dot{\chi}_k\rangle = |\dot{\chi}_k|, \\ \langle\dot{\omega}_k|\omega_k\rangle &= \beta_k, \quad \langle\dot{\omega}_k|\dot{\omega}_k\rangle = |\dot{\omega}_k|, \end{aligned} \tag{37}$$

we can rewrite Eqs. (35) and (36) as

$$\langle\dot{\Phi}|\dot{\Phi}\rangle = p_c\left(\sum_{k=1}^{N}|\dot{\chi}_k| + \sum_{\substack{j,k=1\\ j\neq k}}^{N}\alpha_k\alpha_j^*\right) + (1-p_c)\left(\sum_{k=1}^{N}|\dot{\omega}_k| + \sum_{\substack{j,k=1\\ j\neq k}}^{N}\beta_k\beta_j^*\right),$$

and

$$\langle\dot{\Phi}|\Phi\rangle = p_c\sum_{k=1}^{N}\alpha_k + (1-p_c)\sum_{k=1}^{N}\beta_k, \tag{38}$$

respectively. Finally, we calculate the QFI, corresponding to an arbitrary product probe state as shown in Eq. (4) of main text.

B Proof of Theorem 1

$$\begin{aligned}
&\mathcal{Q}_{\text{avg}} - \mathcal{Q}^A \\
&= 4\left[p_c\left(\sum_{k=1}^{N}|\dot{\chi}_k| + (N-1)\sum_{k=1}^{N}\alpha_k\alpha_k^*\right) + (1-p_c)\left(\sum_{k=1}^{N}|\dot{\omega}_k| + (N-1)\sum_{k}\beta_k\beta_k^*\right) - N\sum_{k=1}^{N}\left|p_c\alpha_k + (1-p_c)\beta_k\right|^2\right] \\
&\quad -4\left[p_c\left(\sum_{k=1}^{N}|\dot{\chi}_k| + \sum_{k=1}^{N}\sum_{j\neq k}^{N}\alpha_k\alpha_j^*\right) + (1-p_c)\left(\sum_{k=1}^{N}|\dot{\omega}_k| + \sum_{k=1}^{N}\sum_{j\neq k}^{N}\beta_k\beta_j^*\right) - \left|p_c\sum_{k=1}^{N}\alpha_k + (1-p_c)\sum_{k=1}^{N}\beta_k\right|^2\right] \\
&= 4\left[p_c\left((N-1)\sum_{k=1}^{N}\alpha_k\alpha_k^* - \sum_{k=1}^{N}\sum_{j\neq k}^{N}\alpha_k\alpha_j^*\right) + (1-p_c)\left((N-1)\sum_{k=1}^{N}\beta_k\beta_k^* - \sum_{k=1}^{N}\sum_{j\neq k}^{N}\beta_k\beta_j^*\right)\right. \\
&\quad \left.+\left|p_c\sum_{k=1}^{N}\alpha_k + (1-p_c)\sum_{k=1}^{N}\beta_k\right|^2 - N\sum_{k=1}^{N}\left|p_c\alpha_k + (1-p_c)\beta_k\right|^2\right] \\
&= 2p_c\sum_{k=1}^{N}\sum_{j=1}^{N}\left(\alpha_k\alpha_k^* + \alpha_j\alpha_j^* - \alpha_k\alpha_j^* - \alpha_j\alpha_k^*\right) + 2(1-p_c)\sum_{k=1}^{N}\sum_{j=1}^{N}\left(\beta_k\beta_k^* + \beta_j\beta_j^* - \beta_k\beta_j^* - \beta_j\beta_k^*\right) \\
&\quad +4\sum_{k=1}^{N}\sum_{j=1}^{N}\left(p_c^2\alpha_k\alpha_j^* + (1-p_c)^2\beta_k\beta_j^* + p_c(1-p_c)\beta_k\alpha_j^* + p_c(1-p_c)\alpha_k\beta_j^*\right) \\
&\quad -4N\sum_{k=1}^{N}\left(p_c^2\alpha_k\alpha_k^* + (1-p_c)^2\beta_k\beta_k^* + p_c(1-p_c)\alpha_k\beta_k^* + p_c(1-p_c)\beta_k\alpha_k^*\right) \\
&= 4Np_c(1-p_c)\sum_{k=1}^{N}\alpha_k\alpha_k^* + 4Np_c(1-p_c)\sum_{k=1}^{N}\beta_k\beta_k^* - 4p_c(1-p_c)\sum_{k=1}^{N}\sum_{j=1}^{N}\alpha_k\alpha_j^* - 4p_c(1-p_c)\sum_{k=1}^{N}\sum_{j=1}^{N}\beta_k\beta_j^* \\
&\quad +4p_c(1-p_c)\sum_{k=1}^{N}\sum_{j=1}^{N}\alpha_k\beta_j^* + 4p_c(1-p_c)\sum_{k=1}^{N}\sum_{j=1}^{N}\beta_k\alpha_j^* - 4Np_c(1-p_c)\sum_{k=1}^{N}\alpha_k\beta_k^* - 4Np_c(1-p_c)\sum_{k=1}^{N}\beta_k\alpha_k^* \\
&= 4Np_c(1-p_c)\left(\sum_{k=1}^{N}\alpha_k\alpha_k^* + \sum_{k=1}^{N}\beta_k\beta_k^* - \sum_{k=1}^{N}\alpha_k\beta_k^* - \sum_{k=1}^{N}\beta_k\alpha_k^*\right) \\
&\quad +4p_c(1-p_c)\left(\sum_{k=1}^{N}\sum_{j=1}^{N}\alpha_k\beta_j^* + \sum_{k=1}^{N}\sum_{j=1}^{N}\beta_k\beta_j^* - \sum_{k=1}^{N}\sum_{j=1}^{N}\alpha_k\alpha_j^* - \sum_{k=1}^{N}\sum_{j=1}^{N}\beta_k\beta_j^*\right) \\
&= 4Np_c(1-p_c)\sum_{k=1}^{N}|\alpha_k - \beta_k|^2 - 4p_c(1-p_c)\left|\sum_{k=1}^{N}\alpha_k - \beta_k\right|^2 \geq 0
\end{aligned} \tag{39}$$

where the last inequality is due to Cauchy–Schwarz inequality.

C Proof of $A + B \leq 1$ in Theorem 3

The sum of coefficients A and B is calculated as

$$A + B$$
$$= 1 - n_3^2(1-2p_s)^2 + 4\{n_1\cos\theta_s + n_2(2p_c-1)\sin\theta_s\}\left[n_3(1-2p_s)\sqrt{p_s(p_s-1)}\{n_1\cos\theta_s + n_2(2p_c-1)\sin\theta_s\}\right]. \tag{40}$$

To maximize $A + B$ with respect to the parameter p_c we calculate

$$\frac{\partial(A+B)}{\partial p_c} = 8n_2\sin\theta_s\left[n_3(1-2p_s)\sqrt{p_s(1-p_s)} - 2p_s(1-p_s)\{n_1\cos\theta_s + n_2(-1+2p_c)\sin\theta_s\}\right]. \tag{41}$$

Solving $\frac{\partial(A+B)}{\partial p_c} = 0$ we get

$$p_c = \frac{1}{2} - \frac{1}{4n_2}\left\{2n_1\cot\theta_s + \frac{n_3(-1+2p_s)\csc\theta_s}{\sqrt{p_s(1-p_s)}}\right\} = p_c^o, \tag{42}$$

which is the only critical point. Now, the 2nd order derivative is given by

$$\frac{\partial^2(A+B)}{\partial {p_c}^2} = -32n_2^2 p_s(1-p_s)\sin^2\theta_s. \tag{43}$$

From Eq. (43) it is clear that for $-1 \leq n_2 \leq 1$, $0 < p_s < 1$ and $\theta_s \neq 2m\pi$ with $m \in \mathbb{Z}$, i.e., for all valid region of $n_1, n_2, n_3, p_s, \theta_s, p_c$ where $A \neq 0$, we have $\frac{\partial^2(A+B)}{\partial {p_c}^2} < 0$. This implies that $A + B$ is strictly concave in p_c and $A + B|_{p_c=p_c^o} = 1$ is the local maximum. Therefore, the global maximum value of $A + B$ is given by $A + B|_{p_c=p_c^o} = 1$.

D Axis Estimation with ITDM

The 2×2 unitary matrix is given by

$$\begin{aligned} U &= \exp\left\{-i\frac{\theta}{2}(\sin\phi\cos\xi\ \sigma_x + \sin\phi\sin\xi\ \sigma_y + \cos\phi\ \sigma_z)\right\}, \\ &= \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) - i\cos\phi\sin\left(\frac{\theta}{2}\right) & -i\cos\xi\sin\phi\sin\left(\frac{\theta}{2}\right) - \sin\phi\sin\xi\sin\left(\frac{\theta}{2}\right) \\ -i\cos\xi\sin\phi\sin\left(\frac{\theta}{2}\right) + \sin\phi\sin\xi\sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) + i\cos\phi\sin\left(\frac{\theta}{2}\right) \end{bmatrix}, \end{aligned} \tag{44}$$

where θ is the phase and ϕ, ξ are the parameters of the axis of rotation of the unitary matrix. We aim to estimate ϕ. According to corollary 2, we initialize the input state as $\left|\Psi^S\right\rangle = |\psi_c\rangle\,|\psi\rangle^{\otimes N}$ where $|\psi\rangle = \sqrt{p_s}\,|0\rangle + e^{i\theta_s}\sqrt{1-p_s}\,|1\rangle$ with $\theta_s \in [0, 2\pi]$ and $p_s \in [0, 1]$ and $|\psi_c\rangle = \sqrt{p_c}\,|0\rangle + e^{i\theta_c}\sqrt{1-p_c}\,|1\rangle$. In this scenario, we find that

$$
\begin{aligned}
\alpha = \langle\dot{\chi}|\chi\rangle &= i\Big[(1/2-p_s)\sin\phi\sin\theta\sqrt{p_s(1-p_s)}\cos\theta_s\Big(\sin\xi-\cos\theta\sin\xi\ +\cos\phi\cos\xi\sin\theta\Big)\\
&\quad +\sqrt{p_s(1-p_s)}\sin\theta_s\Big(-\cos\xi+\cos\theta\cos\xi+\cos\phi\sin\xi\sin\theta\Big)\Big],\\
\beta = \langle\dot{\omega}|\omega\rangle &= i\Big[(1/2-p_s)\sin\phi\sin\theta+\sqrt{ps(1-ps)}\cos\theta_s\Big(\cos\theta\sin\phi-\sin\xi+\cos\phi\cos\xi\sin\theta\Big)\\
&\quad +\sqrt{ps(1-ps)}\sin\theta_s\Big(-\cos\xi+\cos\theta\cos\xi-\cos\phi\sin\xi\sin\theta\Big)\Big],\\
|\dot{\chi}| = \langle\dot{\chi}|\dot{\chi}\rangle &= \sin^2\frac{\theta}{2},\quad |\dot{\omega}| = \langle\dot{\omega}|\dot{\omega}\rangle = \sin^2\frac{\theta}{2},
\end{aligned}
$$

where the dot represents the derivative with respect to ϕ. This gives a QFI of $\mathcal{Q}_\phi = AN^2 + BN$, where

$$
A = 4p_c(1-p_c)|\alpha-\beta|^2,\quad B = 4\Big[p_c\Big(|\dot{\chi}|-|\alpha|^2\Big)+(1-p_c)\Big(|\dot{\omega}|-|\beta|^2\Big)\Big]. \tag{45}
$$

By optimizing A with respect to the state and unitary parameters we get $\max_{p_c,p_s,\theta_s,\xi,\theta} A = 4$ at $p_c = 1/2,\ p_s = 1/2,\ \theta_s = \pi,\ \xi = \pi/2,\ \theta = \pi$ which in turn gives $B = 0$. Thus, we get a quantum Fisher information of $\mathcal{Q}_\phi = 4N^2$. It is easy to see that for any $AN^2 + BN$ and $CN^2 + DN$ with $A \geq C$ implies $AN^2 + BN \geq CN^2 + DN$ for all positive integer value of N iff $A + B \geq C + D$. Numerically optimizing $A + B$ indeed gives us a value of 4. Therefore the particular parameters employed provide us with the maximum value of QFI, $\mathcal{Q}_\phi = 4N^2$ with optimal pure product states and optimal unitary parameters. In a similar manner, we can also calculate that optimal $\mathcal{Q}_\xi$ scales as N^2.

E Spectrum of $U_\theta^\dagger U_\theta^T$

$U_\theta^\dagger U_\theta^T$ is a unitary matrix, satisfying $U_\theta^\dagger U_\theta^T\,|u(\hat{n},\theta)^\pm\rangle = \exp\left(\pm i\mathrm{f}(n_2,\theta)\right)|u(\hat{n},\theta)^\pm\rangle$. Here, $\mathrm{f}(n_2,\theta)$ is given by

$$
\mathrm{f}(n_2,\theta) = \tan^{-1}\left(\frac{\sqrt{4n_2^2\sin^2\frac{\theta}{2}\left(1-n_2^2\sin^2\frac{\theta}{2}\right)}}{1-2n_2^2\sin^2\frac{\theta}{2}}\right). \tag{46}
$$

The eigenvectors can be expressed as

$$
|u(\hat{n},\theta)^\pm\rangle = \frac{\mathrm{g}^\mp(\hat{n},\theta)\,|0\rangle+|1\rangle}{\sqrt{1+|\mathrm{g}^\mp(\hat{n},\theta)|^2}}, \tag{47}
$$

where

$$
\mathrm{g}^\mp(\hat{n},\theta) = \frac{2n_1n_2\sin^2\frac{\theta}{2}\mp\sqrt{4n_2^2\sin^2\frac{\theta}{2}\left(1-n_2^2\sin^2\frac{\theta}{2}\right)}}{-2n_2n_3\sin^2\frac{\theta}{2}-in_2\sin\theta}. \tag{48}
$$

F FI for small values of θ

For small values of $\theta = \tilde{\theta}$, we consider up to 1st order correction, i.e., by ignoring $\mathcal{O}(\theta^2)$. This approximates Eqs. (24) and (25) in the main text as $r(\boldsymbol{\nu}) \approx 1 - n_2^2\tilde{\theta}^2(1-s_y^2)/2$ and $f(\boldsymbol{\nu}) \approx n_2 s_y\tilde{\theta} -$

$n_2\tilde{\theta}^2(n_3s_x - n_1s_z)/2$. Consequently, we have $\dot{r}(\boldsymbol{\nu}) = n_2^2\tilde{\theta}(1-s_y^2)$ and $\dot{f}(\boldsymbol{\nu}) = n_2s_y - n_2\tilde{\theta}(n_3s_x - n_1s_z)$. Putting these values alongwith $\theta_c \neq 0$ in Eq. (19) in the main text we get

$$\begin{aligned}
\mathcal{F}_{\tilde{\theta}} &\approx N^2\Big(1-\frac{n_2^2(1-s_y^2)\tilde{\theta}^2}{2}\Big)^{2N-2}\Big[-\tilde{\theta}n_2^2(1-s_y^2)\cos\Big(N\{n_2s_y\tilde{\theta}+\mathcal{O}(\theta^2)\}+\theta_c\Big)-\Big(1-\frac{n_2^2(1-s_y^2)\tilde{\theta}^2}{2}\Big)\\
&\quad\times\Big(n_2s_y+\mathcal{O}(\theta)\Big)\sin\Big(N\{n_2s_y\tilde{\theta}+\mathcal{O}(\theta^2)\}+\theta_c\Big)\Big]^2\Big/\Big[1-\Big(1-\frac{n_2^2(1-s_y^2)\tilde{\theta}^2}{2}\Big)^{2N}\\
&\quad\times\cos^2\Big(N\{n_2s_y\tilde{\theta}+\mathcal{O}(\theta^2)\}+\theta_c\Big)\Big],\\
&\approx N^2\Big[-\tilde{\theta}n_2^2(1-s_y^2)\Big(\cos\theta_c-\{Nn_2s_y\tilde{\theta}+\mathcal{O}(\theta^2)\}\sin\theta_c\Big)-\Big(1-\frac{n_2^2(1-s_y^2)\tilde{\theta}^2}{2}\Big)\Big(n_2s_y+\mathcal{O}(\theta)\Big)\\
&\quad\times\Big(\sin\theta_c+N\{n_2s_y\tilde{\theta}+\mathcal{O}(\theta^2)\}\cos\theta_c-\frac{\sin\theta_c}{2}N^2\{n_2s_y\tilde{\theta}+\mathcal{O}(\theta^2)\}^2\Big)\Big]^2\Big/\Big[1-\Big(1-Nn_2^2(1-s_y^2)\tilde{\theta}^2\Big)\\
&\quad\times\Big(\cos^2\theta_c-N\{n_2s_y\tilde{\theta}+\mathcal{O}(\theta^2)\}\sin 2\theta_c-N^2\{n_2s_y\tilde{\theta}+\mathcal{O}(\theta^2)\}^2\cos 2\theta_c\Big)\Big],\\
&\approx \frac{N^2n_2^2s_y^2\sin^2\theta_c}{1-\cos^2\theta_c}=s_y^2n_2^2N^2. \qquad (49)
\end{aligned}$$

On the other hand, with $\theta_c = 0$ we have

$$\begin{aligned}
\mathcal{F}_{\tilde{\theta}} &\approx N^2\Big(1-\frac{n_2^2(1-s_y^2)\tilde{\theta}^2}{2}\Big)^{2N-2}\Big[-\tilde{\theta}n_2^2(1-s_y^2)\cos\Big(Nn_2s_y\tilde{\theta}+\mathcal{O}(\theta^2)\Big)-\Big(1-\frac{n_2^2(1-s_y^2)\tilde{\theta}^2}{2}\Big)\\
&\quad\times\Big(n_2s_y+\mathcal{O}(\theta)\Big)\sin\Big(Nn_2s_y\tilde{\theta}+\mathcal{O}(\theta^2)\Big)\Big]^2\Big/\Big[1-\Big(1-\frac{n_2^2(1-s_y^2)\tilde{\theta}^2}{2}\Big)^{2N}\cos^2\Big(N\{n_2s_y\tilde{\theta}+\mathcal{O}(\theta^2)\}\Big)\Big],\\
&\approx \frac{N^2\left(-\tilde{\theta}n_2^2(1-s_y^2)-Nn_2^2s_y^2\tilde{\theta}-\mathcal{O}(\theta^2)\right)^2}{1-(1-Nn_2^2(1-s_y^2)\tilde{\theta}^2)(1-N^2n_2^2s_y^2\tilde{\theta}^2)},\\
&\approx \frac{N^2n_2^4\tilde{\theta}^2\left((1-s_y^2)+Ns_y^2\right)^2}{Nn_2^2\tilde{\theta}^2((1-s_y^2)+Ns_y^2)}=Nn_2^2\left((1-s_y^2)+Ns_y^2\right)=s_y^2n_2^2N^2+(1-s_y^2)n_2^2N. \qquad (50)
\end{aligned}$$

Comparing Eqs. (49) and (50), it is evident that $\theta_c = 0$ is the optimal scenario. However, $\theta_c \neq 0$ scenario also attains Heisenberg scaling.

G Proof of Theorem 6

A general N-qubit state can be written as

$$|\Psi\rangle = \sum_{i=1}^{2^N}\Lambda_i|\lambda_i\rangle, \qquad (51)$$

where $(i-1)$ is the decimal equivalent number of the ith basis vector $|\lambda_i\rangle$, represented in binary number. For example, the decimal equivalent number of the basis state $|\lambda_2\rangle = |\mathbf{0}\rangle^{\otimes N-1}\otimes|\mathbf{1}\rangle$ is 1, where $|\mathbf{0}\rangle$ and $|\mathbf{1}\rangle$ is just the representation of qubit basis. Decomposing $|\lambda_i\rangle$ as $\bigotimes_{j=1}^N|\psi_j^i\rangle$ where $|\psi_j^i\rangle$ is the ith basis state for jth qubit with $\langle\psi_j^i|\psi_j^k\rangle = \delta_{ik}$ and $\sum_{i=1}^2|\psi_j^i\rangle\langle\psi_j^i| = I_2$ we rewrite

$$|\Psi\rangle = \sum_{i=1}^{2^N}\Lambda_i\bigotimes_{j=1}^N|\psi_j^i\rangle. \qquad (52)$$

The encoded state can be written as

$$|\Phi\rangle = \sum_{i=1}^{2^N} \Lambda_i \left(\sqrt{p_c}\,|0\rangle \bigotimes_{j=1}^{N} \left|\chi_j^i\right\rangle + e^{i\theta_c}\sqrt{1-p_c}\,|1\rangle \bigotimes_{j=1}^{N} \left|\omega_j^i\right\rangle \right), \tag{53}$$

where $\left|\chi_j^i\right\rangle = U_\theta \left|\psi_j^i\right\rangle$ and $\left|\omega_j^i\right\rangle = U_\theta^T \left|\psi_j^i\right\rangle$. To calculate the QFI, we find the derivative of Eq. (53) above with respect to parameter θ which is given by

$$|\dot{\Phi}\rangle = \sum_{i=1}^{2^N} \Lambda_i \left[\sqrt{p_c}|0\rangle \otimes \left(\sum_{m=1}^{N} |\dot{\chi}_m^i\rangle \bigotimes_{j=1 j\neq m}^{N} |\chi_j^i\rangle \right) + e^{i\theta_c}\sqrt{1-p_c}|1\rangle \otimes \left(\sum_{m=1}^{N} |\dot{\omega}_m^i\rangle \bigotimes_{\substack{j=1 \\ j\neq m}}^{N} |\omega_j^i\rangle \right) \right].$$

Therefore, we have

$$\begin{aligned}
\langle\dot{\Phi}|\dot{\Phi}\rangle &= \sum_{i=1}^{2^N}\sum_{k=1}^{2^N} \Lambda_i\Lambda_k^* \Bigg[p_c \sum_{m=1}^{N} \Big\{ \langle\dot{\chi}_m^k|\dot{\chi}_m^i\rangle \prod_{j\neq m} \langle\chi_j^k|\chi_j^i\rangle + \sum_{l\neq m} \langle\dot{\chi}_m^k|\chi_m^i\rangle\langle\chi_l^k|\dot{\chi}_l^i\rangle \prod_{j\neq m,l} \langle\chi_j^k|\chi_j^i\rangle \Big\} \\
&\quad + (1-p_c) \sum_{m=1}^{N} \Big\{ \langle\dot{\omega}_m^k|\dot{\omega}_m^i\rangle \prod_{j\neq m} \langle\omega_j^k|\omega_j^i\rangle + \sum_{l\neq m} \langle\dot{\omega}_m^k|\omega_m^i\rangle\langle\omega_l^k|\dot{\omega}_l^i\rangle \prod_{j\neq m,l} \langle\omega_j^k|\omega_j^i\rangle \Big\} \Bigg], \\
&= \sum_{i=1}^{2^N}\sum_{k=1}^{2^N} \Lambda_i\Lambda_k^* \Bigg[p_c \sum_{m=1}^{N} \Big(\dot{\chi}_m^{ki} \prod_{j\neq m} a_j^{ki} + \sum_{l\neq m} \alpha_m^{ki}\alpha_l^{ik*} \prod_{j\neq m,l} a_j^{ki} \Big) \\
&\quad + (1-p_c) \sum_{m=1}^{N} \Big(\dot{\omega}_m^{ki} \prod_{j\neq m} b_j^{ki} + \sum_{l\neq m} \beta_m^{ki}\beta_l^{ik*} \prod_{j\neq m,l} b_j^{ki} \Big) \Bigg].
\end{aligned}$$

where

$$\begin{aligned}
\langle\dot{\chi}_m^k|\dot{\chi}_m^i\rangle &= \dot{\chi}_m^{ki}, \quad \langle\chi_j^k|\chi_j^i\rangle = a_j^{ki}, \\
\langle\dot{\omega}_m^k|\dot{\omega}_m^i\rangle &= \dot{\omega}_m^{ki}, \quad \langle\omega_j^k|\omega_j^i\rangle = b_j^{ki}, \\
\langle\dot{\chi}_m^k|\chi_m^i\rangle &= \alpha_m^{ki}, \quad \langle\dot{\omega}_j^k|\omega_j^i\rangle = \beta_j^{ki}.
\end{aligned} \tag{54}$$

Now, $|\langle\dot{\Phi}|\Phi\rangle|^2$ being a positive number we have

$$\begin{aligned}
\mathcal{Q}_\theta &\leq 4\sum_{i=1}^{2^N}\sum_{k=1}^{2^N} \Lambda_i\Lambda_k^* \Bigg[p_c \sum_{m=1}^{N} \Big(\dot{\chi}_m^{ki} \prod_{j\neq m} a_j^{ki} + \sum_{l\neq m} \alpha_m^{ki}\alpha_l^{ik*} \prod_{j\neq m,l} a_j^{ki} \Big) \\
&\quad + (1-p_c) \sum_{m=1}^{N} \Big(\dot{\omega}_m^{ki} \prod_{j\neq m} b_j^{ki} + \sum_{l\neq m} \beta_m^{ki}\beta_l^{ik*} \prod_{j\neq m,l} b_j^{ki} \Big) \Bigg], \\
&= \tilde{\mathcal{Q}}_\theta.
\end{aligned}$$

To simplify $\tilde{\mathcal{Q}}_\theta$, we observe that the terms $\prod_{j\neq m} a_j^{ki}$, and $\prod_{j\neq m} b_j^{ki}$ are the products of the inner products of $N-1$ of the qubits. Since U_θ and U_θ^T are both unitaries they preserve inner products, i.e., $a_j^{ki} = b_j^{ki} = \left\langle\psi_j^k\middle|\psi_j^i\right\rangle = \delta_{ki}$. Hence, given a value of m, $\prod_{j\neq m} a_j^{ki} = \prod_{j\neq m} b_j^{ki} = 0$, if the qubits in $|\lambda_i\rangle$ and $\langle\lambda_k|$ differ in more than one position. Similarly, if the qubits in $|\lambda_i\rangle$ and $\langle\lambda_k|$ differ in more than two positions then the terms $\prod_{j\neq m,l} b_j^{ki} = \prod_{j\neq m,l} a_j^{ki} = 0$ for a given pair of values m and l.

We define $\tilde{\mathcal{Q}}_\theta^X$, which involves contribution from N-qubit basis states where qubits differ in X number of positions only. Therefore, we can write $\tilde{\mathcal{Q}}_\theta = \sum_{X=0}^{N} \tilde{\mathcal{Q}}_\theta^X$. However, from the above discussions, we have $\tilde{\mathcal{Q}}_\theta^X = 0\ \forall X \geq 3$ which implies $\tilde{\mathcal{Q}}_\theta = \sum_{X=0}^{2} \tilde{\mathcal{Q}}_\theta^X$. So we have to calculate only $\tilde{\mathcal{Q}}_\theta^0$, $\tilde{\mathcal{Q}}_\theta^1$ and $\tilde{\mathcal{Q}}_\theta^2$.

- **Upper bound of $\tilde{\mathcal{Q}}_\theta^2$:** A pair of N-qubit basis states, $|\lambda_i\rangle$ and $|\lambda_j\rangle$ which differ in exactly two qubits is given by the condition $j = i \oplus 2^k \oplus 2^l$ with $k \neq l$. Now, we can write

$$\tilde{\mathcal{Q}}_\theta^2 = 4\sum_{i=1}^{2^N}\sum_{k=1}^{N}\sum_{\substack{l=1\\l\neq k}}^{N}\Lambda_i\Lambda^*_{i\oplus 2^k\oplus 2^l}\Big[p_c\Big(\alpha_k^{i\oplus 2^k,i}\alpha_l^{i,i\oplus d^l *}+\alpha_l^{i\oplus d^l,i}\alpha_k^{i,i\oplus 2^k *}\Big) +(1-p_c)\Big(\beta_k^{i\oplus 2^k,i}\beta_l^{i,i\oplus d^l *}+\beta_l^{i\oplus d^l,i}\beta_k^{i,i\oplus 2^k *}\Big)\Big].$$

Since for any complex number, z, we have $z + z^* \leq 2|z|$ and $|\sum_i z_i| \leq \sum_i |z_i|$, we can write

$$\begin{aligned}\tilde{\mathcal{Q}}_\theta^2 &\leq 4\sum_{i=1}^{2^N}\sum_{k=1}^{N}\sum_{\substack{l=1\\l\neq k}}^{N}|\Lambda_i\Lambda^*_{i\oplus 2^k\oplus 2^l}|p_c\Big(\Big|\alpha_k^{i\oplus 2^k,i}\alpha_l^{i,i\oplus 2^l *}\Big|+\Big|\alpha_l^{i\oplus 2^l,i}\alpha_k^{i,i\oplus 2^k *}\Big|\Big)\\ &+ 4\sum_{i=1}^{2^N}\sum_{k=1}^{N}\sum_{\substack{l=1\\l\neq k}}^{N}|\Lambda_i\Lambda^*_{i\oplus 2^k\oplus 2^l}|(1-p_c)\Big(\Big|\beta_k^{i\oplus 2^k,i}\beta_l^{i,i\oplus 2^l *}\Big|+\Big|\beta_l^{i\oplus 2^l,i}\beta_k^{i,i\oplus 2^k *}\Big|\Big). \qquad (55)\end{aligned}$$

Now, if $\max|\alpha_k^{i\oplus 2^k,i}| = \mathcal{A}$ and $\max|\beta_k^{i\oplus 2^k,i}| = \mathcal{B}$, where maximization is done over $|\lambda_i\rangle, \langle\lambda_{i\oplus 2^k}|$ and parameters of the unitary by which encoding is performed, then

$$\tilde{\mathcal{Q}}_\theta^2 \leq 8(p_c\mathcal{A}^2+(1-p_c)\mathcal{B}^2)\sum_{k=1}^{N}\sum_{\substack{l=1\\l\neq k}}^{N}\Big|\sum_{i=1}^{2^N}\Lambda_i\Lambda^*_{i\oplus 2^k\oplus 2^l}\Big|. \qquad (56)$$

Defining $\vec{\Lambda} = \{\Lambda_1, \Lambda_2, \ldots, \Lambda_N\}$ as the coefficient vector and $\{k,l\}$-permuted coefficient vector as $\vec{\Lambda}_{kl} = \{\Lambda_{1\oplus 2^k\oplus 2^l}, \Lambda_{2\oplus 2^k\oplus 2^l}, \ldots, \Lambda_{N\oplus 2^k\oplus 2^l}\}$ we can write $\left|\sum_{i=1}^{2^N}\Lambda_i\Lambda^*_{i\oplus 2^k\oplus d^l}\right| = |\vec{\Lambda}_{k,l}\cdot\vec{\Lambda}| \leq \sqrt{|\vec{\Lambda}\cdot\vec{\Lambda}||\vec{\Lambda}_{k,l}\cdot\vec{\Lambda}_{k,l}|} = 1$. This implies

$$\tilde{\mathcal{Q}}_\theta^2 \leq 8(p_c\mathcal{A}^2+(1-p_c)\mathcal{B}^2)\sum_{k=1}^{N}\sum_{\substack{l=1\\l\neq k}}^{N}1 = 8(p_c\mathcal{A}^2+(1-p_c)\mathcal{B}^2)N(N-1). \qquad (57)$$

- **Upper bound of $\tilde{\mathcal{Q}}_\theta^1$:** A pair of N-qubit basis states, $|\lambda_i\rangle$ and $|\lambda_j\rangle$ which differ in exactly one qubit is characterized by the condition $j = i \oplus 2^k$ which simplifies $\tilde{\mathcal{Q}}_\theta^1$ as

$$\begin{aligned}\tilde{\mathcal{Q}}_\theta^1 &= 4\sum_{i=1}^{2^N}\sum_{k=1}^{N}\Lambda_i\Lambda^*_{i\oplus 2^k}\Big[p_c\Big\{\dot{\chi}_k^{i\oplus 2^k,i}+\sum_{l\neq k}\alpha_k^{i\oplus 2^k,i}\alpha_l^{i,i\oplus 2^k *}+\sum_{m\neq k}\alpha_m^{i\oplus 2^k,i}\alpha_k^{i,i\oplus 2^k *}\Big\}\\ &+(1-p_c)\Big\{\dot{\omega}_k^{i\oplus 2^k,i}+\sum_{l\neq k}\beta_k^{i\oplus 2^k,i}\beta_l^{i,i\oplus 2^k *}+\sum_{m\neq k}\beta_m^{i\oplus 2^k,i}\beta_k^{i,i\oplus 2^k *}\Big\}\Big]. \qquad (58)\\ &\leq 4\Big[\sum_{i=1}^{2^N}\sum_{k=1}^{N}|\Lambda_i\Lambda^*_{i\oplus 2^k}|p_c\Big|\dot{\chi}_k^{i\oplus 2^k,i}\Big|+\sum_{i=1}^{2^N}\sum_{k=1}^{N}|\Lambda_i\Lambda^*_{i\oplus 2^k}|p_c\sum_{l\neq k}\Big|\alpha_k^{i\oplus 2^k,i}\alpha_l^{i,i\oplus 2^k *}\Big|\\ &+\sum_{i=1}^{2^N}\sum_{k=1}^{N}|\Lambda_i\Lambda^*_{i\oplus 2^k}|p_c\sum_{m\neq k}\Big|\alpha_m^{i\oplus 2^k,i}\alpha_k^{i,i\oplus 2^k *}\Big|+\sum_{i=1}^{2^N}\sum_{k=1}^{N}|\Lambda_i\Lambda^*_{i\oplus 2^k}|(1-p_c)\Big|\dot{\omega}_k^{i\oplus 2^k,i}\Big|\\ &+\sum_{i=1}^{2^N}\sum_{k=1}^{N}|\Lambda_i\Lambda^*_{i\oplus 2^k}|(1-p_c)\sum_{l\neq k}\Big|\beta_k^{i\oplus 2^k,i}\beta_l^{i,i\oplus 2^k *}\Big|+\sum_{i=1}^{2^N}\sum_{k=1}^{N}|\Lambda_i\Lambda^*_{i\oplus 2^k}|(1-p_c)\sum_{m\neq k}\Big|\beta_m^{i\oplus 2^k,i}\beta_k^{i,i\oplus 2^k *}\Big|\Big]. \qquad (59)\end{aligned}$$

Similar to Eqs. (56) and (57), we can write

$$\tilde{\mathcal{Q}_\theta}^{(1)} \leq 4N(p_c\mathcal{X} + (1-p_c)\mathcal{W}) + 8(p_c\mathcal{A}^2 + (1-p_c)\mathcal{B}^2)N(N-1), \tag{60}$$

where $\mathcal{X}$ and $\mathcal{W}$ are upper bounds of $\left|\dot{\chi}_k^{i\oplus 2^k,i}\right|$ and $\left|\dot{\omega}_k^{i\oplus 2^k,i}\right|$ respectively.

- **Upper bound of $\tilde{\mathcal{Q}}_\theta^0$:** Since the basis states which differ in no position is that basis state itself, the value of $\tilde{\mathcal{Q}}_\theta^0$ is upper bounded by

$$\tilde{\mathcal{Q}_\theta}^{0} \leq \left(A'N^2 + B'N\right)\sum_{i=1}^{2^N}\Lambda_i\Lambda_i^* = \left(A'N^2 + B'N\right), \tag{61}$$

where $A'N^2 + B'N$ is the upper bound of $\langle\dot{\Phi}|\dot{\Phi}\rangle$ for a N-qubit product-probe state as evident from Sec. 2.2 in the main text.

Finally, we have

$$\tilde{\mathcal{Q}}_\theta = \tilde{\mathcal{Q}}_\theta^0 + \tilde{\mathcal{Q}}_\theta^1 + \tilde{\mathcal{Q}}_\theta^2 \leq \gamma N^2 + \zeta N \tag{62}$$

where both γ and ζ are some function of $\mathcal{X}, \mathcal{W}, \mathcal{A}, \mathcal{B}, A', B'$ only which are independent of N. Hence, we prove that no entangled state can beat Heisenberg scaling in ITDM. Note that, from the above calculation, it trivially follows that this scaling is also true for general N-qudit states.

H Demonstration of advantage in $N = 1$ for nonoptimal encoding

In the case of NITDM by choosing $\rho_i = \rho\forall i$ as probe state and measuring in the $\{|+\rangle, |-\rangle\}$ basis on the control qubit, the FI is given by

$$\mathcal{F}_\theta = \frac{N^2 r^{2N-2}(\boldsymbol{\nu})\left\{\dot{r}(\boldsymbol{\nu})\cos\left(Nf(\boldsymbol{\nu}) + \theta_c\right) - r(\boldsymbol{\nu})\dot{f}(\boldsymbol{\nu})\sin\left(Nf(\boldsymbol{\nu}) + \theta_c\right)\right\}^2}{1 - r^{2N}(\boldsymbol{\nu})\cos^2\left(Nf(\boldsymbol{\nu}) + \theta_c\right)}. \tag{63}$$

where $r(\boldsymbol{\nu}) = \left|\mathrm{Tr}\left(\sum_j K_j^T\rho K_j^\dagger\right)\right|$ and $f(\boldsymbol{\nu}) = \arg\left[\mathrm{Tr}\left(\sum_j K_j^T\rho K_j^\dagger\right)\right]$. In case of $N = 1$, Eq. (63) reduces to

$$\mathcal{F}_\theta = \frac{\left\{\dot{r}(\boldsymbol{\nu})\cos\left(f(\boldsymbol{\nu}) + \theta_c\right) - r(\boldsymbol{\nu})\dot{f}(\boldsymbol{\nu})\sin\left(f(\boldsymbol{\nu}) + \theta_c\right)\right\}^2}{1 - r(\boldsymbol{\nu})\cos^2\left(f(\boldsymbol{\nu}) + \theta_c\right)} = \frac{\left\{\frac{d}{d\theta}\Big(r(\boldsymbol{\nu})\cos\left(f(\boldsymbol{\nu}) + \theta_c\right)\Big)\right\}^2}{1 - \Big(r(\boldsymbol{\nu})\cos\left(f(\boldsymbol{\nu}) + \theta_c\right)\Big)^2} \tag{64}$$

From

$$r(\boldsymbol{\nu}) = \left|\mathrm{Tr}\left(\sum_j K_j^T\rho K_j^\dagger\right)\right| = \left|\frac{1+q}{2} - n_2^2 q(1-\cos\theta) + iqn_2\Big((1-\cos\theta)(n_1r_3 - r_1n_3) + r_2\sin\theta\Big)\right|, \tag{65}$$

we have

$$r(\boldsymbol{\nu})\cos\left(f(\boldsymbol{\nu}) + \theta_c\right) = \mathrm{Re}\left\{\mathrm{Tr}\left(\sum_j K_j^T\rho K_j^\dagger\right)\right\} = \frac{1+q}{2} - n_2^2 q(1-\cos\theta). \tag{66}$$

Putting Eq. (66) in Eq. (64) we get

$$\mathcal{F}_\theta^q = \frac{n_2^4 q^2\sin^2\theta}{1 - \left\{\frac{1+q}{2} - n_2^2 q(1-\cos\theta)\right\}^2}. \tag{67}$$

Proof of optimality of $\hat{n}_2 = \hat{y}$ and comparison with "switched" strategy. Note that $\mathcal{F}_\theta^q$ can be written as

$$\mathcal{F}_\theta^q = \frac{n_2^4 q^2 \sin^2\theta}{1 - \left[\left(\frac{1+q}{2}\right)^2 + n_2^2\{n_2^2 q^2(1-\cos\theta)^2 - q(1+q)(1-\cos\theta)\}\right]}.$$

From Eq. (68) it is clear that the numerator is maximized at $n_2 = \pm 1$ where the denominator is minimized. Hence, $n_2 = \pm 1$ is the optimal encoding axis in this scenario.

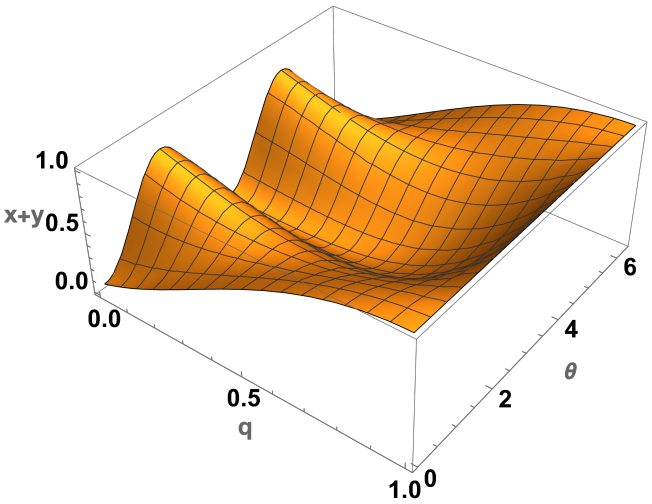

Figure 5: **Value of $x+y$ (ordinate) against noise strength q (abscissa), and paramater θ (abscissa)** The figure demonstrates the value of $x + y$ never goes above 1. All axes are dimensionless.

To compare with "switched" strategy let us evaluate the difference between $\mathcal{F}_\theta^q|_{\hat{n}=\hat{y}}$ and the "switched" strategy, defined as $\delta = \mathcal{F}_\theta^q|_{\hat{n}=\hat{y}} - \mathcal{F}_\theta^{q,S}$. Now, δ can be explicitly calculated as $\delta = 1 - x - y$ where $x = \frac{1}{16}\left\{(1+q)^2 - 4(-1+q)q\cos\theta\right\}^2$ and $y = (1-q)^2\sin^2\theta\left\{1 - \frac{1}{4}\left(-1+q-2q\cos\theta\right)^2\right\}$. Our numerical investigation suggests (see Fig. 5) that $x + y \leq 1$ which proves $\delta \geq 0\,\forall q, \theta$.

Comparing NITDM with "switched" and regular strategy with non-optimal encoding. Here, we demonstrate that, beyond the optimal encoding procedure, NITDM can also be advantageous in certain scenarios even under non-optimal encoding (see Fig. 6), outperforming both the "switched" and regular strategies. Specifically, when selecting $\hat{n} = [0.8, 0.6, 0]$ instead of $n_2 = 1$, NITDM proves to be more effective than the "switched" strategy in the region $q \gtrsim 0.73$, corresponding to the low-noise regime.

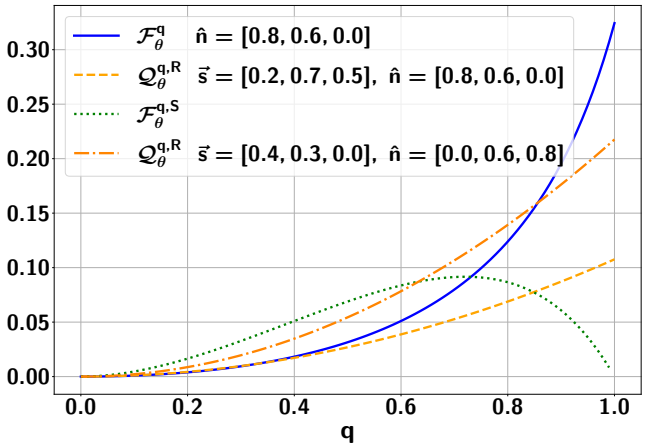

Figure 6: **Fisher information, $\mathcal{F}_\theta^q, \mathcal{F}_\theta^{q,S}$ and QFI, $\mathcal{Q}_\theta^{q,R}$ (ordinate) with respect to noise parameter q (abscissa). Comparison between Flip, Switch, and regular strategies for $n_2 \neq 1$.** We choose $\theta = \pi/4$. The figure demonstrates the advantage of time flip over "switched" and regular strategies in non-ideal encoding scenarios. All axes are dimensionless.

Conversely, maintaining n_2 constant while varying n_1, n_3, and the probe state can reduce the advantage of the time-flip approach compared to regular strategy, as illustrated in Fig. 6. While for $\hat{n} = [0.8, 0.6, 0]$ and $\hat{s} = [0.2, 0.7, 0.5]$ NITDM always excels regular strategy, in case of $\hat{n} = [0, 0.6, 0.8]$ and $\hat{s} = [0.4, 0.3, 0.0]$ only for $q \gtrsim 0.85$ $\mathcal{F}_\theta^q$ is higher. Nevertheless, in many cases within the low-noise regime, NITDM exhibits superior performance.

I θ-averaged performance of single qubit-NITDM

Since both $\mathcal{F}_\theta^q$ and $\mathcal{F}_\theta^{q,S}$ are θ-dependent, to compare the performance of protocols with the aid of switch and flip, we evaluate the difference of the average FI obtained through these strategies, i.e., $\langle \mathcal{F}_\theta^q \rangle - \langle \mathcal{F}_\theta^{q,S} \rangle = \frac{\sqrt{3}}{8}(1-q)\sqrt{(1-q)(3+5q)} + \frac{1}{8}\sqrt{(5+6q-3q^2)(5-2q+5q^2)} - \frac{1}{4}\sqrt{(3+q)(3+q-4qn_2^2)} - \frac{1}{4}\sqrt{(1-q)(1-q+4qn_2^2)}$. Analyzing this particular equation is nontrivial because of the associated square roots. Hence, we opt for a numerical approach instead. One thing we know is that for $n_2 = \pm 1$, NITDM always outperforms the "switched" one at all q and θ values. Therefore, it always beats the switched method on average for all q values. For a given value of q we analyze the range of n_2 values $\pm n_2^{\min} \leq n_2 \leq \pm 1$ where flip strategy is better, i.e., $\langle \mathcal{F}_\theta^q \rangle > \langle \mathcal{F}_\theta^{q,S} \rangle$ and fit $n_2^{\min}$ with, $n_2^{\min} = n_2^0[1 - q^g]^{h^{-1}}$. Fig. 7 represents the χ^2-fitted [85, 86] curve of $n_2^{\min}$ with $g = 1.14147 \pm (0.48\%)$ and $h = 2.38455 \pm (0.25\%)$ as fitting parameters and the constant $n_2^0 = 0.945742$. Although the magnitude of the advantage decreases significantly as n_2 and q deviate from unity, in low-noise regimes, metrology employing time-flip exhibits superior performance compared to that using switch over a broad range of n_2.

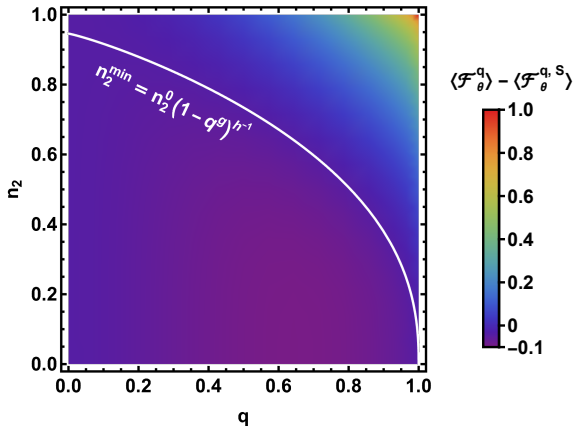

Figure 7: **Comparison between Flip and switch strategy on average. Variation of $\langle \mathcal{F}_\theta^q \rangle - \langle \mathcal{F}_\theta^{q,S} \rangle$ (heatmap) against n_2 (vertical axis) and noise strength q (horizontal axis).** The white contour is the χ^2-fitted line of $n_2^{\min}$ where the difference $\langle \mathcal{F}_\theta^q \rangle - \langle \mathcal{F}_\theta^{q,S} \rangle = 0$. Fitted curve is $n_2^{\min} = n_2^0[1-q^g]^{h^{-1}}$ with $g = 1.14147 \pm (0.48\%)$ and $h = 2.38455 \pm (0.25\%)$ as fitting parameters and a constant $n_2^0 = 0.945742$. In the regions above the curve flip outperforms switch. All axes are dimensionless.

J θ-averaged performance of multiqubit-NITDM

Similar to the case where θ is fixed, in θ-averaged scenario we define $\{\langle \tilde{\mathcal{F}}_\theta^q \rangle, \tilde{N}\}$ corresponding to which $\langle \mathcal{F}_\theta^q / N \rangle$ is maximized at $N = \tilde{N}$. Given $q = 0.99$ and 0.95, in Fig. 8 we demonstrate the behavior of θ-averaged FI with respect to the number of qubits and corresponding quantities are listed in Table 1.

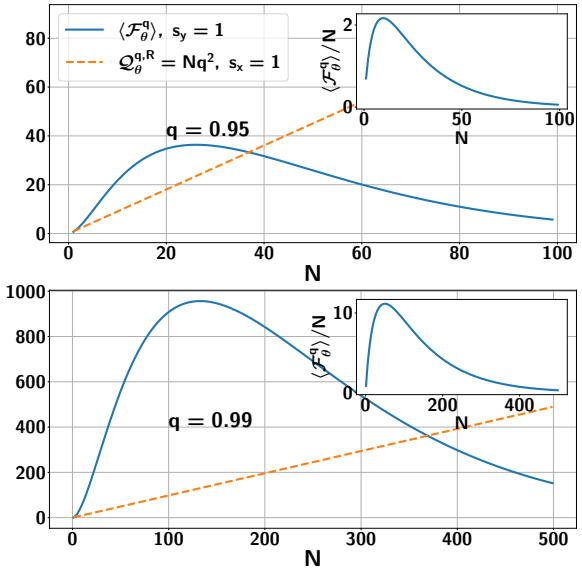

Figure 8: **Advantage of average performance of multi qubit NITDM and regular strategy.** Average FI, $\langle \mathcal{F}_\theta^q \rangle$ (blue lines) with respect to number or qubits N (abscissa) for different noise strength q. QFI in regular strategy, $\mathcal{Q}_\theta^{q,R}$ (orange dashed lines) against N. Inset: Average Fisher information per qubit, $\langle \mathcal{F}_\theta^q \rangle / N$ (ordinate) against number or qubits N (abscissa). Here, we fix the parameter value $\theta = \pi/4$ and the encoding axis as $\hat{n}_2 = 1$. All axes are dimensionless.

From Table 1, it is evident that given $N = 1000$ and $q = 0.99$, it is better to use 20 number of 50-qubit NITDM circuits to get $20\langle \tilde{\mathcal{F}}_\theta^{0.99} \rangle = 11160$ where $Nq^2 = 980.1$. Similarly, for $q = 0.95$ we can use 10-qubit circuits 100 times which leads to $100\langle \tilde{\mathcal{F}}_\theta^{0.95} \rangle = 2170$, but $Nq^2 = 902.5$.

q	$\langle \tilde{\mathcal{F}}_\theta^q \rangle / \tilde{N}$	$\tilde{N}$	$\max_N \langle \mathcal{F}_\theta^q \rangle$	N corresponding to $\max_N \langle \mathcal{F}_\theta^q \rangle$
0.95	2.17	10	36.4	26
0.99	11.16	50	956.4	133

Table 1: Given two values of q, the maximum value of both θ-averaged FI, $\max_N \langle \mathcal{F}_\theta^q \rangle$ and θ-averaged FI per qubit, $\langle \tilde{\mathcal{F}}_\theta^q \rangle / \tilde{N}$, and the corresponding number of qubits for each maximum value are listed from Fig. 8.