Complete Information Balance in Quantum Measurement

Seung-Woo Lee[1,2], Jaewan Kim[3], and Hyunchul Nha[4]

[1]Center for Quantum Information, Korea Institute of Science and Technology, Seoul, 02792, Korea

[2]Quantum Universe Center, Korea Institute for Advanced Study, Seoul 02455, Korea

[3]School of Computational Sciences, Korea Institute for Advanced Study, Seoul 02455, Korea

[4]Department of Physics, Texas A&M University at Qatar, Education City, P.O.Box 23874, Doha, Qatar

Quantum measurement is a basic tool to manifest intrinsic quantum effects from fundamental tests to quantum information applications. While a measurement is typically performed to gain information on a quantum state, its role in quantum technology is indeed manifold. For instance, quantum measurement is a crucial process element in measurement-based quantum computation. It is also used to detect and correct errors thereby protecting quantum information in error-correcting frameworks. It is therefore important to fully characterize the roles of quantum measurement encompassing information gain, state disturbance and reversibility, together with their fundamental relations. Numerous efforts have been made to obtain the trade-off between information gain and state disturbance, which becomes a practical basis for secure information processing. However, a complete information balance is necessary to include the reversibility of quantum measurement, which constitutes an integral part of practical quantum information processing. We here establish all pairs of trade-off relations involving information gain, disturbance, and reversibility, and crucially the one among all of them together. By doing so, we show that the reversibility plays a vital role in completing the information balance. Remarkably, our result can be interpreted as an information-conservation law of quantum measurement in a nontrivial form. We completely identify the conditions for optimal measurements that satisfy the conservation for each tradeoff relation with their potential applications. Our work can provide a useful guideline for designing a quantum measurement in accordance with the aims of quantum information processors.

1 Introduction

Since the early days of quantum mechanics, quantum measurement has been one of the central topics in quantum theory [1]. Unlike the measurement in classical world that only passively provides information on a physical state, quantum measurement has a critical role to reveal quantum phenomena from fundamental tests to quantum technology applications. For instance, incompatible quantum measurements are essential to demonstrate quantum nonlocality [2, 3]. Quantum measurement is also a crucial tool in quantum information protocols for not only extracting information but also processing computation in one-way quantum computation [4], protecting quantum information by detecting and correcting errors in quantum error-correcting protocols, *etc.* [5–12]. That is, the role of quantum measurement is manifold in quantum technology, which requires a thorough characterization encompassing all relevant aspects of measurement.

A well-known effect of quantum measurement is its inevitable disturbance on quantum states, which constitutes a profound basis for secure quantum information processing, e.g. quantum cryptography [13, 14]. It has been a topic of fundamental and practical importance to identify information balance in quantum measurement, particularly the trade-off relation between information gain and state disturbance [15–26]. On the other hand, the usual notion that quantum measurement is irreversible has been substantially reassessed in recent years [27–44]. It turns out that a quantum measurement weakly interacting with a system can be reversed even to faithfully recover the input state with a non-zero success probability [27, 28]. This issue of reversibility thus becomes an integral part to consider in characterizing a quantum measurement in its entirety [27–30]. The measurement reversibility has also widely applied to range of applications, e.g. quantum error corrections [31], gate operations [32, 33] and decoherence suppression [34–36]. Reversing the disturbance of quantum measurement was experimentally demonstrated with superconducting [37], trapped-ion [38] and photonic qubits [35, 36, 39]. Along this line, a trade-off of the reversibility against information gain in quantum measurement was rigorously assessed [40–42]. Trade-offs among information gain, state disturbance and reversibility were also quantitatively studied at the level of a single measurement outcome [43, 44].

Seung-Woo Lee: swleego@gmail.com

However, there does not exist a unified framework yet to deal with all the information contents together as universal quantities in general quantum measurement [26], which is essential to draw a complete picture on information balance in quantum measurement. This is very important for a broad applicability, e.g. designing a quantum measurement to be adapted to the aim of quantum information processing at hand. Here we establish such a complete information balance in quantum measurement. We derive the full trade-off relations among those three information contents, information gain $\mathcal{G}$, disturbance $\mathcal{D}$, and reversibility $\mathcal{R}$. We particularly show that the reversibility $\mathcal{R}$ plays a crucial role in information balance accounting for the gap between $\mathcal{G}$ and $\mathcal{D}$. That is, we find that the global trade-off relation ($\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$) tightens the trade-off between the gain and disturbance ($\mathcal{G}$-$\mathcal{D}$) only [20]. On the other hand, the trade-off between disturbance and reversibility ($\mathcal{D}$-$\mathcal{R}$) compensates the gain and reversibility relation ($\mathcal{G}$-$\mathcal{R}$). Consequently, our result shows that the total information is balanced in an ideal quantum measurement process, which can be interpreted as a conservation of information contents in quantum measurement.

We fully obtain the conditions to saturate all of the trade-off relations and thereby define an *optimal* quantum measurement providing maximal information contents quantum mechanics fundamentally allows. Our framework can offer useful guidelines for designing measurement-based quantum information protocols in quantum computation [9], teleportation [10], quantum metrology [11], and quantum error corrections [12], *etc.*. As all the information contents defined here are directly measurable, our results are readily testable and applicable to any quantum information platforms. Our work may contribute to deepening our fundamental understanding of quantum measurement and provide a rigorous practical benchmark to optimize measurement-based quantum information protocols.

2   General framework

We begin with the general framework for addressing the information changes by a quantum measurement $\mathbf{M}$. Assume that an arbitrary quantum state ρ is prepared to convey information. We perform a quantum measurement $\mathbf{M}$ to extract the information, which disturbs and changes the input state ρ to another state. We then apply a subsequent reversing operation $\mathbf{R}$ to characterize the reversibility of $\mathbf{M}$ (see Fig. 1(a)). The reversing operation is assumed to be chosen to recover the input state ρ, i.e., $(\mathbf{R} \circ \mathbf{M})(\rho) \propto \rho$.

Without loss of generality, we assume here that the input information is encoded onto pure states $\rho = |\psi\rangle\langle\psi|$ in d-dimensional Hilbert space, but the results derived in what follows are valid for any mixed
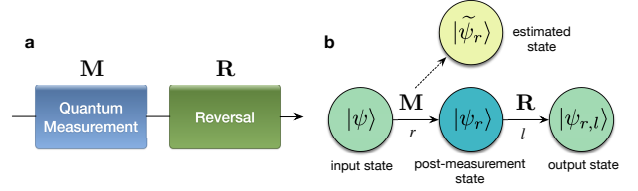

Figure 1: (a) Quantum measurement $\mathbf{M}$ and reversal process $\mathbf{R}$. (b) The input, the post-measurement, the estimated and the output state in a single trial, when the outcomes of the measurement and the reversal are r and l, respectively, are denoted by $|\psi\rangle$, $|\psi_r\rangle$, $|\widetilde{\psi_r}\rangle$, and $|\psi_{r,l}\rangle$, respectively. The closeness of those states on average determines the information contents of a given quantum measurement $\mathbf{M}$.

input states. Consider a quantum measurement described by a set of operators $\mathbf{M} = \{\hat{M}_r | r = 1, \dots, n\}$, satisfying the completeness relation $\sum_r \hat{M}_r^\dagger \hat{M}_r = \hat{\mathbb{1}}$ [6, 7]. When the measurement outcome is r, the input state is changed to $|\psi_r\rangle = \hat{M}_r|\psi\rangle/\sqrt{p(r,\psi)}$ where $p(r,\psi) = \langle\psi|\hat{M}_r^\dagger \hat{M}_r|\psi\rangle$. For each outcome r, we may estimate the input state as $|\widetilde{\psi_r}\rangle$ by a certain estimation strategy. We then apply a subsequent reversing operation described by a set of operators $\mathbf{R} = \{\hat{R}_{r,l} | l = 1, \dots, m\}$, satisfying the completeness relation $\sum_l \hat{R}_{r,l}^\dagger \hat{R}_{r,l} = \hat{\mathbb{1}}$ for each r. The final output state is then given by $|\psi_{r,l}\rangle = \hat{R}_{r,l}\hat{M}_r|\psi\rangle/\sqrt{p(r,l,\psi)}$, where $p(r,l,\psi) = \langle\psi|\hat{M}_r^\dagger \hat{R}_{r,l}^\dagger \hat{R}_{r,l} \hat{M}_r|\psi\rangle$. The reversing operation $\mathbf{R}$ here is assumed to be appropriately chosen according to the measurement outcome r. We note that the whole process cannot be described by a unitary operation as it can be probabilistic and conditional on the result of $\mathbf{M}$ and observer's choice of $\mathbf{R}$. Note that our framework differs from the Petz recovery map [45, 46].

3   Information contents

Let us define the information contents within the above general framework. For a given quantum measurement, the changes of information due to the measurement $\mathbf{M}$ and the reversing operation $\mathbf{R}$ can be defined in terms of the closeness between the input $|\psi\rangle$, the post-measurement $|\psi_r\rangle$, the estimated $|\widetilde{\psi_r}\rangle$, and the output $|\psi_{r,l}\rangle$ states. See Fig. 1(b) together with the details of the definitions in Appendix A.

(i) *Information Gain*: The amount of information obtained by $\mathbf{M}$ can be quantified based on the overlap between the input $|\psi\rangle$ and the estimated state $|\widetilde{\psi_r}\rangle$. The estimation fidelity is obtained by averaging $|\langle\widetilde{\psi_r}|\psi\rangle|^2$ over all input states and possible measurement outcomes, i.e., $\int d\psi \sum_{r=1}^n p(r,\psi)|\langle\widetilde{\psi_r}|\psi\rangle|^2$, which has different values depending on the estimation strategies. We define the *information gain* as the maximum estimation fidelity over all possible strate-

gies, which can be obtained as (see Appendix A.2),

$$\mathcal{G} = \frac{1}{d(d+1)}\left[d + \sum_{r=1}^{n}(\lambda_0^r)^2\right], \qquad (1)$$

where λ_0^r is the largest singular value of $\hat{M}_r$ in the singular value decomposition (Appendix A.1) [20]. The information gain thus lies in the range $1/d \leq \mathcal{G} \leq 2/(d+1)$, where the upper bound is reached by a von Neumann measurement and the lower bound by a unitary operation or a random guess.

(ii) *Disturbance*: In order to quantify the amount of disturbance by **M**, we consider the *operation fidelity* first. The operation fidelity of **M** is given by averaging the overlap between the input $|\psi\rangle$ and the post-measurement state $|\psi_r\rangle$, i.e., $\int d\psi \sum_{r=1}^{n} p(r,\psi)|\langle\psi_r|\psi\rangle|^2$. We can then evaluate the maximum of the operation fidelity (see Appendix A.3), resulting in

$$\mathcal{F} = \frac{1}{d(d+1)}\left[d + \sum_{r=1}^{n}\left(\sum_{i=0}^{d-1}\lambda_i^r\right)^2\right]. \qquad (2)$$

In this case, all the singular values λ_i^r of $\hat{M}_r$ are involved in determining the average fidelity of the disturbed states. We may then define the *disturbance* induced by **M** as the minimum operation infidelity given by

$$\mathcal{D} = 1 - \mathcal{F}, \qquad (3)$$

scaled in the range $0 \leq \mathcal{D} \leq (d-1)/(d+1)$.

Similarly, the operation fidelity of the overall process **R** ∘ **M** without postselection can be obtained by evaluating the average fidelity between the input $|\psi\rangle$ and output state $|\psi_{r,l}\rangle$, i.e., $\int d\psi \sum_{r=1}^{n}\sum_{l=1}^{m} P(r,l,\psi)|\langle\psi_{r,l}|\psi\rangle|^2$. Its maximum can be obtained as (see Appendix A.4)

$$\mathcal{F}(\mathbf{R} \circ \mathbf{M}) = \frac{1}{d(d+1)}\left[d + \sum_{r=1}^{n}\sum_{l=1}^{m}\left(\sum_{i=0}^{d-1}\lambda_i^r\lambda_i^{r,l}\right)^2\right], \qquad (4)$$

where the singular values $\lambda_i^{r,l}$ of $\hat{R}_{r,l}$ are additionally involved in determining the output fidelity. The operation fidelity lies in the range $2/(d+1) \leq \mathcal{F} \leq 1$, where the upper bound is reached by a unitary operation and the lower bound by a von Neumann measurement.

(iii) *Reversibility*: We consider the success event of the reversing operation **R**, which faithfully recovers the input state, i.e., $|\psi_{r,l}\rangle \propto |\psi\rangle$, to evaluate the reversibility. Assume that the operators $\hat{R}_{r,l}$ for $l = 1,\ldots,s < m$ are associated with the success events s.t. $\hat{R}_{r,l}\hat{M}_r|\psi\rangle = \eta_{r,l}|\psi\rangle$, where $|\eta_{r,l}|^2$ is the success probability of the reversing operation when the outcome is r. The *reversibility* is then obtained as the maximum overall success probability,

$$\mathcal{R} = \max_{\{\hat{R}_{r,l}\}} \int d\psi \sum_{r=1}^{n}\sum_{l=1}^{s<m} |\langle\psi|\hat{R}_{r,l}\hat{M}_r|\psi\rangle|^2 = \sum_{r=1}^{n}(\lambda_{d-1}^r)^2, \qquad (5)$$

where λ_{d-1}^r is the smallest singular value of $\hat{M}_r$ [40]. The reversibility is scaled as $0 \leq \mathcal{R} \leq 1$.

The optimal reversing operator can be defined in the following context. Assume that the measurement operator $\hat{M}_r$ is represented in the singular value decomposition as $\hat{M}_r = \hat{V}_r\hat{D}_r$ with a unitary operator $\hat{V}_r$ and a diagonal matrix $\hat{D}_r = \sum_{i=0}^{d-1}\lambda_i^r|i\rangle\langle i|$ (see Appendix A.1). Its optimal reversing operator can then be written by $\hat{R}_{r,1} = \lambda_{d-1}^r\hat{D}_r^{-1}\hat{V}_r^{\dagger}$, where $\hat{D}_r^{-1} = \sum_{i=0}^{d-1}(\lambda_i^r)^{-1}|i\rangle\langle i|$ with nonzero λ_i^r (Here, we set $s = 1$ without loss of generality). For example, a quantum measurement described by $\hat{M}_1 = \sqrt{\eta}|1\rangle\langle 1|$ and $\hat{M}_2 = |0\rangle\langle 0| + \sqrt{1-\eta}|1\rangle\langle 1|$ with measurement strength $0 \leq \eta \leq 1$, can be optimally reversed by $\hat{R}_{2,1} = \sqrt{1-\eta}|0\rangle\langle 0| + |1\rangle\langle 1|$ and $\hat{R}_{2,2} = \sqrt{\eta}|0\rangle\langle 0|$ for the measurement outcome $r = 2$ and the reversal process $l = 1$, respectively. In this case, the reversibility is given by $\mathcal{R} = 1 - \eta$ from Eq. (5). Note that a unitary operation (no measurement) is deterministically reversible $\mathcal{R} = 1$ while a von Neumann measurement is completely irreversible $\mathcal{R} = 0$.

We now have three information contents characterizing a quantum measurement, i.e., information gain $\mathcal{G}$, disturbance $\mathcal{D}$, and reversibility $\mathcal{R}$. These are universal quantities averaged over all input states of a given dimension d and have a clear operational meaning in terms of quantum fidelity, fulfilling the requirements for the information contents in quantum measurement [26]. Note that the information contents evaluated here with pure input states $\rho = |\psi\rangle\langle\psi|$ and the accompanying consequences in information balance are generally valid for arbitrary mixed input states ρ, since the maximum averaged in the space of pure states must represent the maximum in the space of convex combination of pure states, i.e. mixed states.

4 Information balance in quantum measurement

In this section, we now derive the trade-off relations among the information contents of quantum measurement. Before presenting those trade-off relations, we first introduce and prove a useful inequality on the relation between the reversibility $\mathcal{R}$ and the overall operation fidelity $\mathcal{F}(\mathbf{R} \circ \mathbf{M})$ after the measurement followed by the reversing operation defined in the previous section:

Lemma 1. For any quantum measurement **M** and subsequent reversing operation **R**, the reversibility $\mathcal{R}$ and the operation fidelity of the overall process $\mathcal{F}(\mathbf{R} \circ \mathbf{M})$ always satisfy an inequality

$$2 + (d-1)\mathcal{R} \leq (d+1)\mathcal{F}(\mathbf{R} \circ \mathbf{M}), \qquad (6)$$

in arbitrary finite dimension $d \geq 2$.

Detailed proof of Lemma 1 is given in Appendix B. We define a vector $\vec{u}_i^l = (\lambda_i^{r=1}\lambda_i^{r=1,l}, \ldots, \lambda_i^{r=n}\lambda_i^{r=n,l})$ for $i = 0, \ldots, d-1$ in terms of the singular values λ_i^r of the measurement operators $\{\hat{M}_r\}$ and $\lambda_i^{r,l}$ of the reversing operators $\{\hat{R}_{r,l}\}$. The equality in (6) holds if and only if the quantum measurement satisfies

$$\vec{u}_i^l \cdot \vec{u}_j^l = \delta_{ij}|\vec{u}_i^l|^2, \ \forall l \neq 1. \tag{7}$$

We will use Lemma 1 to derive the trade-off relations presented in what follows.

Before presenting our main results, we introduce two previously known trade-off relations between information contents, denoted by $\mathcal{G}$-$\mathcal{D}$ [20] and $\mathcal{G}$-$\mathcal{R}$ [40] as below.

($\mathcal{G}$-$\mathcal{D}$: Information gain and Disturbance trade-off) The trade-off relation between the information gain $\mathcal{G}$ and disturbance $\mathcal{D} = 1 - F$ was derived in Ref. [20] as

$$\sqrt{\mathcal{F} - \frac{1}{d+1}} \leq \sqrt{\mathcal{G} - \frac{1}{d+1}} + \sqrt{(d-1)\left(\frac{2}{d+1} - \mathcal{G}\right)}. \tag{8}$$

This is the quantitative proof of the heuristic knowledge 'the more information a quantum measurement obtains, the more disturbed the quantum state becomes.' It gives the lower bound of diturbance for a given amount of information that a quantum measurement extracts.

($\mathcal{G}$-$\mathcal{R}$: Information gain and Reversibility trade-off) The trade-off relation between the information gain $\mathcal{G}$ and the reversibility $\mathcal{R}$ was derived in Ref. [40] as

$$d(d+1)\mathcal{G} + (d-1)\mathcal{R} \leq 2d, \tag{9}$$

which was the first information-theoretic approach introducing the role of the reversibility in quantum measurement. It captures the idea that 'the more information a quantum measurement obtains, the less reversible the quantum measurement is.'

However, a global trade-off relation including all three information contents has been missing so far. The full quantitative links among the three information contents have also not been completed (see Fig. 2). Let us now derive this global trade-off relation, including all the information contents, $\mathcal{G}$, $\mathcal{D}$, and $\mathcal{R}$, aiming to complete the total information balance as follows.

Theorem 1. ($\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$: Global trade-off relation) The information gain $\mathcal{G}$, the disturbance $\mathcal{D}$ and the reversibility $\mathcal{R}$ of quantum measurement always satisfy an inequality

$$\sqrt{\mathcal{F} - \frac{1}{d+1}} \leq \sqrt{\mathcal{G} - \frac{1}{d+1}} + \sqrt{\frac{\mathcal{R}}{d(d+1)}} \tag{10}$$

$$+ \sqrt{(d-2)\left(\frac{2}{d+1} - \mathcal{G} - \frac{\mathcal{R}}{d(d+1)}\right)},$$

with $\mathcal{F} = 1 - \mathcal{D}$, in an arbitrary dimension $d \geq 2$.

The inequality in (10) determines the quantitative relation among the three information contents, $\mathcal{G}$, $\mathcal{D}$, and $\mathcal{R}$ (see Appendix C for the details of the proof). This relation can be interpreted in various ways. It draws the upper bound of $\mathcal{F}$ or equivalently the lower bound of $\mathcal{D}$ with respect to both $\mathcal{G}$ and $\mathcal{R}$. Or, it indicates the maximum possible reversibility for a given pair of the information amounts $\mathcal{G}$ and $\mathcal{D}$.

When $d = 2$, the inequality in (10) is equivalent to $\mathcal{G}$-$\mathcal{D}$ in (8). This is due to the relation between the information gain $\mathcal{G}$ and the reversibility $\mathcal{R}$ for $d = 2$, i.e. $\frac{2}{3} - \mathcal{G} = \frac{\mathcal{R}}{6}$. It is given by the completeness condition $\sum_r \hat{M}_r^\dagger \hat{M}_r = \hat{\mathbb{1}}$ leading to $\sum_{r=1}^n (\lambda_0^r)^2 + \sum_{r=1}^n (\lambda_1^r)^2 = d = 2$ together with Eqs. (1) and (5).

On the other hand, for $d > 2$ in general, the $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ inequality in (10) is fundamentally different from $\mathcal{G}$-$\mathcal{D}$ in (8) and $\mathcal{G}$-$\mathcal{R}$ in (9). Note that $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ provides a tighter bound than $\mathcal{G}$-$\mathcal{D}$ in the relation between the information gain and the disturbance. This becomes clearer with its saturation condition described below.

($\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ saturation condition) The $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ inequality (10) is saturated if and only if the quantum measurement satisfies following conditions: all $\vec{v}_i$ for $i = 0, \cdots, d-1$ are collinear and

$$|\vec{v}_1| = \cdots = |\vec{v}_{d-2}|, \tag{11}$$

where $\vec{v}_i = (\lambda_i^{r=1}, \ldots, \lambda_i^{r=n})$ is a vector defined with the singular values of $\hat{M}_r$.

See Appendix C for the detailed description of the condition. We can see that a quantum measurement satisfying the $\mathcal{G}$-$\mathcal{D}$ saturation condition (i.e., all $\vec{v}_i$ are collinear and $|\vec{v}_1| = \cdots = |\vec{v}_{d-1}|$ [20]) also satisfies the saturation condition of $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ (Table 1), but the converse is not always true. It indicates that the information balance can be more tightly characterized by $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ in a broader set of quantum measurements than $\mathcal{G}$-$\mathcal{D}$. In section 5, we will further discuss on the different sets of quantum measurements classified based on the saturation conditions of the trade-off relations. Besides, we also analytically show that the right-hand side of the inequality (10) is always lower than or equal to the right-hand side of the inequality (8) (see Appendix D), which guarantees that $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ tightens $\mathcal{G}$-$\mathcal{D}$.

We here introduce and prove another useful inequality on the operation fidelity by a reversing operation as below:

Lemma 2. The overall operation fidelity covering all output states without postselection by the quantum measurement and the reversal $\mathbf{R} \circ \mathbf{M}$ is upper bounded by the operation fidelity of the quantum measurement
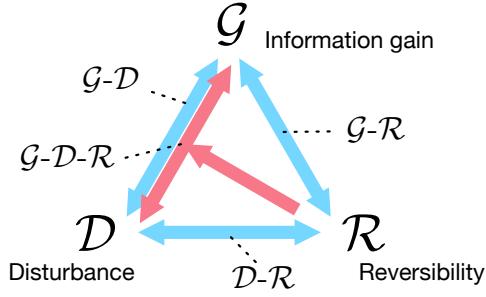
Figure 2: Three information contents of quantum measurement, i.e., information gain $\mathcal{G}$, disturbance $\mathcal{D} = 1 - \mathcal{F}$, and reversibility $\mathcal{R}$, and their quantitative links: ($\mathcal{G}$-$\mathcal{D}$) in Eq. (8) [20], ($\mathcal{G}$-$\mathcal{R}$) in Eq. (9) [40], ($\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$) Theorem 1 in Eq. (10), ($\mathcal{D}$-$\mathcal{R}$) Theorem 2 in Eq. (13).

$\mathbf{M}$, i.e.,

$$\mathcal{F}(\mathbf{R} \circ \mathbf{M}) \leq \mathcal{F}(\mathbf{M}). \tag{12}$$

See Appendix E for the details of the proof. Note that the equality in (12) holds when $\mathbf{M}$ is a unitary operation or a von Neumann measurement. The inequality (12) implies that 'the disturbance in quantum measurement never decreases by any subsequent reversing operation.' We can further generalize this for arbitrary k-times sequential quantum measurements, s.t. $\mathcal{F}(\mathbf{M}_k \circ \cdots \mathbf{M}_2 \circ \mathbf{M}_1) \leq \cdots \leq \mathcal{F}(\mathbf{M}_2 \circ \mathbf{M}_1) \leq \mathcal{F}(\mathbf{M}_1)$.

The inequality in (12) and its extension mentioned above are intuitively plausible by the second law of thermodynamics. These indicates the non-increasing of the average fidelity between the input and output states by reversing operations so that it differs from, but may be fundamentally related to, the data processing inequality [47]. We use (12) to derive a trade-off relation in what follows.

Let us now derive the quantitative relation between the disturbance and the reversibility of quantum measurement using Lemma 1 and Lemma 2.

Theorem 2. ($\mathcal{D}$-$\mathcal{R}$: Disturbance and Reversibility trade-off) The disturbance $\mathcal{D}$ and the reversibility $\mathcal{R}$ of quantum measurements always satisfy

$$(d-1)\mathcal{R} + (d+1)\mathcal{D} \leq d - 1, \tag{13}$$

in an arbitrary dimension d.

Proof– From Lemma 1 and Lemma 2, $(d-1)\mathcal{R} \leq (d+1)\mathcal{F}(\mathbf{R} \circ \mathbf{M}) - 2 \leq (d+1)\mathcal{F} - 2$. By $\mathcal{D} = 1 - \mathcal{F}$, we obtain the inequality (13). $\square$

The inequality in (13) complements other trade-off relations, determining the upper bound of $\mathcal{R}$ by $\mathcal{D}$. It implies that 'the more disturbing a quantum measurement is, the less reversible it becomes.' Here the equality holds for von Neumann measurements or unitary operations.

We have now completed the full quantitative links among the three information contents, i.e., information gain $\mathcal{G}$, disturbance $\mathcal{D}$, and reversibility $\mathcal{R}$ of

quantum measurement, as illustrated in Fig. 2. These clearly show how the total information is balanced during the measurement and the reversal process. Suppose that we choose and modify a quantum measurement, intending to extract information maximally from the input quantum state. Such an optimization inevitably increases the disturbance $\mathcal{D}$ (i.e., decreases $\mathcal{F}$) from the relation $\mathcal{G}$-$\mathcal{D}$. The operation fidelity $\mathcal{F}$ can be interpreted here as the remaining information in the post-measurement state as a part of the initially encoded information into the input state. On the other hand, there exists a hidden part of information, by which we can recover the input state using a subsequent reversing operation on the output state, quantified by $\mathcal{R}$. It turns out that $\mathcal{R}$ fills the gap between $\mathcal{G}$ and $\mathcal{D}$ and tightens further the trade-off relations beyond $\mathcal{G}$-$\mathcal{D}$ as proved in this section.

As an example, let us consider a quantum measurement with operators $\hat{M}_i = \sqrt{p}|i\rangle\langle i| + \sqrt{(1-p)(3-p)/3}|i+1\rangle\langle i+1| + \sqrt{p(1-p)/3}|i+2\rangle\langle i+2|$ ($i = 0, 1, 2$) for $0.458619 \leq p \leq 1$. The basis states are $|i\rangle \equiv |i \mod 3\rangle$ in the Hilbert space of $d = 3$. Using Eqs. (1),(2) and (5), we can obtain the information gain $\mathcal{G} = (1+p)/4$, the operation fidelity $\mathcal{F} = 1/4 + (\sqrt{p} + \sqrt{(1-p)(3-p)/3} + \sqrt{p(1-p)/3})^2/4$, and the reversibility $\mathcal{R} = p(1-p)$. These turn out to saturate the global $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ tradeoff in Eq. (10) but not the $\mathcal{G}$-$\mathcal{D}$ one in Eq. (8). In Fig. 3 (left), we show the curve (black solid) representing the values of the pair $\mathcal{G}$ and $\mathcal{F}$ corresponding to the considered measurement, together with the curve (blue dashed) representing the bound from the $\mathcal{G}$-$\mathcal{D}$ relation in Eq. (8). We clearly see that the bound by $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ is tighter than the one by $\mathcal{G}$-$\mathcal{D}$.

In view of the $\mathcal{G}$-$\mathcal{D}$ tradeoff, the measurement is not optimal in the sense that it does not maximize the output fidelity $\mathcal{F}$ for a given degree of $\mathcal{G}$. However, in fact, it is an optimal measurement saturating the $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ tradeoff in Eq. (10). The gap between the solid and the dashed curves is due to the reversibility $\mathcal{R}$, which is also in trade-off with $\mathcal{G}$ as plotted in Fig. 3 (right). It thus illustrates the importance of including all information contents $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ to characterize a
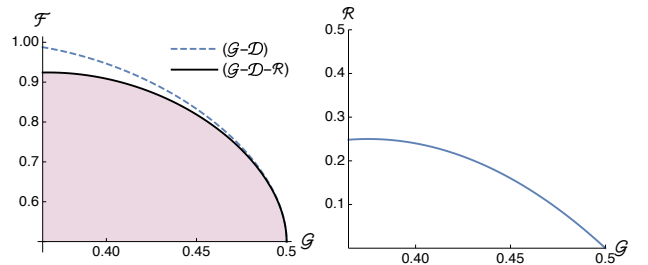

Figure 3: Left: Comparison of the upper bounds by $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ and $\mathcal{G}$-$\mathcal{D}$ relations for the measurement in the main text. The bound by $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ is tighter than the one by $\mathcal{G}$-$\mathcal{D}$. Right: The gap is due to the reversibility $\mathcal{R}$, which is also in trade-off relation with the information gain $\mathcal{G}$.

quantum measurement as optimal or non-optimal.

We note that there would exist some missing parts of the information accounted for by none of $\mathcal{G}$, $\mathcal{F}$, or $\mathcal{R}$. Such a missing part should be due to the non-optimality of quantum measurement or ignorance in the estimation of the input state based on the measurement outcomes. For instance, if we take into account the effect of errors in quantum measurement leading to imperfect reversal, the final output state may not be the same as the input state, i.e., $(\mathbf{R} \circ \mathbf{M})(\rho) \propto \rho' \neq \rho$. Specifically, let us consider the case when the reversing operation succeeds but yields the output state $\rho(\psi, \epsilon)$ as

$$\hat{R}_{r,1} \hat{M}_r |\psi\rangle\langle\psi| \hat{M}_r^\dagger \hat{R}_{r,1}^\dagger = |\eta_{r,1}|^2 \rho(\psi, \epsilon) \qquad (14)$$

which deviates from the original input state due to an error parameterized by $0 \leq \epsilon \leq 1$. The reversibility can then be evaluated as (see Appendix F)

$$\mathcal{R} = \sum_r (\lambda_{d-1}^r)^2 F_s. \qquad (15)$$

Here $F_s = \int d\psi \int_0^1 d\epsilon\, p(\epsilon)\langle\psi|\rho(\psi,\epsilon)|\psi\rangle$ is the average fidelity between the input state and the output state of the successful reversal, where $p(\epsilon)$ denotes the error probability density. It shows that errors tend to decrease the reversibility in proportion to F_s. Note that when $F_s = 1$ without errors the reversibility can reach the maximum value in Eq. (5). Similarly, the amount of information gain can be degraded from the maximum $\mathcal{G}$ in Eq. (1) in the presence of errors. Therefore, imperfection and errors result in a missing part of the total information. As a result, the information contents $\mathcal{G}$, $\mathcal{D}$ and $\mathcal{R}$ of the quantum measurement with errors satisfy but cannot saturate the information trade-off relations.

In this context, we may define the *optimal quantum measurement* as the one that reaches the upper bounds of trade-off relations without any missing part of the total information. This further generalizes the definition of optimal quantum measurement to reach the bound of $\mathcal{G}$-$\mathcal{D}$ with minimal disturbance in [23, 41, 48]. In a sense, the optimal measurement is the one that conserves the total information in a nontrivial form obeying the trade-off relations. We present how optimal quantum measurements can be classified into different sets according to the considered trade-off relations in the next section.

5 Classifying optimal quantum measurements

The saturation condition on each trade-off relation determines the criterion for classifying optimal quantum measurements. Let $\mathcal{S}$ be the set of all possible quantum measurements. We then denote by $\mathcal{S}_{\mathcal{G}\text{-}\mathcal{D}\text{-}\mathcal{R}}$,
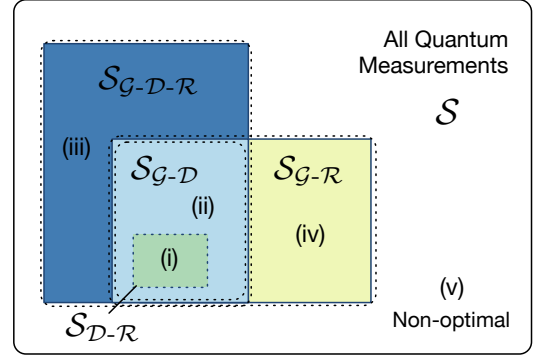

Figure 4: A Venn diagram for quantum measurements based on the optimality to reach the bounds of $\mathcal{G}$, $\mathcal{D}$, and $\mathcal{R}$ by the trade-off relations. The quantum measurements (i)-(iv) presented in the main text are the representative examples of the different types of optimal quantum measurements in each divided region, while the measurement in the region (v) is non-optimal.

$\mathcal{S}_{\mathcal{G}\text{-}\mathcal{D}}$, $\mathcal{S}_{\mathcal{G}\text{-}\mathcal{R}}$, and $\mathcal{S}_{\mathcal{D}\text{-}\mathcal{R}}$ those subsets of quantum measurements saturating the trade-off relations $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$, $\mathcal{G}$-$\mathcal{D}$, $\mathcal{G}$-$\mathcal{R}$, and $\mathcal{D}$-$\mathcal{R}$, respectively. The saturation condition of $\mathcal{G}$-$\mathcal{R}$ is $\hat{M}_r^\dagger \hat{M}_r = a_r |j_r\rangle\langle j_r| + b_r \hat{\mathbb{1}}$ where $j_r \in \{0, \cdots, d-1\}$ and a_r and b_r are non-negative parameters [40]. It was shown in Ref. [41] that a quantum measurement saturating $\mathcal{G}$-$\mathcal{D}$ always satisfies the condition saturating $\mathcal{G}$-$\mathcal{R}$, while the converse is not true, s.t. $\mathcal{S}_{\mathcal{G}\text{-}\mathcal{R}} \supset \mathcal{S}_{\mathcal{G}\text{-}\mathcal{D}}$. In the previous section from Theorem 1, we have observed that $\mathcal{S}_{\mathcal{G}\text{-}\mathcal{D}\text{-}\mathcal{R}} \supset \mathcal{S}_{\mathcal{G}\text{-}\mathcal{D}}$, i.e, a quantum measurement saturating $\mathcal{G}$-$\mathcal{D}$ always saturates $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$, but the converse is not true. We further find that $\mathcal{S}_{\mathcal{G}\text{-}\mathcal{D}\text{-}\mathcal{R}} \cap \mathcal{S}_{\mathcal{G}\text{-}\mathcal{R}} = \mathcal{S}_{\mathcal{G}\text{-}\mathcal{D}}$ from the saturation conditions of $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ and $\mathcal{G}$-$\mathcal{R}$ (Table 1). The detailed proof is given in Appendix G. The elements of $\mathcal{S}_{\mathcal{D}\text{-}\mathcal{R}}$ are either unitary operators or von Neumann measurements by Theorem 2. Now, the Venn diagram of optimal quantum measurement sets can be constructed as illustrated in Fig. 4, classified based on the optimality of quantum measurement to reach the upper bounds of information contents by the trade-off relations.

Let us consider some examples of different types of quantum measurements.

(i) Assume that a von Neumann measurement $\hat{P} = |i\rangle\langle i|$ is performed on arbitrary d-dimensional quan-

Table 1: Saturation condition on the trade-off relations.

	Saturation condition
$\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$	all $\vec{v}_i$ are collinear[a] & $|\vec{v}_1| = \cdots = |\vec{v}_{d-2}|$
$\mathcal{G}$-$\mathcal{D}$	all $\vec{v}_i$ are collinear & $|\vec{v}_1| = \cdots = |\vec{v}_{d-1}|$
$\mathcal{G}$-$\mathcal{R}$	$\hat{M}_r^\dagger \hat{M}_r = a_r |j_r\rangle\langle j_r| + b_r \hat{\mathbb{1}}$[b] with $a_r, b_r \geq 0$
$\mathcal{D}$-$\mathcal{R}$	von Neumann meas. or unitary op.

[a] $\vec{v}_i = (\lambda_i^{r=1}, \ldots, \lambda_i^{r=n})$

[b] $j_r \in \{0, \ldots, d-1\}$

tum states. It allows one to obtain the maximum information $\mathcal{G} = 2/(d+1)$. Therefore, no significant information remains on the post-measurement state $\mathcal{F} = 2/(d+1)$ nor is recoverable $\mathcal{R} = 0$ (we remind that the range of information gain, operation fidelity, and reversibility are $1/d \le \mathcal{G} \le 2/(d+1)$, $2/(d+1) \le \mathcal{F} \le 1$, and $0 \le \mathcal{R} \le 1$, respectively). It is straightforward to see that these quantities saturate all the trade-off relations (Fig. 4).

(ii) Consider a quantum measurement $\hat{M}_i = \sqrt{p}|i\rangle\langle i| + \sqrt{(1-p)/2}(\hat{\mathbb{1}} - |i\rangle\langle i|)$ $(i = 0,1,2)$ for $1/3 \le p \le 1$. It becomes a von Neumann measurement when $p = 1$ and a unitary operator when $p = 1/3$. It causes a partial collapse of the input state and can be reversed for $1/3 < p < 1$. For each outcome i, the optimal reversing operation is defined by $\hat{R}_{i,1} = \sqrt{(1-p)/2p}|i\rangle\langle i| + (\hat{\mathbb{1}} - |i\rangle\langle i|)$ (success) and $\hat{R}_{i,2} = \sqrt{(3p-1)/2p}|i\rangle\langle i|$ (non-success). We obtain the information contents as $\mathcal{G} = (1+p)/4$, $\mathcal{F} = (3 - p + 2\sqrt{2p(1-p)})/4$, and $\mathcal{R} = 3(1-p)/2$, which are saturating $\mathcal{G}$-$\mathcal{D}$, $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$, and $\mathcal{G}$-$\mathcal{R}$ relations, i.e., $\{\hat{M}_r\} \in S_{\mathcal{G}\text{-}\mathcal{D}\text{-}\mathcal{R}}$, $\{\hat{M}_r\} \in S_{\mathcal{G}\text{-}\mathcal{D}}$ and $\{\hat{M}_r\} \in S_{\mathcal{G}\text{-}\mathcal{R}}$.

(iii) Consider a quantum measurement $\hat{M}_i = \sqrt{p}|i\rangle\langle i| + \sqrt{2(1-p)/3}|i+1\rangle\langle i+1| + \sqrt{(1-p)/3}|i+2\rangle\langle i+2|$ $(i = 0,1,2)$ for $2/5 \le p \le 1$ with $|i\rangle \equiv |i \bmod 3\rangle$ in 3-dimensional Hilbert space. It can be optimally reversed by $\hat{R}_{i,1} = \sqrt{(1-p)/3p}|i\rangle\langle i| + \sqrt{1/2}|i+1\rangle\langle i+1| + |i+2\rangle\langle i+2|$ (success) and $\hat{R}_{i,2} = \sqrt{(4p-1)/3p}|i\rangle\langle i| + \sqrt{1/2}|i+1\rangle\langle i+1|$ (non-success). We obtain $\mathcal{G} = (1+p)/4$, $\mathcal{F} = (3 + \sqrt{2}(1-p) + (\sqrt{3} + \sqrt{6})\sqrt{p(1-p)})/6$, and $\mathcal{R} = 1 - p$. These saturate $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ but do not saturate $\mathcal{G}$-$\mathcal{D}$ and $\mathcal{G}$-$\mathcal{R}$, i.e., $\{\hat{M}_r\} \in S_{\mathcal{G}\text{-}\mathcal{D}\text{-}\mathcal{R}}$, $\{\hat{M}_r\} \notin S_{\mathcal{G}\text{-}\mathcal{D}}$ and $\{\hat{M}_r\} \notin S_{\mathcal{G}\text{-}\mathcal{R}}$.

(iv) Consider a quantum measurement $\hat{M}_i = \sqrt{1/3}|i\rangle\langle i| + \sqrt{p/6}(|i+1\rangle\langle i+1| + |i+2\rangle\langle i+2|)$ $(i = 0,1)$ and $\hat{M}_2 = \sqrt{(3-p)/3}|2\rangle\langle 2| + \sqrt{(4-p)/6}(|0\rangle\langle 0| + |1\rangle\langle 1|)$ for $0 \le p \le 1$. It can be optimally reversed by $\hat{R}_{i,1} = \sqrt{p/2}|i\rangle\langle i| + |i+1\rangle\langle i+1| + |i+2\rangle\langle i+2|$ (success) and $\hat{R}_{i,2} = \sqrt{(1-p)/2}|i\rangle\langle i|$ (non-success) for $i = 0,1$, and $\hat{R}_{2,1} = \sqrt{(4-p)/2(3-p)}|2\rangle\langle 2| + |0\rangle\langle 0| + |1\rangle\langle 1|$ (success) and $\hat{R}_{2,2} = \sqrt{(2-p)/2(3-p)}|2\rangle\langle 2|$ (non-success). We obtain $\mathcal{G} = (14 - p)/36$, $\mathcal{F} = (32 - p + 4\sqrt{2p} + 2\sqrt{6 - 2p}\sqrt{4-p})/72$, and $\mathcal{R} = (4+p)/6$. Notably, these saturate $\mathcal{G}$-$\mathcal{R}$ but do not saturate $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ and $\mathcal{G}$-$\mathcal{D}$, i.e., $\{\hat{M}_r\} \in S_{\mathcal{G}\text{-}\mathcal{R}}$, $\{\hat{M}_r\} \notin S_{\mathcal{G}\text{-}\mathcal{D}\text{-}\mathcal{R}}$ and $\{\hat{M}_r\} \notin S_{\mathcal{G}\text{-}\mathcal{D}}$.

(v) Consider a weak quantum measurement with operators $\hat{M}_1 = \sqrt{p}|1\rangle\langle 1|$ and $\hat{M}_2 = |0\rangle\langle 0| + \sqrt{1-p}|1\rangle\langle 1| + |2\rangle\langle 2|$, performed on an arbitrary quantum state $|\psi\rangle$ in 3-dimensional Hilbert space. While the input state $|\psi\rangle$ is completely collapsed on $|1\rangle$ when the outcome $r = 1$, it is partially collapsed when $r = 2$ so that the measurement is reversible for $p < 1$. The optimal reversing operators are given by $\hat{R}_{2,1} = \sqrt{1-p}|0\rangle\langle 0| + |1\rangle\langle 1| + \sqrt{1-p}|2\rangle\langle 2|$ (success) and $\hat{R}_{2,2} = \sqrt{p}|0\rangle\langle 0| + \sqrt{p}|2\rangle\langle 2|$ (non-success). The amount of obtained information is $\mathcal{G} = (4+p)/12$,

while the remaining and the reversible ones are $\mathcal{F} = (2 + \sqrt{1-p})/3$ and $\mathcal{R} = 1 - p$, respectively. While these quantities satisfy all the trade-off relations, they do not saturate any of them.

We have observed that all of the above examples satisfy the trade-off inequalities. The examples (i) - (iv) are optimal quantum measurements that saturate at least one of the trade-off relations. In an optimal quantum measurement, the total information is divided into $\mathcal{G}$, $\mathcal{D}$, and $\mathcal{R}$ and balanced by the change of the parameter p. On the other hand, the example (v) is non-optimal which satisfies all the trade-off inequalities but does not saturate any of them. Each of the examples (i) - (v) given above represents each divided region in the diagram of Fig. 4.

6 Discussion

We have established the complete information balance in quantum measurement by deriving the full quantitative trade-off relations among information gain, disturbance, and reversibility. Under a quantum measurement, the initial information contained in the input state is divided into i) the obtained information $\mathcal{G}$, ii) the remaining information $\mathcal{F}$ in the post-measurement state, and iii) the reversible information $\mathcal{R}$. Our result clearly shows that the three quantities are balanced by the trade-off relations. The reversibility $\mathcal{R}$ turns out to play an essential role in completing the information balance, filling the gap between the information gain $\mathcal{G}$ and disturbance $\mathcal{D} = 1 - \mathcal{F}$. Note that the three information contents are defined to be universal for a given dimension by covering all possible input states and have clear operational meanings, fulfilling a general requirement as an information content in quantum measurement [26].

While all quantum measurements, including noisy or weak measurements, must satisfy the trade-off relations, the conditions to saturate them define optimal quantum measurements. Those optimal measurements may be said to conserve the total information in a nontrivial form according to the trade-off relation. We have classified all quantum measurements into different sets based on their optimality to reach the upper bounds of $\mathcal{G}$, $\mathcal{D}$, and $\mathcal{R}$. We then thoroughly analyzed their relations resulting in the Venn diagram in Fig. 4. These may provide useful guidelines for designing a quantum measurement according to its aim in quantum information protocol. For example, a maximum information gain with minimal disturbance is desirable for estimating or discriminating quantum information [41], whereas a maximum reversibility suits the aim of transmitting [10] or protecting [12] quantum information.

Our results imply that total information does not increase in quantum measurement and reversal process and may hint at the extension of the information conservation law. It should be obviously the case with
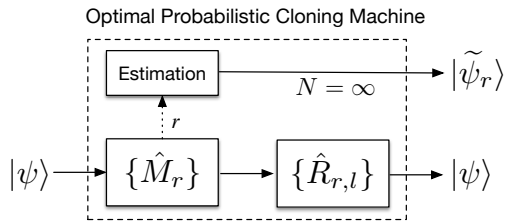
Figure 5: The state estimation by a quantum measurement $\{\hat{M}_r\}$ followed by a reversal process is equivalent to an $1 \rightarrow N+1$ asymmetric probabilistic cloning machine as $N \rightarrow \infty$. The output at the bottom yields exactly the same state as the input state with the probability given by $\mathcal{R}$, while the upper estimation part yields an infinite number of approximated copies with the fidelity given by $\mathcal{G}$.

a unitarity process in line with the second law of thermodynamics [49, 50]. Extending to the non-unitary quantum measurement and selective processes, our optimal measurements saturating trade-off relations may be considered as defining the information conservation.

Our results can also be interpreted as a quantitative refinement of the no-cloning theorem [51] in the context of quantum measurement. As a direct application, we can consider a universal cloning machine, which has been analyzed so far either as a deterministic process with a lower fidelity [52–57] or a probabilistic one for an exact cloning [58, 59]. By contrast, our results allow us to optimize further the protocol by compromising the output fidelity and the success probability of the cloning. For example, consider a $1 \rightarrow N+1$ asymmetric, probabilistic, cloning machine yielding

$$|\psi\rangle \rightarrow \eta_{r,l}|\tilde{\psi}_r\rangle^{\otimes N}|\psi\rangle, \qquad (16)$$

where $|\tilde{\psi}_r\rangle^{\otimes N}$ are the approximate N copies and $|\psi\rangle$ a perfect copy. Employing the optimal measurement and the reversal process can realize this cloning machine when $N \rightarrow \infty$, as illustrated in Fig. 5, with $|\eta_{r,l}|^2$ the success probability, and r and l the outcomes of the measurement and the reversal process, respectively. Compared to the deterministic version [55–57], our results provide a method to optimize the cloning process further with enhanced fidelities at the expense of success probability. Note that the trade-off relations among $\mathcal{G}$, $\mathcal{D}$, and $\mathcal{R}$ determine the quantitative upper bound of the performance of the cloning machine.

An important path for further study is exploring practical applications, e.g. measurement-based quantum processor, teleportation, metrology, and quantum error correction. Characterizing the information flow in sequential quantum measurements would also be interesting, in which the uncertainty relation between incompatible measurements may be crucial. It may also be valuable to translate our results into the context of quantum thermodynamics [60, 61]. As the information contents $\mathcal{G}$, $\mathcal{D}$, and $\mathcal{R}$ are defined as mea-

surable quantities, the derived trade-off relations are ready to be tested in any quantum information platform, e.g., superconducting [37], ion trap [38], and photonic qubits [35, 36, 39, 41, 42]. Our work establishes fundamental criteria for characterizing quantum measurement and offers useful guidelines for designing optimal measurement-based quantum processors.

Acknowledgments

References

[1] W. Heisenberg, *Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik*, Z. Phys. **43**, 172–198 (1927).

[2] M. T. Quintino, T. Vértesi, and N. Brunner, Joint Measurability, Einstein-Podolsky-Rosen Steering, and Bell Nonlocality, Phys. Rev. Lett. **113**, 160402 (2014).

[3] R. Uola, T. Moroder, and O. Gühne, Joint Measurability of Generalized Measurements Implies Classicality, Phys. Rev. Lett. **113**, 160403 (2014).

[4] R. Raussendorf and H. J. Briegel, A One-Way Quantum Computer, Phys. Rev. Lett. **86**, 5188 (2001).

[5] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[6] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, 2009).

[7] K. Jacobs, *Quantum Measurement Theory and its Applications* (Cambridge University Press, 2014).

[8] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels, Phys. Rev. Lett. **70**, 1895 (1993).

[9] H. J. Briegel, D. E. Browne, W. Dür, R. Raussendorf, and M. Van den Nes, Measurement-based quantum computation, Nature Physics **5**, 19–26 (2009).

[10] S. Pirandola, J. Eisert, C. Weedbrook, A. Furusawa, and S. L. Braunstein, Advances in quantum teleportation, Nature Photonics **9**, 641–652 (2015).

[11] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nature Photonics **5**, 222–229 (2011).

[12] D. Lidar and T. Brun, *Quantum Error Correction* (Cambridge University Press, 2013).

[13] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Quantum cryptography, Rev. Mod. Phys. **74**, 145 (2002).

[14] C. H. Bennett and G. Brassard, *Quantum cryptography: Public key distribution and coin tossing*, Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, 175 (1984).

[15] H. J. Groenewold, A problem of information gain by quantal measurements, Int. J. Theoretical Phys. **4**, 327–338 (1971).

[16] G. Lindblad, An entropy inequality for quantum measurements, Commun. Math. Phys. **28**, 245–249 (1972).

[17] M. Ozawa, On information gain by quantum measurements of continuous observables, J. Math. Phys. **27**, 759 (1986).

[18] C. A. Fuchs and A. Peres, Quantum-state disturbance versus information gain: Uncertainty relations for quantum information, Phys. Rev. A **53**, 2038 (1996).

[19] C. A. Fuchs and K. A. Jacobs, Information-tradeoff relations for finite-strength quantum measurements, Phys. Rev. A **63**, 062305 (2001).

[20] K. Banaszek, Fidelity balance in quantum operations, Phys. Rev. Lett. **86**, 1366 (2001).

[21] K. Banaszek and I. Devetak, Fidelity trade-off for finite ensembles of identically prepared qubits, Phys. Rev. A **64**, 052307 (2001).

[22] G. M. D'ariano, On the Heisenberg principle, namely on the information-disturbance trade-off in a quantum measurement, Fortschr. Phys. **51**, 318 (2003).

[23] M. F. Sacchi, Information-disturbance tradeoff in estimating a maximally entangled state, Phys. Rev. Lett. **96**, 220502 (2006).

[24] F. Buscemi, M. Hayashi, and M. Horodecki, Global information balance in quantum measurements, Phys. Rev. Lett. **100**, 210504 (2008).

[25] S. Luo, Information conservation and entropy change in quantum measurements, Phys. Rev. A **82**, 052103 (2010).

[26] M. Berta, J. M. Renes, and M. M. Wilde, Identifying the information gain of a quantum measurement, IEEE Transactions on Information Theory **60**, 7987 (2014).

[27] M. Ueda and M. Kitagawa, Reversibility in quantum measurement processes, Phys. Rev. Lett. **68**, 3424 (1992).

[28] A. Royer, Reversible quantum measurements on a spin 1/2 and measuring the state of a single system, Phys. Rev. Lett. **73**, 913 (1994).

[29] M. Ueda, N. Imoto, and H. Nagaoka, Logical reversibility in quantum measurement: General theory and specific examples, Phys. Rev. A **53**, 3808 (1996).

[30] A. N. Jordan and A. N. Korotkov, Uncollapsing the wavefunction by undoing quantum measurements, Contemporary Physics **51**, 125 (2010).

[31] M. Koashi and M. Ueda, Reversing measurement and probabilistic quantum error correction, Phys. Rev. Lett. **82**, 2598 (1999).

[32] H. Terashima and M. Ueda, Nonunitary quantum circuit, International Journal of Quantum Information **3**, 633 (2005).

[33] A. N. Korotkov and A. N. Jordan, Undoing a weak quantum measurement of a solid-state qubit, Phys. Rev. Lett. **97**, 166805 (2006).

[34] A. N. Korotkov and K. Keane, Decoherence suppression by quantum measurement reversal, Phys. Rev. A **81**, 040103(R) (2010).

[35] Y.-S. Kim, Y.-W. Cho, Y.-S. Ra, and Y.-H. Kim, Reversing the weak quantum measurement for a photonic qubit, Optics Express **17**, 11978 (2009).

[36] Y.-S. Kim, J.-C. Lee, O. Kwon, and Y.-H. Kim, Protecting entanglement from decoherence using weak measurement and quantum measurement reversal, Nature Physics **8**, 117 (2012).

[37] N. Katz, M. Neeley, M. Ansmann, R. C. Bialczak, M. Hofheinz, E. Lucero, A. O'Connell, H. Wang, A. N. Cleland, J. M. Martinis, and A. N. Korotkov, Reversal of the weak measurement of a quantum state in a superconducting phase qubit, Phys. Rev. Lett. **101**, 200401 (2008).

[38] P. Schindler, T. Monz, D. Nigg, J. T. Barreiro, E. A. Martinez, M. F. Brandl, M. Chwalla, M. Hennrich, and R. Blatt, Undoing a Quantum Measurement, Phys. Rev. Lett. **110**, 070403 (2013).

[39] J.-C. Lee, Y.-C. Jeong, Y.-S. Kim, and Y.-H. Kim, Experimental demonstration of decoherence suppression via quantum measurement reversal, Optics Express **19**, 16309 (2011).

[40] Y. W. Cheong and S.-W. Lee, Balance between information gain and reversibility in weak measurement, Phys. Rev. Lett. **109**, 150402 (2012).

[41] H.-T. Lim, Y.-S. Ra, K.-H. Hong, S.-W. Lee, and Y.-H. Kim, Fundamental bounds in measurements for estimating quantum states, Phys. Rev. Lett. **113**, 020504 (2014).

[42] G. Chen, Y. Zou, X.-Y. Xu, J.-S. Tang, Y.-L. Li, J.-S. Xu, Y.-J. Han, C.-F. Li, G.-C. Guo, H.-Q. Ni, Y. Yu, M.-F. Li, G.-W. Zha, Z.-C. Niu, and Y. Kedem, Experimental Test of the State Estimation-Reversal Tradeoff Relation in General Quantum Measurements, Phys. Rev. X **4**, 021043 (2014).

[43] H. Terashima, Information, fidelity, and reversibility in single-qubit measurements, Phys. Rev. A **83**, 032114 (2011).

[44] H. Terashima, Information, fidelity, and reversibility in general quantum measurements, Phys. Rev. A **93**, 022104 (2016).

[45] D. Petz, Sufficient subalgebras and the relative

entropy of states of a von Neumann algebra, Commun. Math. Phys. **105**, 123 (1986).

[46] D. Petz, Sufficiency of channels over von Neumann algebras, The Quarterly Journal of Mathematics **39**, 97(1988).

[47] M. Junge, R. Renner, D. Sutter, M. M. Wilde, and A. Winter, Universal recovery maps and approximate sufficiency of quantum relative entropy, Annales Henri Poincare **19**, 2955 (2018).

[48] F. Sciarrino, M. Ricci, F. De Martini, R. Filip, and L. Mišta, Jr., Realization of a minimal disturbance quantum measurement, Phys. Rev. Lett. **96**, 020408 (2006).

[49] F. Buscemi, S. Das, and M. M. Wilde, Approximate reversibility in the context of entropy gain, information gain, and complete positivity, Phys. Rev. A **93**, 062314 (2016).

[50] H. Kwon and M. S. Kim, Fluctuation theorems for a quantum channel, Phys. Rev. X **9**, 031029 (2019).

[51] W. Wootters and W. Zurek, A Single Quantum Cannot be Cloned, Nature **299**, 802–803 (1982).

[52] V. Bužek and M. Hillery, Quantum copying: Beyond the no-cloning theorem, Phys. Rev. A **54**, 1844 (1996).

[53] N. Gisin and S. Massar, Optimal quantum cloning machines, Phys. Rev. Lett. **79**, 2153 (1997).

[54] V. Scarani, S. Iblisdir, N. Gisin, and A. Acín, Quantum cloning, Rev. Mod. Phys. **77**, 1225 (2005).

[55] D. Bruss, A. Ekert, and C. Macchiavello, Optimal universal quantum cloning and state estimation, Phys. Rev. Lett. **81**, 2598 (1998).

[56] S. Iblisdir, A. Acín, N. J. Cerf, R. Filip, J. Fiurášek, and N. Gisin, Multipartite asymmetric quantum cloning, Phys. Rev. A **72**, 042328 (2005).

[57] J. Bae and A. Acín, Asymptotic quantum cloning is state estimation, Phys. Rev. Lett. **97**, 030402(2006).

[58] L. M. Duan and G. C. Guo, Probabilistic cloning and identification of linearly independent quantum states, Phys. Rev. Lett. **80**, 4999 (1998).

[59] L. M. Duan and G. C. Guo, A probabilistic cloning machine for replicating two non-orthogonal states, Phys. Lett. A. **243**, 261 (1998).

[60] T. Sagawa and M. Ueda, Second law of thermodynamics with discrete quantum feedback control, Phys. Rev. Lett. **100**, 080403 (2008).

[61] K. Jacobs, Second law of thermodynamics and quantum feedback control: Maxwell's demon with weak measurements, Phys. Rev. A **80**, 012322 (2009).

A Definitions

A.1 Singular value decomposition

By a singular value decomposition, a measurement operator $\hat{M}_r$ can be represented as $\hat{M}_r = \hat{V}_r \hat{D}_r \hat{W}_r$ in terms of unitary operators $\hat{W}_r$ and $\hat{V}_r$ and a diagonal matrix $\hat{D}_r = \sum_{i=0}^{d-1} \lambda_i^r |i\rangle\langle i|$. Without loss of generality, we set $\hat{W}_r = \hat{\mathbb{1}}$ and also assume that the singular values are defined here in decreasing order, $\lambda_0^r \geq \lambda_1^r \geq \ldots \geq \lambda_{d-1}^r \geq 0$. Note that the singular values satisfy the completeness relation

$$\sum_r \sum_i (\lambda_i^r)^2 = d. \tag{17}$$

Similarly, the operator of reversing operation can be represented as $\hat{R}_{r,l} = \hat{V}_{r,l} \hat{D}_{r,l} \hat{W}_{r,l}$ with unitary operators $\hat{W}_{r,l}$ and $\hat{V}_{r,l}$ and a diagonal matrix $\hat{D}_{r,l} = \sum_{i=0}^{d-1} \lambda_i^{r,l} |i\rangle\langle i|$. From the completeness relation $\sum_l \hat{R}_{r,l}^\dagger \hat{R}_{r,l} = \hat{\mathbb{1}}$, the singular values for the reversing operators satisfy

$$\sum_l \sum_i (\lambda_i^{r,l})^2 = d. \tag{18}$$

Without loss of generality, we can set $\hat{V}_{r,l} = \hat{\mathbb{1}}$ so that the completeness relation can be written by $\sum_l (\lambda_i^{r,l})^2 = 1$.

A.2 Information Gain

For a given $\mathbf{M} = \{\hat{M}_r\}$, the fidelity between the input and the estimated states $|\langle \psi | \widetilde{\psi}_r \rangle|^2$ can be averaged over all possible input states and the measurement outcomes r as

$$\bar{\mathcal{G}}(\mathbf{M}) = \int d\psi \sum_{r=1}^n p(r, \psi) |\langle \widetilde{\psi}_r | \psi \rangle|^2. \tag{19}$$

Let us introduce the Schur's lemma, which can be used for any operator $\hat{O}$ in $d \times d$ Hilbert space;

$$\int_G dg \left(\hat{U}_g^\dagger \otimes \hat{U}_g^\dagger \right) \hat{O} \left[\hat{U}_g \otimes \hat{U}_g \right] = \alpha_1 \hat{\mathbb{1}} \otimes \hat{\mathbb{1}} + \alpha_2 \hat{S}, \tag{20}$$

$$\alpha_1 = \frac{d^2 \mathrm{Tr}[\hat{O}] - d \mathrm{Tr}[\hat{O}\hat{S}]}{d^2(d^2-1)}, \quad \alpha_2 = \frac{d^2 \mathrm{Tr}[\hat{O}\hat{S}] - d \mathrm{Tr}[\hat{O}]}{d^2(d^2-1)},$$

where $\hat{U}_g$ is a unitary representation of d-dimensional unitary group $g \in G = \mathrm{U}(d)$ such that $\int_G dg = 1$ and $\hat{S}$ is a swap operator defined as $\hat{S}|i\rangle \otimes |j\rangle = |j\rangle \otimes |i\rangle$. Then, Eq. (19) is written as

$$\sum_{r=1}^n \int d\psi \langle \psi | \otimes \langle \psi | \left(\hat{M}_r^\dagger \hat{M}_r \otimes |\widetilde{\psi}_r\rangle\langle\widetilde{\psi}_r| \right) |\psi\rangle \otimes |\psi\rangle,$$

which can be recast by the Schur's lemma into

$$\bar{\mathcal{G}}(\mathbf{M}) = \frac{1}{d(d+1)} \left[d + \sum_{r=1}^n \langle \widetilde{\psi}_r | \hat{M}_r^\dagger \hat{M}_r | \widetilde{\psi}_r \rangle \right]. \tag{21}$$

The second term can be written again as $\langle \widetilde{\psi}_r | \hat{M}_r^\dagger \hat{M}_r | \widetilde{\psi}_r \rangle = \langle \widetilde{\psi}_r | \hat{D}_r^\dagger \hat{V}_r^\dagger \hat{V}_r \hat{D}_r | \widetilde{\psi}_r \rangle = \sum_i \langle \widetilde{\psi}_r | \hat{D}_r^\dagger | i \rangle \langle i | \hat{D}_r | \widetilde{\psi}_r \rangle = \sum_i (\lambda_i^r)^2 |\langle i | \widetilde{\psi}_r \rangle|^2$. Then, the *information gain* can be defined as its maximum value, obtained when the estimate state is $|\widetilde{\psi}_r\rangle = |0\rangle$ for outcome r so that

$$\mathcal{G} \equiv \bar{\mathcal{G}}_{\max}(\mathbf{M}) = \frac{1}{d(d+1)} \left[d + \sum_{r=1}^n (\lambda_0^r)^2 \right]. \tag{22}$$

Note that this is valid for arbitrary input states ρ, since the maximum evaluated in the space of pure states is also the maximum evaluated in the space of mixed states due to the convex structure of quantum states.

A.3 Disturbance

The fidelity between the input and the output states $|\langle\psi_r|\psi\rangle|^2$, for a given quantum measurement $\mathbf{M} = \{\hat{M}_r\}$, can be averaged over all possible input states and measurement outcomes r as

$$\bar{\mathcal{F}}(\mathbf{M}) = \int d\psi \sum_{r=1}^{n} p(r,\psi)|\langle\psi_r|\psi\rangle|^2 = \int d\psi \sum_{r=1}^{n} P(r,\psi)\frac{\langle\psi|\hat{M}_r|\psi\rangle\langle\psi|\hat{M}_r^\dagger|\psi\rangle}{P(r,\psi)} = \int d\psi \sum_{r=1}^{n} \left|\langle\psi|\hat{M}_r|\psi\rangle\right|^2. \quad (23)$$

By the Schur's lemma, we can rewrite it as

$$\bar{\mathcal{F}}(\mathbf{M}) = \frac{1}{d(d+1)}\left[d + \sum_{r=1}^{n}\left|\mathrm{Tr}\hat{M}_r\right|^2\right]. \quad (24)$$

Since $|\mathrm{Tr}\hat{M}_r| = |\sum_i \langle i|\hat{V}_r\hat{D}_r|i\rangle| = |\sum_i \langle i|\hat{V}_r\hat{D}_r|i\rangle| = |\sum_i \lambda_i^r\langle i|\hat{V}_r|i\rangle| \leq \sum_i \lambda_i^r|\langle i|\hat{V}_r|i\rangle| \leq \sum_i \lambda_i^r$, the maximum operation fidelity is then given as

$$\mathcal{F} \equiv \bar{\mathcal{F}}_{max}(\mathbf{M}) = \frac{1}{d(d+1)}\left[d + \sum_{r=1}^{n}\left(\sum_{i=0}^{d-1}\lambda_i^r\right)^2\right], \quad (25)$$

from which the *disturbance* can also be defined as the minimum operation infidelity $\mathcal{D} = 1 - \mathcal{F}$.

A.4 Maximum Operation Fidelity after Reversal

For given $\mathbf{M} = \{\hat{M}_r\}$ and $\mathbf{R} = \{\hat{R}_{r,l}\}$, the operation fidelity including all output states without postselection is

$$\bar{\mathcal{F}}(\mathbf{R}\circ\mathbf{M}) = \int d\psi \sum_{r=1}^{n}\sum_{l=1}^{m} P(r,l,\psi)|\langle\psi_{r,l}|\psi\rangle|^2, \quad (26)$$

where the right hand side can be written as

$$\int d\psi \sum_{r=1}^{n}\sum_{l=1}^{m} P(r,l,\psi)\frac{\langle\psi|\hat{R}_{r,l}\hat{M}_r|\psi\rangle\langle\psi|\hat{M}_r^\dagger\hat{R}_{r,l}^\dagger|\psi\rangle}{P(r,l,\psi)} = \int d\psi \sum_{r=1}^{n}\sum_{l=1}^{m}\left|\langle\psi|\hat{R}_{r,l}\hat{M}_r|\psi\rangle\right|^2. \quad (27)$$

From Eq. (24), it can be written again as

$$\bar{\mathcal{F}}(\mathbf{R}\circ\mathbf{M}) = \frac{1}{d(d+1)}\left[d + \sum_{r=1}^{n}\sum_{l=1}^{m}\left|\mathrm{Tr}\hat{R}_{r,l}\hat{M}_r\right|^2\right]. \quad (28)$$

Since

$$\left|\mathrm{Tr}\hat{R}_{r,l}\hat{M}_r\right| = \left|\sum_i \langle i|\hat{D}_{r,l}\hat{W}_{r,l}\hat{V}_r\hat{D}_r|i\rangle\right| = \left|\sum_i \lambda_i^r\lambda_i^{r,l}\langle i|\hat{W}_{r,l}\hat{V}_r|i\rangle\right| \leq \sum_i \lambda_i^r\lambda_i^{r,l}\left|\langle i|\hat{W}_{r,l}\hat{V}_r|i\rangle\right|, \quad (29)$$

the maximum of $\bar{\mathcal{F}}(\mathbf{R}\circ\mathbf{M})$ is obtained when $\hat{W}_{r,l}\hat{V}_r = \hat{\mathbb{1}}$. Therefore,

$$\bar{\mathcal{F}}_{\max}(\mathbf{R}\circ\mathbf{M}) = \frac{1}{d(d+1)}\left[d + \sum_{r=1}^{n}\sum_{l=1}^{m}\left(\sum_{i=0}^{d-1}\lambda_i^r\lambda_i^{r,l}\right)^2\right]. \quad (30)$$

A.5 Reversibility

We assume that $\hat{R}_{r,l}$ for $l = 1, 2, ..., s$ are associated with the success reversal events such that

$$\hat{R}_{r,l}\hat{M}_r|\psi\rangle = \eta_{r,l}|\psi\rangle, \quad (31)$$

with a complex variable $\eta_{r,l}$. Since $\hat{\mathbb{1}} - \sum_{l=1}^{s}\hat{R}_{r,l}^\dagger\hat{R}_{r,l}$ is positive definite from the completeness relation,

$$\sup_{|\phi\rangle}\langle\phi|\sum_{l=1}^{s}\hat{R}_{r,l}^\dagger\hat{R}_{r,l}|\phi\rangle \leq 1 \quad (32)$$

is satisfied for arbitrary quantum state $|\phi\rangle$. Then,

$$\sup_{|\phi\rangle}\langle\phi|\sum_{l=1}^{s}\hat{R}_{r,l}^{\dagger}\hat{R}_{r,l}|\phi\rangle \geq \sup_{|\psi_r\rangle}\langle\psi_r|\sum_{l=1}^{s}\hat{R}_{r,l}^{\dagger}\hat{R}_{r,l}|\psi_r\rangle = \sup_{|\psi\rangle}\frac{\langle\psi|\hat{M}_r^{\dagger}\sum_{l=1}^{s}\hat{R}_{r,l}^{\dagger}\hat{R}_{r,l}\hat{M}_r|\psi\rangle}{p(r,\psi)} = \frac{\sum_{l=1}^{s}|\eta_{r,l}|^2}{\inf_{|\psi\rangle}p(r,\psi)}, \quad (33)$$

so that $\sum_{l=1}^{s}|\eta_{r,l}|^2 \leq \inf_{|\psi\rangle}p(r,\psi)$ is satisfied. For an arbitrary input state $|\psi\rangle = \sum_{i=0}^{d-1}\alpha_i|i\rangle$ where $\sum_{i=0}^{d-1}|\alpha_i|^2 = 1$, $\inf_{|\psi\rangle}p(r,\psi)$ is obtained when $\alpha_{d-1} = 1$ and all other α_i are zero, because the singular values of $\hat{M}_r$ are assumed to be defined in decreasing order, so that

$$\sum_{l=1}^{s}|\eta_{r,l}|^2 \leq \inf_{|\psi\rangle}p(r,|\psi\rangle) = (\lambda_{d-1}^r)^2. \quad (34)$$

Therefore, the *reversibility* can be defined as its maximum

$$\mathcal{R} \equiv \max_{\{\hat{R}_{r,l}\}}\int d\psi\sum_{r=1}^{n}\sum_{l=1}^{s}|\langle\psi|\hat{R}_{r,l}\hat{M}_r|\psi\rangle|^2 = \sum_{r=1}^{n}(\lambda_{d-1}^r)^2. \quad (35)$$

B Proof of Lemma 1

For an optimally chosen $\mathbf{R} = \{\hat{R}_{r,l}\}$ to reverse $\mathbf{M} = \{\hat{M}_r\}$, the singular values are given by $\lambda_i^{r,l=1} = \lambda_{d-1}^r/\lambda_i^r$. If we define $\vec{u}_i^l = (\lambda_i^{r=1}\lambda_i^{r=1,l},\ldots,\lambda_i^{r=n}\lambda_i^{r=n,l})$, the second term of $\mathcal{F}(\mathbf{R}\circ\mathbf{M})$ in Eq. (4)

$$\sum_{r=1}^{n}\sum_{l=1}^{m}\Big(\sum_{i=0}^{d-1}\lambda_i^r\lambda_i^{r,l}\Big)^2 = \sum_{i,j}\sum_l \vec{u}_i^l\ldots\vec{u}_j^l = \sum_{i,j}(\vec{u}_i^{l=1}\cdot\vec{u}_j^{l=1} + \sum_{l\neq 1}\vec{u}_i^l\cdot\vec{u}_j^l) = d^2\sum_r(\lambda_{d-1}^r)^2 + \sum_{i,j}\sum_{l\neq 1}\vec{u}_i^l\cdot\vec{u}_j^l, \quad (36)$$

and $\sum_{i,j}\sum_{l\neq 1}\vec{u}_i^l\cdot\vec{u}_j^l \geq \sum_i\sum_{l\neq 1}|\vec{u}_i^l|^2$, where the equality is reached with a condition $\vec{u}_i^{l\neq 1}\cdot\vec{u}_j^{l\neq 1} = \delta_{i,j}|\vec{u}_i^l|^2$. From the completeness relation of the overall measurements,

$$\sum_{r=1}^{n}\sum_{l=1}^{m}\sum_{i=0}^{d-1}(\lambda_i^r\lambda_i^{r,l})^2 = d\sum_r(\lambda_{d-1}^r)^2 + \sum_i\sum_{l\neq 1}|\vec{u}_i^l|^2 = d.$$

Therefore,

$$\mathcal{F}(\mathbf{R}\circ\mathbf{M}) = \frac{1}{d(d+1)}\Big[d + d^2\sum_r(\lambda_{d-1}^r)^2 + \sum_{i,j}\sum_{l\neq 1}\vec{u}_i^l\cdot\vec{u}_j^l\Big]$$

$$\geq \frac{1}{d(d+1)}\big[2d + (d^2-d)\mathcal{R}\big], \quad (37)$$

where $\mathcal{R} = \sum_r(\lambda_{d-1}^r)^2$. As a result, we obtain

$$2 + (d-1)\mathcal{R} \leq (d+1)\mathcal{F}(\mathbf{R}\circ\mathbf{M}), \quad (38)$$

with the equality condition $\vec{u}_i^{l\neq 1}\cdot\vec{u}_j^{l\neq 1} = \delta_{i,j}|\vec{u}_i^l|^2$. $\square$

C Proof of Theorem 1

Let us first define $\vec{v}_i = (\lambda_i^{r=1},\ldots,\lambda_i^{r=n})$. Then, $g \equiv d(d+1)\mathcal{G} - d = \sum_r(\lambda_0^r)^2 = |\vec{v}_0|^2$, $f \equiv d(d+1)\mathcal{F} - d = \sum_r(\sum_i\lambda_i^r)^2 = \sum_{i,j}\vec{v}_i\cdot\vec{v}_j$, and $\mathcal{R} = \sum_r(\lambda_{d-1}^r)^2 = |\vec{v}_{d-1}|^2$. The completeness relation in Eq. (17) can be written as $\sum_r\sum_i(\lambda_i^r)^2 = \sum_i|\vec{v}_i|^2 = d$.

From the Schwarz inequality,

$$f \leq \sum_{i,j=0}^{d-1}|\vec{v}_i||\vec{v}_j| = \Big(\sum_i|\vec{v}_i|\Big)^2 = \Big(\sqrt{g} + \sqrt{\mathcal{R}} + \sum_{i=1}^{d-2}|\vec{v}_i|\Big)^2, \quad (39)$$

where the equality can be reached when all the vectors $\vec{v}_i$ are collinear. Then, from the inequality of arithmetic and quadratic means,

$$\sum_{i=1}^{d-2} |\vec{v}_i| \leq \sqrt{(d-2) \sum_{i=1}^{d-2} |\vec{v}_i|^2}, \tag{40}$$

where the equality can be reached when $|\vec{v}_1| = \cdots = |\vec{v}_{d-2}|$. Here, the right hand side can be rewritten by the completeness relation as $\sqrt{(d-2)(d-g-\mathcal{R})}$. Therefore, we obtain

$$\sqrt{f} \leq \sqrt{g} + \sqrt{\mathcal{R}} + \sqrt{(d-2)(d-g-\mathcal{R})}, \tag{41}$$

and equivalently

$$\sqrt{\mathcal{F} - \frac{1}{d+1}} \leq \sqrt{\mathcal{G} - \frac{1}{d+1}} + \sqrt{\frac{\mathcal{R}}{d(d+1)}} + \sqrt{(d-2)\left(\frac{2}{d+1} - \mathcal{G} - \frac{\mathcal{R}}{d(d+1)}\right)}. \tag{42}$$

$\square$

D Proof that $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ tightens $\mathcal{G}$-$\mathcal{D}$

Assume that we have $d-1$ non-negative real numbers x_i where $i = 1, \cdots, d-1$. From the inequality between arithmetic and quadratic mean,

$$\left(\frac{1}{d-1} \sum_{i=1}^{d-1} x_i\right)^2 \leq \frac{1}{d-1} \sum_{i=1}^{d-1} x_i^2. \tag{43}$$

where the equality holds if and only if $x_1 = x_2 = \cdots = x_{d-1}$. By letting $x_1 = \sqrt{\mathcal{R}}$ and $x_2 = \cdots = x_{d-1} = \sqrt{(d-g-\mathcal{R})/(d-2)}$ with $g \equiv d(d+1)\mathcal{G} - d$, we obtain

$$\sqrt{\mathcal{R}} + \sqrt{(d-2)(d-g-\mathcal{R})} \leq \sqrt{(d-1)(d-g)}, \tag{44}$$

which indicates that $\mathcal{G}$-$\mathcal{D}$-$\mathcal{R}$ is tighter than $\mathcal{G}$-$\mathcal{D}$. $\square$

E Proof of Lemma 2

If we define a vector $\vec{w}_i^r = (\lambda_i^r \lambda_i^{r,l=1}, \ldots, \lambda_i^r \lambda_i^{r,l=m})$, the second term of $\mathcal{F}(\mathbf{R} \circ \mathbf{M})$ in Eq. (4) can be written by

$$\sum_{r=1}^{n} \sum_{l=1}^{m} \left(\sum_{i=0}^{d-1} \lambda_i^r \lambda_i^{r,l}\right)^2 = \sum_{r=1}^{n} \sum_{i,j} \vec{w}_i^r \cdot \vec{w}_j^r. \tag{45}$$

Then, by the Schwarz inequality,

$$\sum_{r=1}^{n} \sum_{i,j} \vec{w}_i^r \cdot \vec{w}_j^r \leq \sum_{r=1}^{n} \sum_{i,j} |\vec{w}_i^r||\vec{w}_j^r| = \sum_{r=1}^{n} \left(\sum_i |\vec{w}_i^r|\right)^2 = \sum_{r=1}^{n} \left(\sum_i \lambda_i^r \sqrt{\sum_l (\lambda_i^{r,l})^2}\right)^2, \tag{46}$$

where the equality holds if all the vectors $\vec{w}_i^r$ are collinear. From the completeness relation for the reversing operators in Eq. (18), we arrive at

$$\sum_{r=1}^{n} \sum_{l=1}^{m} \left(\sum_{i=0}^{d-1} \lambda_i^r \lambda_i^{r,l}\right)^2 \leq \sum_{r=1}^{n} \left(\sum_{i=0}^{d-1} \lambda_i^r\right)^2, \tag{47}$$

and, from Eq. (2),

$$\mathcal{F}(\mathbf{R} \circ \mathbf{M}) \leq \mathcal{F}(\mathbf{M}). \tag{48}$$

$\square$

F   Reversibility with errors

Using a non-optimal quantum measurement or operation, the reversing process may not be able to exactly reverse the quantum measurement due to errors so that the output state is not exactly the same as the input state, s.t. $(\mathbf{R} \circ \mathbf{M})(\rho) \propto \rho' \neq \rho$. Let us consider the case that the reversing operation succeeds after measurement as

$$\hat{R}_{r,1} \hat{M}_r |\psi\rangle \langle\psi| \hat{M}_r^\dagger \hat{R}_{r,1}^\dagger = |\eta_{r,1}|^2 \rho(\psi, \epsilon) \tag{49}$$

but the resulting output state $\rho(\psi, \epsilon)$ deviates from the original input state. Here, we consider an arbitrary error model parameterized by $0 \leq \epsilon \leq 1$ and the error density function $p(\epsilon)$, s.t. $\int_0^1 p(\epsilon) d\epsilon = 1$. We assume that the error occurs independently of the measurement outcome for simplicity.

From the definition of the reversibility in Eq. (5), we can evaluate

$$\begin{aligned}
\mathcal{R}(\epsilon) &= \max_{\{\hat{R}_{r,1}\}} \int d\psi \sum_r \langle\psi|\hat{R}_{r,1} \hat{M}_r|\psi\rangle \langle\psi|\hat{M}_r^\dagger \hat{R}_{r,1}^\dagger|\psi\rangle \\
&= \max_{\{\hat{R}_{r,1}\}} \int d\psi \sum_r |\eta_{r,1}|^2 \langle\psi|\rho(\psi, \epsilon)|\psi\rangle \\
&= \sum_r (\lambda_{d-1}^r)^2 \int d\psi \langle\psi|\rho(\psi, \epsilon)|\psi\rangle.
\end{aligned} \tag{50}$$

By averaging over the whole error region, we obtain

$$\mathcal{R} = \int_0^1 d\epsilon \, p(\epsilon)\mathcal{R}(\epsilon) = \sum_r (\lambda_{d-1}^r)^2 F_s, \tag{51}$$

where $F_s = \int d\psi \int_0^1 d\epsilon \, p(\epsilon)\langle\psi|\rho(\psi, \epsilon)|\psi\rangle$ is defined as the average fidelity between the input state and the output state when the reversing operation succeeds.

G   Proof of $\mathcal{S}_{\mathcal{G}\text{-}\mathcal{D}\text{-}\mathcal{R}} \cap \mathcal{S}_{\mathcal{G}\text{-}\mathcal{R}} = \mathcal{S}_{\mathcal{G}\text{-}\mathcal{D}}$

We can prove that a quantum measurement satisfying both saturation conditions of $\mathcal{G}\text{-}\mathcal{D}\text{-}\mathcal{R}$ and $\mathcal{G}\text{-}\mathcal{R}$ always saturates $\mathcal{G}\text{-}\mathcal{D}$. From the condition to saturate $\mathcal{G}\text{-}\mathcal{D}\text{-}\mathcal{R}$ described in Eq. (11), the singular values satisfy $\lambda_0^r \geq \lambda_1^r = \cdots = \lambda_{d-2}^r \geq \lambda_{d-1}^r$. The measurement operator is then written by $\hat{M}_r = \lambda_0^r |0\rangle\langle 0| + \lambda_1^r (|1\rangle\langle 1| + \cdots + |d-2\rangle\langle d-2|) + \lambda_{d-1}^r |d-1\rangle\langle d-1|$ so that

$$\hat{M}_r^\dagger \hat{M}_r = \{(\lambda_0^r)^2 - (\lambda_{d-1}^r)^2\}|0\rangle\langle 0| + \{(\lambda_1^r)^2 - (\lambda_{d-1}^r)^2\}(|1\rangle\langle 1| + \cdots + |d-2\rangle\langle d-2|) + (\lambda_{d-1}^r)^2 \hat{\mathbb{1}}. \tag{52}$$

If it satisfies the saturation condition of $\mathcal{G}\text{-}\mathcal{R}$, i.e., $\hat{M}_r^\dagger \hat{M}_r = a_r |j_r\rangle\langle j_r| + b_r \hat{\mathbb{1}}$ with non-negative a_r and b_r where $j_r \in \{0, \ldots, d-1\}$ [40], we find that

$$\lambda_1^r = \lambda_{d-1}^r. \tag{53}$$

As a result, the quantum measurement satisfies the saturation condition of $\mathcal{G}\text{-}\mathcal{D}$, i.e., all $\vec{v}_i$ are collinear and $|\vec{v}_1| = \cdots = |\vec{v}_{d-1}|$, where $\vec{v}_i = (\lambda_i^{r=1}, \ldots, \lambda_i^{r=n})$ [20]. $\square$