Balance Between Information Gain and Reversibility in Weak Measurement

Yong Wook Cheong[1] and Seung-Woo Lee[2,*]

[1]Gyeonggi Science High School, Suwon, Gyeonggi-Do 440-210, Korea

[2]Center for Macroscopic Quantum Control, Department of Physics and Astronomy, Seoul National University, Seoul, 151-742, Korea

(Received 26 March 2012; revised manuscript received 4 July 2012; published 9 October 2012)

We derive a tight bound between the quality of estimating a quantum state by measurement and the success probability of undoing the measurement in arbitrary dimensional systems, which completely describes the tradeoff relation between the information gain and reversibility. In this formulation, it is clearly shown that the information extracted from a weak measurement is erased through the reversing process. Our result broadens the information-theoretic perspective on quantum measurement as well as provides a standard tool to characterize weak measurements and reversals.

DOI: 10.1103/PhysRevLett.109.150402 PACS numbers: 03.65.Ta, 03.67.Mn

Since Heisenberg discussed the γ-ray microscope gedanken experiment [1], the disturbance induced by measurement has become one of the fundamental issues in quantum mechanics. A heuristic statement, 'the more information is obtained from a quantum system, the more its state is disturbed by measurement' is widely believed nowadays, and numerous efforts have been devoted to proving this in a quantitative manner [2–6].

However, the general belief in the irreversibility of quantum measurement has been shown to be not always true in the sense that the input state can be retrieved with a nonzero success probability by reversing operations on the post-measurement state [7,8]. This is because the quantum state is not fully perturbed by measurement when the interaction between the system and measurement apparatus is weak. The measurement that induces a partial collapse of a quantum state is called a "weak measurement," and its reversing process has been studied theoretically [9–11] and realized experimentally [12,13]. It has attracted much attention due to its potential applications in quantum information processing [14,15].

In an information-theoretic point of view, reversibility can be understood as a degree of preserved information in the measurement process, and thus should be quantitatively related to the extracted one [16]. In fact, the information that is not extracted from a measurement is transferred to the remainder of the whole Hilbert space describing the measurement process (see the Appendix). Even after some information is extracted through a measurement, its state can be retrieved with a probability equal to the degree of ignorance of the state [8]. This concept was also proved in another context as the 'no-hiding theorem' of information [17]. In this sense, the extracted information should be more tightly related to the possibility of undoing the measurement [4] rather than the closeness between input and post-measurement states as used in previous works [2]. Recently, an entropic tradeoff relation was derived based on the concept of information conservation in the measurement process [18], and a degree of information gain was investigated by changing the reversibility in a single measurement outcome level [19]. However, a clear and direct quantitative relation between information gain and reversibility in quantum measurement has so far been missing.

In this Letter, we derive a tight bound between the amount of information gain and reversibility in arbitrary d-level systems, which are quantified by the average estimation fidelity [2] and the reversal probability [8], respectively. In particular, it shows a sharp tradeoff relation between them with a monotonic equation for qubit (2-level) systems. To our knowledge, this is the first direct and quantitative link between information gain and reversibility. Moreover, since both the estimation fidelity [5,20,21] and reversal probability [8,11–13] are measurable quantities, its demonstration is experimentally feasible. Our result provides a fundamental insight on the quantum measurement as well as a useful tool to characterize reversals of weak measurements potentially used in quantum information processing [14,15].

Quantum measurement—An ideal measurement can be described by a set of operators $\{\hat{A}_r | r = 1, \ldots, N\}$, satisfying the completeness relation

$$\sum_{r=1}^{N} \hat{A}_r^\dagger \hat{A}_r = \hat{\mathbb{1}}, \tag{1}$$

where the index r indicates the obtained classical information. A measurement performed on a system transforms its input state $|\psi\rangle$ to

$$|\psi_r\rangle = \frac{\hat{A}_r |\psi\rangle}{\sqrt{p(r, |\psi\rangle)}}, \tag{2}$$

which is the *post-measurement* state, where $p(r, |\psi\rangle) = \langle\psi|\hat{A}_r^\dagger \hat{A}_r|\psi\rangle$ is the probability that the outcome is r.

A measurement operator $\hat{A}_r$ can be written by the singular-value decomposition: $\hat{A}_r = \hat{V}_r \hat{D}_r \hat{U}_r$, where $\hat{U}_r$ and $\hat{V}_r$ are unitary operators, and $\hat{D}_r$ is a diagonal matrix with non-negative entries. We assume $\hat{V}_r = \mathbb{1}$

without loss of generality, and $\hat{U}_r$ can be written by $\hat{U}_r = \sum_{i=0}^{d-1} |v_i^r\rangle\langle w_i^r|$ with two orthonormal bases $\{|v_i^r\rangle | i = 0, \ldots, d-1\}$ and $\{|w_i^r\rangle | i = 0, \ldots, d-1\}$ for d-level measurements. The diagonal matrix can be also written by $\hat{D}_r = \sum_{i=0}^{d-1} \lambda_i^r |v_i^r\rangle\langle v_i^r|$, with non-negative diagonal elements λ_i^r (i.e., singular values) put in decreasing order such that $\lambda_0^r \geq \lambda_1^r \geq \ldots \geq \lambda_{d-1}^r$. Thus, each measurement operator can be represented by

$$\hat{A}_r = \sum_{i=0}^{d-1} \lambda_i^r |v_i^r\rangle\langle w_i^r|, \tag{3}$$

and due to the completeness relation in Eq. (1) their singular values λ_i^r satisfy

$$\sum_{r=1}^{N} \sum_{i=0}^{d-1} (\lambda_i^r)^2 = d. \tag{4}$$

Information gain—In order to quantify the obtained information through a measurement, we employ the estimation fidelity [2]. When the measured outcome is r, one can make a guess on the input state $|\psi\rangle$ and select a state $|\widetilde{\psi_r}\rangle$. The quality of the guess can be quantified with the help of overlap between them $|\langle\psi|\widetilde{\psi_r}\rangle|^2$. Then, the mean estimation fidelity is obtained by averaging $|\langle\psi|\widetilde{\psi_r}\rangle|^2$ over all possible measurement outcomes r and input states $|\psi\rangle$:

$$G = \int d\psi \sum_{r=1}^{N} p(r, |\psi\rangle) |\langle\widetilde{\psi_r}|\psi\rangle|^2, \tag{5}$$

which gives different values depending on the guess strategy. We reformulate it by

$$\sum_{r=1}^{N} \int d\psi \langle\psi| \otimes \langle\psi| (\hat{A}_r^\dagger \hat{A}_r \otimes |\widetilde{\psi_r}\rangle\langle\widetilde{\psi_r}|) |\psi\rangle \otimes |\psi\rangle, \tag{6}$$

and use the Schur's lemma [22] that leads to the identity,

$$\begin{aligned} &\int_G dg [\hat{U}^\dagger(g) \otimes \hat{U}^\dagger(g)] \hat{O} [\hat{U}(g) \otimes \hat{U}(g)] \\ &\quad = \alpha_1 \hat{\mathbb{1}} \otimes \hat{\mathbb{1}} + \alpha_2 \hat{S}, \\ \alpha_1 &= \frac{d^2 \mathrm{Tr}(\hat{O}) - d\mathrm{Tr}(\hat{O}\hat{S})}{d^2(d^2-1)}, \\ \alpha_2 &= \frac{d^2 \mathrm{Tr}(\hat{O}\hat{S}) - d\mathrm{Tr}(\hat{O})}{d^2(d^2-1)}, \end{aligned}$$

for any operator $\hat{O}$ acting on the $d \times d$ Hilbert space. Here dg is Haar invariant measure on the d-dimensional unitary group $G = U(d)$ such that $\int_G dg = 1$, $\hat{U}(g)$ is an irreducible unitary representation of $g \in G$, and $\hat{S}$ is a swap operator defined as $\hat{S}|i\rangle \otimes |j\rangle = |j\rangle \otimes |i\rangle$. A simpler form is then obtained as

$$\frac{1}{d(d+1)} \left(d + \sum_{r=1}^{N} \mathrm{Tr}[\hat{A}_r^\dagger \hat{A}_r \otimes |\widetilde{\psi_r}\rangle\langle\widetilde{\psi_r}| \hat{S}] \right), \tag{7}$$

and by using Eq. (3), its second term is rewritten by

$$\sum_{r=1}^{N} \sum_{i=0}^{d-1} (\lambda_i^r)^2 |\langle\widetilde{\psi_r}|w_i^r\rangle|^2, \tag{8}$$

which gives a maximum value when the estimated state $|\widetilde{\psi_r}\rangle$ is equivalent to $|w_0^r\rangle$. Then, we define the measure of *information gain* as the maximal value of the mean estimation fidelity,

$$G_{\max} = \frac{1}{d(d+1)} \left[d + \sum_{r=1}^{N} (\lambda_0^r)^2 \right], \tag{9}$$

which is a function of the maximal singular value λ_0^r of the measurement operators. Note that it is scaled in the range $1/d \leq G_{\max} \leq 2/(d+1)$, where the upper bound $2/(d+1)$ is reachable by a von Neumann measurement and the lower bound $1/d$ is obtained by a unitary measurement or equivalently by a random guess. The result in Eq. (9) is valid for arbitrary input states $\hat{\rho}$ as a mixed state degrades the estimation fidelity by averaging over the input probability so that its maximum is always obtained in the space of pure states.

Reversibility—A reversing operator $\hat{R}^{(r)}$ can be defined for a physically reversible measurement $\hat{A}_r$ [8] to recover the input state as $\hat{R}^{(r)}|\psi_r\rangle \propto |\psi\rangle$. Thus, a subsequent measurement of reversing operator $\hat{R}^{(r)}$ after the first measurement $\hat{A}_r$ leads to a successful reversal, independently on the input state $|\psi\rangle$, as

$$\hat{R}^{(r)} \hat{A}_r |\psi\rangle = \eta_r |\psi\rangle, \tag{10}$$

where η_r is a nonzero complex number.

Since $\hat{R}^{(r)}$ can be regarded as an element of a complete measurement set, $\mathbb{1} - \hat{R}^{(r)\dagger}\hat{R}^{(r)}$ is positive, semidefinite, and equivalently,

$$\sup_{|\phi\rangle} \langle\phi|\hat{R}^{(r)\dagger}\hat{R}^{(r)}|\phi\rangle \leq 1, \tag{11}$$

for the arbitrary (normalized) quantum state $|\phi\rangle$. Simultaneously [8],

$$\begin{aligned} \sup_{|\phi\rangle} \langle\phi|\hat{R}^{(r)\dagger}\hat{R}^{(r)}|\phi\rangle &\geq \sup_{|\psi_r\rangle} \langle\psi_r|\hat{R}^{(r)\dagger}\hat{R}^{(r)}|\psi_r\rangle \\ &= \sup_{|\psi\rangle} \frac{\langle\psi|\hat{A}_r^\dagger \hat{R}^{(r)\dagger}\hat{R}^{(r)}\hat{A}_r|\psi\rangle}{p(r, |\psi\rangle)} \\ &= \frac{|\eta_r|^2}{\inf_{|\psi\rangle} p(r, |\psi\rangle)}, \end{aligned} \tag{12}$$

so that $|\eta_r|^2 \leq \inf_{|\psi\rangle} p(r, |\psi\rangle)$ is satisfied. As the input state can be written with an arbitrary orthonormal basis $\{|w_i\rangle | i = 0, \ldots, d-1\}$ as $|\psi\rangle = \sum_{i=0}^{d-1} \alpha_i |w_i\rangle$ where

$\sum_{i=0}^{d-1}|\alpha_i|^2 = 1$ and the singular values are defined in decreasing order,

$$|\eta_r|^2 \leq \inf_{|\psi\rangle} p(r, |\psi\rangle) = \inf_{\{\alpha_i\}} |\alpha_i \lambda_i^r|^2 = (\lambda_{d-1}^r)^2, \quad (13)$$

is obtained when $\alpha_{d-1} = 1$ and all other α_i are zero.

Therefore, the reversal probability for each measurement outcome r has the upper limit as

$$P_{\rm rev}(r) = |\langle\psi|\hat{R}^{(r)}|\psi\rangle_r|^2 = \frac{|\eta_r|^2}{P(r, |\psi\rangle)} \leq \frac{(\lambda_{d-1}^r)^2}{P(r, |\psi\rangle)}. \quad (14)$$

We then define the *reversibility* as the maximal mean value of reversal probability over all the outcomes r [10],

$$P_{\rm rev} = \max \sum_{r=1}^{N} P_{\rm rev}(r) P(r, |\psi\rangle) = \sum_{r=1}^{N} (\lambda_{d-1}^r)^2, \quad (15)$$

which notably does not depend on the input state $|\psi\rangle$ but is given as a function of the minimal singular value of measurement operators, λ_{d-1}^r. Its maximum value $P_{\rm rev} = 1$ is obtained by a unitary measurement, meaning that the input state can be deterministically retrieved with appropriate reversing unitary operation, while the minimum value $P_{\rm rev} = 0$ is given by a von Neumann measurement, implying that full extraction of information frustrates the reversing process.

Assuming arbitrary mixed input states $\hat{\rho}$, we can obtain the same reversibility with the form in Eq. (15) [10] as $\inf_{\hat{\rho}} p(r, \hat{\rho}) = \inf_{\hat{\rho}} {\rm Tr}[\hat{\rho}\hat{A}_r^\dagger \hat{A}_r] = (\lambda_{d-1}^r)^2$ and $P_{\rm rev}(r) \leq (\lambda_{d-1}^r)^2 / P(r, \hat{\rho})$ so that $P_{\rm rev} = \sum_{r=1}^{N} (\lambda_{d-1}^r)^2$. Therefore, the result in Eq. (15) is valid for arbitrary input states.

It may be considerable to quantify the disturbance of quantum states by using the reversibility of measurement. For instance, we can define a measure of disturbance by the quantity $1 - P_{\rm rev}$. It shows that the higher the reversal probability is, the less the state is disturbed, which satisfies the requirements for measures of state disturbance listed in Ref. [4].

Tradeoff relation—We now derive a tradeoff relation between the information gain and reversibility from the representation obtained above. An inequality

$$\sum_{r=1}^{N} \{(\lambda_0^r)^2 + (d-1)(\lambda_{d-1}^r)^2\} \leq d, \quad (16)$$

is derived from the completeness relation in Eq. (4) and the nonincreasing order of the singular values ($\lambda_0^r \geq \lambda_1^r \geq \ldots \geq \lambda_{d-1}^r$). From Eqs. (9), (15), and (16), we can finally obtain a bound inequality for $G_{\rm max}$ and $P_{\rm rev}$ as

$$d(d+1)G_{\rm max} + (d-1)P_{\rm rev} \leq 2d, \quad (17)$$

where $1/d \leq G_{\rm max} \leq 2/(d+1)$, which is the main result of this Letter, showing a tradeoff relation between information gain and reversibility.

We can find a measurement that is maximally reversible for a fixed amount of information gain, which saturates the inequality in Eq. (17). The necessary and sufficient condition to reach the equality sign is that each measurement operator has the form satisfying $\hat{A}_r^\dagger \hat{A}_r = a_r |w_0^r\rangle\langle w_0^r| + b_r \hat{\mathbb{1}}$ for certain nonnegative parameters a_r and b_r. It is thus guaranteed that the inequality in Eq. (17) is tight and can not be further improved. Interestingly, the maximal reversibility in our result does not necessarily correspond to the minimal disturbance, which is defined by the closeness of the transformed state from the input state $\int d\psi \sum_{r=1}^{N} |\langle\psi|\hat{A}_r|\psi\rangle|^2$ [2], while the converse is true. This implies that our tradeoff relation differs from the one proposed by Banaszek [2].

For qubit (2-level) systems, a particulary interesting tradeoff relation is obtained. In this case, the inequality of Eq. (17) is reduced to a monotonic equation

$$6G_{\rm max} + P_{\rm rev} = 4, \quad (18)$$

where $1/2 \leq G_{\rm max} \leq 2/3$. We emphasize that $G_{\rm max}$ and $P_{\rm rev}$ for any ideal measurement should satisfy this equation. Therefore, we come to a heuristic statement about quantum measurement 'the more information is obtained from a quantum system, the less possible it is to retrieve the input state of the system'.

Erasing information—The tradeoff relation in (17) and (18) implicate the possibility of erasing information by reversing operation. One may ask whether it is possible to erase the information already obtained and possibly recorded somewhere else. The answer is 'yes' for any partial information obtained by weak measurement, while any full information by von Neumann measurement is not erasable. In order to describe the erasing process, we will consider two weak measurements, saying $\{\hat{A}_r\}$ and $\{\hat{B}_\mu\}$, performed one after the other on an unknown system. Then, the erasure of information is simply understood as a collection of the opposite information by $\{\hat{B}_\mu\}$ that makes the information already obtained by $\{\hat{A}_r\}$ less certain [10].

Let us assume that one element of the second measurement set is given by $\hat{B}_1 = \hat{R}^{(r)}$. If the results of two measurements are given in turn as r and 1, the total measurement operation performed on the state is described by $\hat{B}_1\hat{A}_r = \hat{R}^{(r)}\hat{A}_r$. From Eq. (10), it satisfies $\hat{R}^{(r)}\hat{A}_r|\psi\rangle = \eta_r|\psi\rangle$ independently on the input state $|\psi\rangle$, meaning that no information is obtained about the state. Therefore, we conclude that the information obtained through a measurement is erased by its reversal.

Since a measurement operator $\hat{A}_r$ is decomposable into $\hat{A}_r = \hat{D}_r\hat{U}_r$, its optimal reversing operation is given from Eq. (10) as $\hat{R}^{(r)} = \eta_r \hat{U}_r^\dagger \hat{D}_r^{-1}$ where $\hat{U}_r^\dagger = \sum_{i=0}^{d-1} |w_i^r\rangle\langle v_i^r|$ and $\hat{D}_r^{-1} = \sum_{i=0}^{d-1} \frac{1}{\lambda_i^r} |w_i^r\rangle\langle w_i^r|$, with an assumption that each λ_i^r is nonzero. Then, we can define the *erasing operator* for an arbitrary measurement operator $\hat{A}_r$ as

$$\hat{E}^{(r)} = \lambda_{d-1}^r \hat{D}_r^{-1} = \sum_{i=0}^{d-1} \frac{\lambda_{d-1}^r}{\lambda_i^r} |w_i^r\rangle\langle w_i^r|. \quad (19)$$

It transforms the post-measurement state $|\psi_r\rangle$ to

$$\hat{E}^{(r)}|\psi_r\rangle = \sqrt{P_{er}}\hat{U}_r|\psi\rangle, \tag{20}$$

where $P_{er} = \langle\psi_r|\hat{E}^{(r)}\hat{E}^{(r)}|\psi_r\rangle = (\lambda^r_{d-1})^2/P(r,|\psi\rangle)$, from which the input state $|\psi\rangle$ can be retrieved deterministically by unitary operation [16], meaning that at this stage the information obtained by $\{\hat{A}_r\}$ is erased.

Examples—(i) Assume the case when a von Neumann measurement with two operators $\hat{A}_1 = |0\rangle\langle 0|$ and $\hat{A}_2 = |1\rangle\langle 1|$ is performed on an arbitrary qubit. Then, the degree of information gain has the maximal value $G_{\max} = 2/3$ with a zero reversibility ($P_{\rm rev} = 0$) irrespectively on the input state. It shows that the von Neumann measurement can not be reversed in any case (the information can not be erased).

(ii) Consider a weak measurement described by two operators $\hat{A}_1 = \sqrt{\eta}|1\rangle\langle 1|$ and $\hat{A}_2 = |0\rangle\langle 0| + \sqrt{1-\eta}|1\rangle\langle 1|$ where η is defined as the probability of detecting $|1\rangle$ state (as implemented in Ref. [13]). If the measurement outcome is $r = 1$ the state collapses on the state $|1\rangle$, while when $r = 2$ the input state collapses partially and can be retrieved. The degree of information gain is $G_{\max} = (3+\eta)/6$ and the reversibility is $P_{\rm rev} = 1-\eta$, satisfying the tradeoff relation (18).

The information obtained by this measurement can be erased by properly choosing another measurement. From Eq. (19), the erasing operator for $\hat{A}_2$ [where $\hat{D}^{-1} = |0\rangle\langle 0| + (1/\sqrt{1-\eta})|1\rangle\langle 1|$ and $\lambda_1^2 = \sqrt{1-\eta}$] is given as

$$\hat{E}^{(2)} = \lambda_1^2\hat{D}^{-1} = \sqrt{1-\eta}|0\rangle\langle 0| + |1\rangle\langle 1|. \tag{21}$$

A measurement $\{\hat{B}_\mu|\mu = 1, 2\}$ can be then defined with two operators $\hat{B}_1 = \hat{E}^{(2)}$ and $\hat{B}_2 = \sqrt{\eta}|0\rangle\langle 0|$, satisfying the completeness relation $\hat{B}_1^2 + \hat{B}_2^2 = \mathbb{1}$. Thus, the information extracted from the result $r = 2$ of the first measurement $\{\hat{A}_r\}$ is erased probabilistically by the subsequent measurement $\{\hat{B}_\mu\}$ when its outcome is $\mu = 1$, since the result of the second measurement makes the information obtained from the first measurement uncertain. Our formalism is generally applicable to any examples of weak measurements and reversals in Refs. [9–15].

Remarks—Our result provides a useful framework for generalizing the quantum teleportation [23]. Suppose that Alice performs a joint measurement (assumed here as a projection $\{|w^r\rangle_{ab}\langle w^r|\}$ for simplicity) on an unknown input $|\psi\rangle_a$ and one party of an entangled channel $|\Psi\rangle_{bc}$. Here a, b, and c denote the input, Alice's, and Bob's modes, respectively. The teleportation can then be described as a reversible measurement with operators $\{_{ab}\langle w^r|\Psi\rangle_{bc}\}$ performed on $|\psi\rangle_a$ so that, based on our formalism, the extracted information of $|\psi\rangle_a$ during the teleportation and its reversibility are certainly in the tradeoff relation. As the reversibility here indicates the success probability of teleportation, the result is rephrased as 'the less information about the input state is disclosed during the teleportation, the higher the teleportation probability.' For example, the standard teleportation [23] is deterministic as Alice cannot obtain any information of $|\psi\rangle_a$ by the Bell measurement with a maximally entangled channel. Within this framework, various tasks of quantum transmission (e.g., from the teleportations using nonmaximally entangled or nonorthogonal measurements with arbitrary entangled channels to the communications in quantum networks [24]) can be characterized. The detailed analysis of generalized teleportation will be presented elsewhere.

Obviously our result manifests the quantum no-cloning theorem in information-theoretic perspective [25], as a perfect copy of a quantum state would violate the bound in Eq. (17), which is a crucial ingredient of quantum cryptography [26]. Another implication of our result is that the success rate of quantum error correction should be bounded by the amount of information loss in the qubit [8], which may lead to further applications in quantum computation.

In summary, we derive a tradeoff relation between the degree of information gain and reversibility in arbitrary-dimensional quantum measurement. It quantitatively shows that 'the more information is obtained from quantum measurement, the less possible it is to undo the measurement.' Simultaneously, it is clearly shown that undoing a quantum measurement erases the same amount of information obtained by the measurement. Our result, as providing an information-theoretic insight on quantum measurement, is expected to widen the potential applications of weak measurements and reversals in quantum information processing.

We thank J. Lee and H.-W. Lee for discussions. Y. W. C. acknowledge supports from the National Research Foundation of Korea Grant funded by the Korea government (Grant No. NRF-2009-351-C00028). S. W. L. acknowledge supports from the National Research Foundation of Korea (NRF) funded by the Korea government (Grant No. 3348-20100018), and the T. J. Park Foundation.

Appendix.—Suppose that an arbitrary input state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ and ancillary n-qubit states $|0\rangle^{\otimes n}$ are prepared for the measurement. A general measurement can be described as the combination of a unitary operation U acting on the total $(n+1)$ qubits and a projection measurement acting on the selected m-qubits out of the $(n+1)$ qubits. The probability that m-qubits are projected on $\hat{P}_{\bar{i}} = |\bar{i}\rangle\langle\bar{i}| = |i_1, \ldots, i_m\rangle\langle i_1, \ldots, i_m|(i_1, \ldots, i_m \in \{0, 1\})$ is given by

$$p_{\bar{i}} = \mathrm{Tr}(\hat{P}_{\bar{i}}\hat{U}|\psi\rangle\langle\psi| \otimes |0\rangle\langle 0|^{\otimes n}\hat{U}^\dagger\hat{P}_{\bar{i}}). \tag{A1}$$

If the probability $p_{\bar{i}}$ of each measurement outcome $\bar{i}$ is independent on the input state $|\psi\rangle$, then no information about $|\psi\rangle$ is obtained through the measurement. In this case, the input state can be retrieved deterministically as shown below.

Let us define $|\psi_j\rangle = \hat{U}|j\rangle \otimes |0\rangle^{\otimes n}$ ($j \in \{0, 1\}$). Since the probability $p_{\bar{i}}$ in Eq. (A1) is invariant for any input state $|\psi\rangle$, we obtain an orthogonal condition

$$\langle\psi_0|\bar{i}\rangle\langle\bar{i}|\psi_1\rangle = 0, \tag{A2}$$

where $\langle\psi_0|\bar{i}\rangle$ is a $(n - m + 1)$-qubit bra vector. By normalizing $\langle\bar{i}|\psi_0\rangle$ and $\langle\bar{i}|\psi_1\rangle$, we obtain two orthonormal vectors, saying $|\varphi_{0\bar{i}}\rangle$ and $|\varphi_{1\bar{i}}\rangle$. Then $|\psi_j\rangle$ can be represented by

$$|\psi_j\rangle = \sum_{\bar{i}} \sqrt{p_{\bar{i}}}\,|\bar{i}\rangle \otimes |\varphi_{j\bar{i}}\rangle, \tag{A3}$$

where $|\bar{i}\rangle$ and $|\varphi_{ji}\rangle$ are a projected m-qubit state and a corresponding $(n - m + 1)$-qubit state, respectively. As $\hat{U}$ is a linear operator, the evolution of total $n + 1$ qubits under $\hat{U}$ is given by

$$\hat{U}|\psi\rangle \otimes |0\rangle^{\otimes n} = \sum_{i} \sqrt{p_i}\,|\bar{i}\rangle \otimes (\alpha|\varphi_{0\bar{i}}\rangle + \beta|\varphi_{1\bar{i}}\rangle). \tag{A4}$$

If the outcome on m-qubit projection is $|\bar{i}\rangle$, then remaining $(n - m + 1)$ qubits are reduced to $\alpha|\varphi_{0\bar{i}}\rangle + \beta|\varphi_{1\bar{i}}\rangle$. Since $|\varphi_{0\bar{i}}\rangle$ and $|\varphi_{1\bar{i}}\rangle$ are orthonormal vectors determined by $\hat{U}$, we can retrieve the input state by performing a proper unitary operation on the remaining state. The reversal is possible for any measurement outcome $\bar{i}$ whenever $p_{\bar{i}}$ is independent of the input state. We thus conclude that if no information is extracted through the measurement, the whole information is preserved in the remaining part of the Hilbert space describing the measurement and the original state can be retrieved deterministically.

*swleego@gmail.com

[1] W. Heisenberg, Z. Phys. **43**, 172 (1927).
[2] K. Banaszek, Phys. Rev. Lett. **86**, 1366 (2001).
[3] K. Banaszek and I. Devetak, Phys. Rev. A **64**, 052307 (2001).
[4] G. M. D'Ariano, Fortschr. Phys. **51**, 318 (2003).
[5] F. Sciarrino, M. Ricci, F. De Martini, R. Filip, and L. Mista, Phys. Rev. Lett. **96**, 020408 (2006); U. L. Andersen, M. Sabuncu, R. Filip, and G. Leuchs, Phys. Rev. Lett. **96**, 020409 (2006).
[6] M. F. Sacchi, Phys. Rev. Lett. **96**, 220502 (2006).
[7] M. Ueda and M. Kitagawa, Phys. Rev. Lett. **68**, 3424 (1992); A. Royer, Phys. Rev. Lett. **73**, 913 (1994); **74**, 1040(E) (1995); M. Ueda, N. Imoto, and H. Nagaoka, Phys. Rev. A **53**, 3808 (1996); H. Terashima and M. Ueda, Phys. Rev. A **74**, 012102 (2006).
[8] M. Koashi and M. Ueda, Phys. Rev. Lett. **82**, 2598 (1999).
[9] A. N. Korotkov and A. N. Jordan, Phys. Rev. Lett. **97**, 166805 (2006).
[10] A. N. Jordan and A. N. Korotkov, Contemp. Phys. **51**, 125 (2010).
[11] Q. Sun, M. Al-Amri, and M. S. Zubairy, Phys. Rev. A **80**, 033838 (2009).
[12] N. Katz *et al.*, Phys. Rev. Lett. **101**, 200401 (2008).
[13] Y.-S. Kim, Y.-W. Cho, Y.-S. Ra, and Y.-H. Kim, Opt. Express **17**, 11978 (2009).
[14] A. N. Korotkov and K. Keane, Phys. Rev. A **81**, 040103(R) (2010); J.-C. Lee, Y.-C. Jeong, Y.-S. Kim, and Y.-H. Kim, Opt. Express **19**, 16309 (2011).
[15] Y.-S. Kim, J.-C. Lee, O. Kwon, and Y.-H. Kim, Nature Phys. **8**, 117 (2011).
[16] M. A. Nielsen and C. M. Caves, Phys. Rev. A **55**, 2547 (1997).
[17] S. L. Braunstein and A. K. Pati, Phys. Rev. Lett. **98**, 080502 (2007).
[18] F. Buscemi, M. Hayashi, and M. Horodecki, Phys. Rev. Lett. **100**, 210504 (2008); S. Luo, Phys. Rev. A **82**, 052103 (2010).
[19] H. Terashima, Phys. Rev. A **83**, 032114 (2011).
[20] M. G. Genoni and M. G. A. Paris, Phys. Rev. A **74**, 012301 (2006).
[21] S.-Y. Baek, Y. W. Cheong, and Y.-H. Kim, Phys. Rev. A **77**, 060308(R) (2008).
[22] R. F. Werner, Phys. Rev. A **40**, 4277 (1989); S. Albeverio, S.-M. Fei, and W.-L. Yang, Phys. Rev. A **66**, 012301 (2002).
[23] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
[24] J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi, Phys. Rev. Lett. **78**, 3221 (1997); S. Ritter, C. Nölleke, C. Hahn, A. Reiserer, A. Neuzner, M. Uphoff, M. Mücke, E. Figueroa, J. Bochmann, and G. Rempe, Nature (London) **484**, 195 (2012).
[25] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).
[26] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).