Classically Replaceable Operation

Guoding Liu, Xingjian Zhang, and Xiongfeng Ma∗
Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, 100084 China

Quantum information science provides powerful technologies beyond the scope of classical physics. In practice, accurate control of quantum operations is a challenging task with current quantum devices. The implementation of high fidelity and multi-qubit quantum operations consumes massive resources and requires complicated hardware design to fight against noise. An approach to improving the performance and efficiency is to replace some quantum operations with tractable classical processing. Despite the common practice of this approach, rigorous criteria to determine whether a given quantum operation is replaceable “classically” are still missing. In this work, we define the classically replaceable operations in four general scenarios. In each scenario, we provide their necessary and sufficient criteria and point out the corresponding classical processing. For a practically favourable case of unitary classically replaceable operations, we show that the replaced classical processing is deterministic. Beyond that, we regard the irreplaceability with classical processing as a quantum resource and relate it to the performance of a channel in a non-local game, as manifested in a robustness measure.

I. INTRODUCTION

Quantum technology realizes information processing tasks that cannot be achieved by classical means. For instance, quantum key distribution (QKD) brings information-theoretic security between two remote parties [1, 2]. In the field of computation, quantum computers have the potential to bring an exponential speed-up in solving certain problems. The recent realizations of controllable quantum systems with dozens of qubits exhibit the “quantum advantage” over their classical counterparts [3–5]. On the other hand, quantum information processing is costly. Due to the unavoidable noise in quantum devices, current realizations of high-fidelity multi-qubit quantum operations are extremely challenging. On the contrary, classical processing is in general easier to implement and has higher accuracy. The tractability of classical processing motivates us to replace some quantum operations in a task with classical ones. In QKD, after the quantum stage of state preparation and measurement, the users would apply certain quantum operations to distill a secret key [6]. In practice, these operations contain a mass of multi-qubit gates, like Control-NOT and Tofolli gates, and hence are very challenging to realize. Thanks to Shor and Preskill’s seminal work to reduce quantum operations to classical post-processing [7], QKD becomes the earliest practical application in quantum information science [8].

The tractability of classical processing motivates us to replace some quantum operations in a quantum circuit with classical ones. In QKD, after the quantum stage of state preparation and measurement, the users would apply certain quantum operations to distill a secret key [6]. In practice, these operations contain a mass of multi-qubit gates, like Control-NOT and Tofolli gates, and hence are very challenging to realize. Thanks to Shor and Preskill’s seminal work to reduce quantum operations to classical post-processing [7], QKD becomes the earliest practical application in quantum information science [8].

The idea of replacing is also conducive to quantum computing. Within the decoherence time of a quantum system, the number of implementable operations is bounded [9], hindering the ability of quantum computing to solve large-scale problems. If we replace some quantum operations in a quantum circuit with classical processing, we can reduce the number of quantum gates. For instance, when realizing a quantum circuit, we shall omit any operation commuting with the observable to be measured, as these operations do not change the measurement results. We may also replace entanglement breaking (EB) channels [10] with measurement, classical processing, and state preparation in the middle of a quantum circuit.

These instances show that certain quantum operations under specific scenarios can be replaced by classical processing. In practice, classical replacement is flexible and can be embedded into many quantum information processing protocols. In particular, it is appealing to combine the idea with the recent quantum-classical-hybrid algorithms and data processing techniques, such as the variational quantum algorithms [11] and shadow tomography [12], where we may further reduce the difficulty in the quantum part of the task.

Here, we are interested to see what kind of quantum operations are classically replaceable, namely, classically replaceable operation (CRO). In the study of certain specific resource theories, such as quantum coherence of states and entanglement of quantum channels [10, 13–15], the issue has been implicitly considered. Nevertheless, at the moment, a systematic study of CROs is still missing. There are neither rigorous definitions nor efficient criteria for a CRO. On the other hand, there are quantum operations that cannot be replaced classically. Then, this “irreplaceability” implies quantumness intractable to classical means, which might exhibit quantum advantages in some information

∗ xma@tsinghua.edu.cn
processing tasks. A mathematical characterization of the irrereplaceability would reveal the boundary between quantum and classical technologies.

In this work, we provide rigorous definitions of CROs in four different scenarios, depending on whether the input and output of the operation are classical or quantum. It is worth noting that, whether a quantum operation can be precisely replaced is what we are concerned with for classical replacement. In the case where the input and output are both classical, classical replacement is essentially classical simulation and any quantum operation is a CRO. We shall focus on the rest three cases where at least either the input or the output is quantum. We characterize the three CRO sets by showing their necessary and sufficient criteria. Furthermore, we establish a resource-theoretic framework to quantify the irrereplaceability of an operation. Interestingly, we can prove the existence of a non-local game to show the quantum advantage of irreplaceable operations.

The paper is organised as follows. In Section II, we clarify the definitions of “classical processing” and “classical replaceability” in this work. In Section III, we show the main result – providing the mathematical characterization of CRO sets by showing their necessary and sufficient criteria. Furthermore, we establish a resource-theoretic framework to quantify the irrereplaceability of an operation. Interestingly, we can prove the existence of a non-local game to show the quantum advantage of irreplaceable operations.

II. CLASSICAL REPLACEABILITY

First, we introduce the notations. A quantum system is represented as a state in a Hilbert space $\mathcal{H}$. For simplicity, we assume the dimension of $\mathcal{H}$ is finite. Denote $d = \dim \mathcal{H}$, we could find $d$ orthonormal vectors $\{|e_i\rangle, i \in [d]\}$ as a basis of $\mathcal{H}$, where $[d] = \{0, 1, \ldots, d - 1\}$. The choices of basis can be infinite and we usually define one basis $\{|i\rangle, i \in [d]\}$ as the computational basis of $\mathcal{H}$. If there are $n$ quantum systems $\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{n-1}$, the computational basis of $\bigotimes_{k=0}^{n-1} \mathcal{H}_k$ is defined as the tensor product of computational basis on each subsystem. Denote $\mathcal{D}(\mathcal{H})$ as the set of all density operators on $\mathcal{H}$, a quantum operation, $O$, is defined as a completely positive and trace preserving (CPTP) map acting on $\mathcal{D}(\mathcal{H})$. That is, $\forall \rho \in \mathcal{D}(\mathcal{H}')$ and $\forall \sigma \in \mathcal{D}(\mathcal{H})$,

$$
\begin{align*}
I' \otimes O(\rho) &\geq 0, \\
\text{tr}(O(\sigma)) &= \text{tr}(\sigma),
\end{align*}
$$

where $\mathcal{H}'$ is an ancillary quantum system with an arbitrary dimension and $I'$ is the identity on $\mathcal{D}(\mathcal{H}')$. In the following, we also denote CPTP as the set of all quantum operations. The multiplication or composition on CPTP, is denoted as $\circ$.

A classical system can be viewed as a random variable. If a random variable has $d$ values, we call it as a dit with dimension $d$. Given a dit taking the value $i$ with probability $p_i$, $i \in [d]$, the state preparation under basis $\{|e_i\rangle, i \in [d]\}$ is transforming the dit to the quantum state $\sigma = \sum_{i=0}^{d-1} p_i |e_i\rangle \langle e_i|$. Reversely, given a quantum state $\rho$, we can perform a measurement under basis $\{|e_i\rangle, i \in [d]\}$ and obtain a random variable taking the value $i$ with probability $\text{tr}(\rho |e_i\rangle \langle e_i|)$.

Now, let us clarify the meaning of “classical processing”. Classical computers can be viewed as Turing machines or circuit models [16, 17]. Any function of the form $f : [d] \rightarrow [d']$ can be realized with a certain number of fixed logical gates in classical circuits [9], where $d, d' \in \mathbb{Z}_+$. If we introduce probabilistic Turing machines [18, 19], we have the ability to compute random functions. In principle, we can use a classical computer to realize any probabilistic function subject to a pre-determined probability distribution. In this work, we do not consider the computability of probabilistic Turing machine. That is, we do not consider the computational complexity to realize a random function.

Mathematically, a random function can be described by a stochastic matrix. Given a dit $s$ with the probability mass function $q_s$, a probabilistic map $O_c : [d] \rightarrow [d']$ transfers it to a new random variable $s' \in [d']$. A specific result $s'$ occurs with probability

$$
\Pr(O_c(s) = s') = \sum_s P_{s's} q_s,
$$

where the probabilistic map is described by the stochastic matrix $P$, satisfying $P_{s's} \geq 0$ and $\sum_s P_{s's} = 1, \forall s$. In general, $d$ and $d'$ can be different. Without loss of generality, we assume the two values are the same, as we can always dilate the random variable with a smaller alphabet by adding trivial elements. The set of probabilistic maps $O_c : [d] \rightarrow [d]$ describes classical processing with dimension $d$.

Now we aim to find out all the quantum operations that can be replaced by classical processing. Here, we require the dimensions of the replaced quantum operation and the replacing classical processing to be the same, which means replacing a CRO acting on a qudit with a classical processing acting on a dit.
We distinguish CROs in four cases. To be specific, we restrict the input and output of the operation to be classical or quantum, and then find out the CROs. They are classical-quantum CROs (cqCRO), quantum-quantum CROs (qqCRO), quantum-classical CROs (qcCRO), and classical-classical CROs (ccCRO). Here, classical input means that we prepare a state based on a classical random variable and set this state as the input of the quantum operation. Classical output means that we further implement a measurement after the action of a quantum operation. The measurement result is the classical output and we do not concern about the remained state after the measurement. The replacement means that the output does not change for any input, even the input is unknown.

**Definition 1** (Classically replaceable operation). There are four cases that a quantum operation is replaceable with classical processing, cqCRO, qqCRO, qcCRO, and ccCRO defined as follows, as shown in Figure 1. Here, we define a fixed computational basis for state preparation and measurement.

1. Consider operations right after a state preparation. A cqCRO can be realized by first preprocessing the input classically and then preparing the output state.
2. Consider operations with quantum input and quantum output. A qqCRO can be realized by first measuring the input, processing the outcomes classically, and then preparing the output state.
3. Consider operations right before a measurement. A qcCRO can be realized by first measuring the input state and then processing the outcomes classically.
4. Consider operations between a state preparation and a measurement. A ccCRO can be realized by processing the input classically.

![Diagram of CROs](image_url)

**FIG. 1.** Four different kinds of CROs. The CROs are the quantum operations in dark blue blocks excluding the state preparation and measurement. Here, all state preparation and measurement are performed on the computational basis. ① Given any classical input $s_{in}$, the output quantum state $\rho_{out}$ of a cqCRO, $O_c$, can be obtained by first preprocessing $s_{in}$ with classical processing $O_{ic}$ and then preparing the output state. ② Given any input quantum state $\rho_{in}$, the output quantum state $\rho_{out}$ of a qqCRO, $O_m$, can be obtained by first measuring the state, applying classical processing $O_{mc}$ to the outcomes, and preparing the state. ③ Given any input state $\rho_{in}$, the measurement result $s_{out}$ after performing a qcCRO, $O_f$, can be obtained by first measuring $\rho_{in}$ and then processing the outcomes classically with $O_{fc}$. ④ Given any classical input $s_{in}$, the classical output $s_{out}$ of a ccCRO, $O_c$, can be obtained by processing $s_{in}$ classically with $O_{ic}$.

It is worth mentioning that in Definition 1, the CRO is the quantum operation _beside_ the state preparation and measurement where the state preparation and measurement are _excluded_. We require the most strict condition for qqCRO as we do not put any restriction on the input and output. Any qqCRO is a cqCRO and a qcCRO. Interestingly, in Section III, we show that the intersection of qqCRO and qcCRO is not qqCRO. Furthermore, the case of ccCRO...
can be viewed as the classical simulation. As all quantum operations can be simulated classically, ccCRO contains all quantum operations so that cqCRO, qqCRO, and qcCRO are all subsets of ccCRO. Note that in this work we might use CRO to represent an replaceable operation or the set of replaceable operations depending on the context.

In Definition 1, a CRO only guarantees that the subsystem it acts on is unchanged before and after classical replacement. The state of the rest part of the system might change. The correlation between the subsystem and the rest one could change as well. The reason is that a CRO might be a part of large evolution on the whole system. Replacing the CRO with classical processing could destroy this evolution so that the joint state of the system and the environment is different before and after replacement. Strictly speaking, the evolution of the subsystem originating from a large evolution on the whole system is not necessarily a CPTP map. Here, we only consider the case that the evolution of the subsystem is a CPTP map. We provide more discussions in Appendix B.

Now, let us consider the special case of unitary CROs, namely, classically replaceable unitary operations (CRU). As an unitary is a closed evolution, the correlation between the subsystem and the rest one remains the same before and after replacing CRU. That means, the joint state of the system and the environment does not change. Moreover, in Section III, we prove that cqCRU, qqCRU, qcCRU can be replaced by deterministic classical processing while this conclusion is not true for ccCRU. A counter-example is that a Hadmard gate between the state preparation and measurement on a qubit can only be replaced with probabilistic classical processing. Figure 2 shows examples of CRUs in a quantum circuit. The state of the whole quantum circuit at any time does not change before and after replacement for any circuit input.

![Diagram](image-url)

**FIG. 2.** Within a quantum circuit, cqCRU, qqCRU, and qcCRU always lie in the beginning, middle, and ending of the circuit, respectively. All state preparation operations and measurements are performed on the computational basis. For any input statistics \((s_{i1}, s_{i2}, s_{i3})\), the state of the whole quantum circuit at any time does not change before and after replacing CRU \(U_{i2}, U_{m1}\), and \(U_{f2}\) with \(U_{ic}, U_{mc}\), and \(U_{fc}\), respectively. The statistics of output \((s_{o1}, s_{o2}, s_{o3})\) also remains the same.

**III. MATHEMATICAL CHARACTERIZATION OF CRO**

In this section, we first provide the mathematical equivalent definitions for four kinds of CROs. As any quantum operation is a ccCRO, we focus on the rest three cases. Then we discuss the extensions of CRO when the state preparation and measurement are not performed on the computational basis.
A. Equivalent Definitions of CRO

In this subsection, we provide the mathematical characterization of the first three kinds of CROs in Definition 1. Given the computational basis on $\mathcal{H}$, the dephasing operation is defined as a map from $\mathcal{D}(\mathcal{H})$ to $\mathcal{D}(\mathcal{H})$ that $\forall \rho \in \mathcal{D}(\mathcal{H})$, 

$$\Delta(\rho) = \sum_{i=0}^{d-1} |i\rangle\langle i| \rho|i\rangle\langle i|.$$

Then, we mathematically characterize the three kinds of CROs.

**Theorem 1.** For any $O \in \text{CPTP}$,

- $O \in \text{cqCRO} \iff O \circ \Delta = \Delta \circ O \circ \Delta \iff \exists O' \in \text{CPTP}, O \circ \Delta = \Delta \circ O' \circ \Delta$;
- $O \in \text{qqCRO} \iff O = \Delta \circ O \circ \Delta \iff \exists O' \in \text{CPTP}, O = \Delta \circ O' \circ \Delta$;
- $O \in \text{qcCRO} \iff \Delta \circ O = \Delta \circ O \circ \Delta \iff \exists O' \in \text{CPTP}, \Delta \circ O = \Delta \circ O' \circ \Delta$.

We provide the detailed proof of Theorem 1 in Appendix C. Here, we provide the sketch of the proof. As the proof for three kinds of CROs are similar, we exhibit the idea with respect to qcCRO. Theorem 1 contains three equivalent criteria for each kind of CRO. We first prove the equivalence of the second criterion and the third, then the first and the second. The equivalence of the second criterion and the third can be directly verified by expressing channels in Choi-matrix representation.

For the next proof, we first find the corresponding classical processing for each operation satisfying $\Delta \circ O = \Delta \circ O \circ \Delta$. Then we prove that if any operation does not satisfy that criterion, then it cannot be replaced. The argument for the irreplaceability of an operation comes from that it can output different results for two inputs but through classical processing we can only get the same results. From this we can conclude there exists a quantum state that the output cannot be obtained after replacing the operation with classical processing.

From Theorem 1, we can define the sets for different CROs,

$$\text{cqCRO} = \{ O \in \text{CPTP} | O \circ \Delta = \Delta \circ O \circ \Delta \}$$

(4)

$$\text{qqCRO} = \{ O \in \text{CPTP} | O = \Delta \circ O \circ \Delta \}$$

(5)

$$\text{qcCRO} = \{ O \in \text{CPTP} | \Delta \circ O = \Delta \circ O \circ \Delta \}$$

(6)

In Appendix C, we show that for any qcCRO $O$, the stochastic matrix $P$ of the corresponding classical processing satisfies $P_{s's'} = \text{tr}(|s'\rangle\langle s'| O(|s\rangle\langle s|))$. If $\exists U$, where $U$ is unitary, such that $\forall \rho \in \mathcal{D}(\mathcal{H}), O(\rho) = U \rho U^\dagger$. Then Eq. (6) requires that

$$\sum_k \langle i|U|j\rangle \langle j|U^\dagger|k\rangle \langle k| = \delta_{ij} \sum_k \langle i|U|k\rangle \langle k|U^\dagger|i\rangle \langle i|,$$

(7)

where $\delta_{ij}$ equals 1 if $i = j$ and 0 otherwise. Then for any $i, j, k$ and $i \neq j$,

$$U_{ki}U^*_{kj} = 0,$$

(8)

where $U_{ki} = \langle k|U|i\rangle$. It means for each row of $U$, there is only one element nonzero. As $P_{s's} = \text{tr}(|s'\rangle\langle s'| U |s\rangle\langle s| U^\dagger) = |U_{s's}|^2$, the value of $P_{s's}$ must be 0 or 1. Then the replacing classical processing is deterministic. Similar arguments apply to unitary operations in cqCRO and qqCRO.

**Corollary 1.** cqCRU, qqCRU and qcCRU can be replaced with deterministic classical processing.

It is interesting that from Theorem 1, we can obtain a relation between CRO and the resource theory of coherence [20]. In the framework of coherence resource theory, the free states, also named incoherent states, are the states diagonal on the computational basis. We can define two sets of operations: maximally incoherent operations (MIO) [20, 21] and coherence non-activating operations (CNAO) [13]. The former is the maximal set of operations that cannot generate coherence from incoherent states. The latter is the maximal set of operations that cannot activate coherence from the input state. The details of coherence theory are shown in Appendix A.

In fact, Eq. (4) can be seen as the equivalent definition of MIO [13]. Similarly, Eq. (5) is a definition for CNAO. Then, we see the equivalence between cqCRO and MIO, qcCRO and CNAO. The intersection of cqCRO and qcCRO is just the intersection of MIO and CNAO, which is a set of dephasing-covariant incoherent operations (DIO) [13, 21], defined as

$$\text{DIO} = \{ O \in \text{CPTP} | \Delta \circ O = O \circ \Delta \}.$$

(9)
From Eq. (5) and Eq. (9), we can see that \( \text{qqCRO} \) is a proper subset of DIO, hence \( \text{qqCRO} \) is not the intersection of \( \text{cqCRO} \) and \( \text{qcCRO} \). An operation belonging to both \( \text{cqCRO} \) and \( \text{qcCRO} \) might not be \( \text{qqCRO} \). For example, CNOT gate is a DIO but not a \( \text{qqCRO} \). At the same time, \( \text{qqCRO} \) is a subset of EB channels [10], since any \( \text{qcCRO} \), \( O \), can be given by

\[
O(\rho) = \sum_{ij} \text{tr}(\rho |i\rangle \langle i|) P_{ij} |j\rangle \langle j|,
\]

where \( P_{ij} \) is the stochastic matrix associated with the replaced classical processing. We visualize the relations among three kinds of CROs, MIO, CNAO and EB channels in Figure 3.

![Diagram of CROs and EB channels](image)

**FIG. 3.** Relations among three kinds of CROs, MIO, CNAO and EB channels. Here, \( \text{cqCRO} = \text{MIO} \), \( \text{qcCRO} = \text{CNAO} \), DIO = \( \text{cqCRO} \cap \text{qcCRO} \), \( \text{qqCRO} \subset \text{DIO} \), \( \text{qqCRO} \subset \text{EB} \).

### B. Extension of CRO with Degeneracy

In Definition 1, we assume the state preparation and measurement to be rank-one, that is, a classical value uniquely corresponds to a basis vector. In general, an observable could have degeneracy which means we might prepare different states for the same inputs or obtain the same measurement results for different quantum states. In this case, the description for CROs could change. Here, we only consider the extension of \( \text{qcCRO} \). For simplicity, we also assume the measurement to be projective value measurements (PVM).

The PVM after the replaced quantum operation is pre-determined. The projectors of this PVM are denoted as \( \{P_n, n \in [d']\} \) where \( d' \leq d \), \( P_n P_m = \delta_{nm} P_n \), \( \sum_n P_n = \mathbb{I}_d \). The coarse graining dephasing operation \( \Delta' \) is defined as: \( \forall \rho \in D(\mathcal{H}) \),

\[
\Delta'(\rho) = \sum_n P_n \text{tr}(P_n \rho).
\]

Now we aim to use the same PVM followed with classical processing to replace an operation before a PVM. The statistics of the final measurement result maintains the same before and after the replacement. The proof is similar with the scenario of \( \text{qcCRO} \) with details shown in Appendix C. The result is

\[
\text{qcCROext} = \{O \in \text{CPTP}|\Delta' \circ O = \Delta' \circ O \circ \Delta'\}.
\]

In fact, \( \text{qcCRO} \) is a special case of \( \text{qcCROext} \) when \( \Delta' = \Delta \).

### C. Extension of CRO with Unitary Transformation

In the discussion above, we fix the basis of state preparation and measurement. Nevertheless, in practice, the state preparation and measurement basis could be selected from multiple choices. When we use classical processing to replace the quantum operation, we can prepare the state or measure the state with one of these choices instead of a fixed one.
For simplicity, we assume the state preparation and measurement is non-degenerate. As we can apply unitary to change bases, the freedom to choose bases is equivalent to the freedom to apply unitary after state preparation and before measurement. The unitary is chosen from an ensemble $U$, which is often limited by quantum devices in practice. The largest unitary ensemble is $U = U_d$, where $U_d$ is the unitary group with dimension $d$. In reality, single-qubit operations are normally easy to implement. In this case, we are often interested in the local unitary ensemble, $U = U_d^\otimes n$, where $n$ is the number of qubits in the quantum system.

Given the ensemble $U$, we denote the extensions of cqCRO, qqCRO and qcCRO as $cqCRO_U$, $qqCRO_U$, and $qcCRO_U$, respectively, as shown in Figure 4. The three kinds of CRO are defined as follows.

1. Consider operations right after the state preparation, a $cqCRO_U$ can be realized by first processing the input classically and then preparing the state under a chosen basis.

2. Consider operations with quantum input and quantum output. A $qqCRO_U$ can be realized by first measuring the input with a basis, processing the outcomes classically, and then preparing the output state under another possibly different basis.

3. Consider operations right before measurement. A $qcCRO_U$ can be realized by measuring the quantum input and processing the outcomes classically.

Note that the classical value can always be discriminated and copied, which means we can choose the state preparation basis based on the classical input and measurement outcomes. This is the origin of control unitary in Figure 4(a) and 4(b). In our model, we suppose the quantum input is unknown and cannot be discriminated before measuring it, so the measurement basis cannot be chosen depending on the quantum input. That means in the first step of replacing $qqCRO_U$ or $qcCRO_U$, we need to select a fixed measurement independent of the input.

From Figure 4(c) we can see that an operation, $O$, is $qcCRO_U$ if and only if there exists unitary $U \in U$ s.t., $O \circ U^{-1} \in qcCRO$. The set of $qcCRO_U$ is the union set of $\{qcCRO_U\}$ where $U \in U$,

$$qcCRO_U = \{O \in CPTP | \exists U \in U, \Delta \circ (O \circ U^{-1}) = \Delta \circ (O \circ U^{-1}) \circ \Delta\}. \ (13)$$
We remark that unlike qcCRO_\(U\), cqCRO_\(U\) or qqCRO_\(U\) is not a simple union of \{cqCRO_\(U\)\} or \{qqCRO_\(U\)\}, \(U \in \mathcal{U}\), due to the possible dependence of state preparation basis on the classical input or the measurement result.

From the operational meaning of qqCRO_\(U\), we can find another representation. Any qqCRO_\(U\) can be realized by measuring the input state \(\rho\) with a basis \{\(e_i\)\}, transforming measurement result \(i\) into \(j\) with probability \(P_{ij}\), and then preparing the state under another possibly different basis \{\(f_j^{'}\)\},

\[
O_m(\rho) = \sum_{i,j} P_{ij} \langle e_i | \rho | e_i \rangle |f_j^{'}\rangle \langle f_j'| |
\]

\[
= \sum_i (\langle e_i | \rho | e_i \rangle) (\sum_j P_{ij} |f_j^{'}\rangle \langle f_j'| ),
\]

where \(P_{ij}\) is the element of a stochastic matrix, \{\(\{f_j^{'}\}\)\} reflects the dependence between the state preparation basis and the measurement result \(i\). From Eq. (14) we can see that any qqCRO_\(U\) is an EB channel [10], which measures the state with a positive operator valued measurement (POVM) followed by state preparation.

A classical input can always be modelled as a quantum input followed with the computational basis measurement as shown in Figure 5. We can see that if an operation, \(O_i\), is a cqCRO_\(U\), then \(O_i \circ \Delta\) is a qqCRO_\(U\). From a similar argument of qqCRO_\(U\), we obtain that any qqCRO_\(U\), \(O_i\), satisfies,

\[
O_i \circ \Delta(\rho) = \sum_{i,j} P_{ij} |i \rangle \langle i| |f_j^{'}\rangle \langle f_j' |
\]

\[
= \sum_i |i \rangle \langle i| (\sum_j P_{ij} |f_j^{'}\rangle \langle f_j' |),
\]

where \(P_{ij}\) is the element of a stochastic matrix, \{\(|i\rangle\)\} is the computational basis, \{\(|f_j^{'}\rangle\)\} is a basis for state preparation.

Interestingly, if \(\mathcal{U} = \mathcal{U}_d\), then any CPTP map is a cqCRO_\(U\). The reason is that we can evaluate the output quantum state \(\rho_{out}\) after reading the classical input \(s_{in}\). With the freedom to choose arbitrary unitary, one can prepare \(\rho_{out}\) directly.

### IV. CHARACTERIZATION OF IRREPLACEABILITY

The replaceability with classical processing of CRO motivates us to view the irreplaceability as a kind of quantum resource [22–24]. We can establish a channel resource theory to quantitatively study the potential quantum advantage brought by irreplaceability. We take CRO as the set of free channels, and any quantum operation outside this set contains the resource of irreplaceability. Depending on the nature of the channel input and output, we can choose different types of CROs and specify a corresponding resource theory. In this work, we take qcCRO as an example.

To establish a channel resource theory for irreplaceability, we first need to specify the free channels and free super operations [23]. The set of free channels is naturally given by Eq. (6). We consider the set of resource non-generating (RNG) super operations to be the set of free super operations \(\mathcal{F}\),

\[
\mathcal{F} = \text{RNG} = \{ \Lambda | \Lambda (M) \in \text{qcCRO}, \forall M \in \text{qcCRO}\}.
\]
To quantify the amount of irreplaceability of a channel, we can utilize the Choi-state representation of the channel. Given a channel $\mathcal{N}$ acting on $\mathcal{D}(\mathcal{H})$, the corresponding Choi-state is

$$\Phi_{\mathcal{N}} = \mathbb{I} \otimes \mathcal{N}(|\Phi^+\rangle \langle \Phi^+|),$$

(17)

where $|\Phi^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$ is the maximally entangled state on $\mathcal{H} \otimes \mathcal{H}$. We define two measures to characterize irreplaceability. One is relative entropy of irreplaceability with details shown in Appendix E. The other is robustness of irreplaceability defined as follows.

**Definition 2 (Robustness of Irreplaceability).** Given a channel $\mathcal{N} \in \text{CPTP}$, the robustness of irreplaceability of $\mathcal{N}$ is

$$R(\mathcal{N}) = \min_{\mathcal{M} \in \text{CPTP}} \left\{ s \geq 0 \left[ \frac{\Phi_{\mathcal{N}} + s\Phi_{\mathcal{M}}}{1 + s} \in \mathcal{F} \right] \right\},$$

(18)

where $\mathcal{F} = \{ \Phi_{\mathcal{M}} | \mathcal{M} \in \text{qcCRO} \}$.

The robustness of irreplaceability in Eq. (18) has several equivalent definitions, as shown by the following lemma.

**Lemma 1.** The robustness of irreplaceability of a channel $\mathcal{N}$ can be equivalently given by

1. **Monotonicity:** $\forall \mathcal{N} \in \text{CPTP}, \forall \Lambda \in \mathcal{F}$,
   
   $$R(\Lambda(\mathcal{N})) \leq R(\mathcal{N}).$$

   (23)

2. **Convexity:** Given an index set $\mathcal{I}$, $\forall \mathcal{N} \in \text{CPTP}, \forall \{p_i\}_{i \in \mathcal{I}}$ such that $p_i \geq 0, \sum_{i \in \mathcal{I}} p_i = 1$,

   $$R\left( \sum_{i \in \mathcal{I}} p_i \mathcal{N}_i \right) \leq \sum_{i \in \mathcal{I}} p_i R(\mathcal{N}_i).$$

   (24)

We leave the proof of the lemma in Appendix D. From Lemma 1, we can see that $R(\mathcal{N}) = R(\Delta \circ \mathcal{N})$. The robustness is a valid measure, which vanishes to zero only for qcCRO and enjoys the properties of monotonicity under free superoperations and convexity under channel mixing, as shown by the following lemma.

**Lemma 2.** The robustness of irreplaceability in Eq. (18), $R(\mathcal{N})$, has the following properties:

1. **Monotonicity:** $\forall \mathcal{N} \in \text{CPTP}, \forall \Lambda \in \mathcal{F}$,

   $$R(\Lambda(\mathcal{N})) \leq R(\mathcal{N}).$$

   (23)

2. **Convexity:** Given an index set $\mathcal{I}$, $\forall \mathcal{N} \in \text{CPTP}, \forall \{p_i\}_{i \in \mathcal{I}}$ such that $p_i \geq 0, \sum_{i \in \mathcal{I}} p_i = 1$,

   $$R\left( \sum_{i \in \mathcal{I}} p_i \mathcal{N}_i \right) \leq \sum_{i \in \mathcal{I}} p_i R(\mathcal{N}_i).$$

We leave the proof of the lemma in Appendix D. Intuitively, these properties originate from the convexity of the set of free channels. The robustness can hence be viewed as a geometric measure for irreplaceability that quantifies the distance between the considered channel and the set of qcCRO.

Interestingly, the robustness of irreplaceability has an operational meaning: it measures the advantage that a channel can provide in a non-local game [25, 26] of state discrimination. To be specific, we define a bipartite non-local game in Box 1.
In the case where Alice sends the state $\sigma_i$, the probability that Bob takes a guess of $\sigma_j$ is $\text{tr}(\sigma_i | j\rangle\langle j|)$. On average, the expected payoff value that Alice and Bob can obtain in this game is

$$p(N, G) = \sum_{i,j} \alpha_{ij} \text{tr}(\sigma_i | j\rangle\langle j|).$$

For simplicity, we call it the performance of $N$ in the non-local game $G$. Without loss of generality, we can require that $\forall M \in \text{qcCRO}$, the payoff values $\alpha_{ij}$ satisfy that the performance of all qcCROs is non-negative and bounded by 1. Under this constraint, we have the following theorem showing the advantage brought by an irreplaceable quantum channel.

**Theorem 2.** The best advantage a quantum channel can provide over all qcCROs and all non-local games is given by $1 + R(N)$:

$$\max_{\{\alpha_{ij}\}, \{\sigma_i\}} \min_{M \in \text{qcCRO}} \frac{p(N, G)}{p(M, G)} = 1 + R(N),$$

s.t. $0 \leq p(M, G) \leq 1, \forall M \in \text{qcCRO}.$

The proof of Theorem 2 follows the idea of duality in conic programming [25, 27]. We leave the detailed proof in Appendix D. Here, we discuss the operational meaning of the result. A qcCRO followed with a computational basis measurement is equivalent to applying a computational basis measurement followed with classical processing. As a result, Bob cannot distinguish the off-diagonal terms of $\{\sigma_i\}$ with a qcCRO. If the considered quantum channel is irreplaceable, when optimizing over all possible non-local games to maximize its advantage, we may choose a set of states $\{\sigma_i\}$ that vary mainly in the off-diagonal terms. The replaceable channels would fail in distinguishing the difference in these elements. On the other hand, a higher value of the robustness of irreplaceability brings a stronger ability of a channel to probe the off-diagonal terms of $\{\sigma_i\}$ and hence a higher performance in the non-local game.

In the non-local game as shown in Box 1, Bob can only perform computational basis measurement to distinguish quantum states. We can generalize the non-local game by considering other measurement settings. In fact, for the case where Bob can perform arbitrary positive operator-valued measurements (POVM) to distinguish quantum states, the best advantage a quantum channel can provide over all qcCROs and all non-local games is also $1 + R(N)$. The result is compatible with the conclusion given by Ref. [25]. Another interesting case is where Bob can choose arbitrary local basis measurements. This is often the scenario in practical implementation of quantum information processing protocols, like shadow tomography [12]. In this case, we could choose the convex hull of qcCRO$_{U_2}$ as the set of free channels, where $U$ equals $U_2^m$ and represents the freedom to choose measurement bases. We expect to obtain similar results like Theorem 2 when considering the advantage brought by an irreplaceable quantum channel and leave it for future works.

**V. CONCLUSION**

In this work, we define the concept of CROs in four scenarios. Depending on the feature of the input and output of an operation, we classify CROs into four types. Among these sets, ccCRO is the largest that composes of all quantum operations, while qqCRO is the smallest, which is a subset of all the other three sets. We provide necessary
and sufficient criteria to determine whether an operation is a CRO and present its corresponding classical processing. Interestingly, for the special cases of unitary operations, namely, cqCRU, qqCRU, and qcCRU, one only needs deterministic classical processing for replacement. Furthermore, we discuss two extensions of CRO, where the bases of state preparation and measurement can be degenerate and different.

From a theoretical view, the clarification of replaceability and irreplaceability manifests a difference between classical and quantum operations. Focusing on the set of qcCRO, we establish a resource theory framework for the study of irreplaceability with classical operations. We propose two measures, namely, robustness and relative entropy, to quantify the resource. An interesting discovery is that the robustness measure quantifies the quantum advantage of an operation over all the qcCROs in a non-local game. Besides, we find that cqCRO and qcCRO are equivalent to MIO and CNAO in the resource theory of coherence, respectively, while qqCRO is a subset of both DIO and EB channels. The relation reveals the connection between irreplaceability, coherence, and entanglement.

Note that, unlike classical simulation, we are only concerned with whether a quantum operation can be replaced regardless of the consumption of classical computing resources. This is the case in many subjects like the security analysis of QKD. It may be practically interesting to study the circuit complexity issues [28] in the classical replacement.

There are some other interesting future directions. One direction is extending the concept of CRO from finite dimensions to infinite dimensions. We expect such studies to benefit continuous-variable quantum information processing and inspire new perspectives on the non-classicality in the continuous-variable regime. Another topic is studying the problem of classical replacement in a non-Markovian evolution, that is, the evolution between two different times may not be a CPTP map. It is also interesting to further explore the resource theory of irreplaceability. As in other resource theories, we expect to witness more valid measures with operational meanings for irreplaceability.

ACKNOWLEDGMENTS

We thank Pei Zeng and Junjie Chen for the helpful discussions. This work was supported by the National Natural Science Foundation of China Grants No. 11875173 and No. 12174216 and by the National Key Research and Development Program of China Grants No. 2019QY0702 and No. 2017YFA0303903.

Appendix A: Resource Theory

Quantum resource theory (QRT) studies how to characterize the resource stored in a quantum state, like entanglement [29] and coherence [20]. It also provides a tool to explore the problem of the interconversion between different resources under specific restrictions. In this section we will review the framework of resource theory [30] and introduce the resource theory of coherence [20] as well as the channel resource theory.

1. Framework of Resource Theory

For any QRT, we first point out the following as three main ingredients: resource, free states, and free operations. The resource like coherence is the quantity we characterize in the resource theory. The term “free” means that this kind of states or operations can be obtained at no cost. The state not free is a resourceful state. The three ingredients are not independent to each other. They satisfy the free operations postulate (FOP) [30], that is, any free operation cannot transform a free state to a resourceful state. Denote the Hilbert space as $\mathcal{H}$, the states on $\mathcal{H}$ as $\mathcal{D}(\mathcal{H})$, the set of free states on $\mathcal{D}(\mathcal{H})$ as $\mathcal{F}$. Then for any free operation $\Lambda$,

$$\Lambda(\sigma) \in \mathcal{F}, \forall \sigma \in \mathcal{F}. \quad (A1)$$

From this postulate, we could define a set of operations named resource non-generating operations (RNG), containing all the operations satisfying FOP:

$$\text{RNG} = \{\Lambda \in \text{CPTP} | \Lambda(\sigma) \in \mathcal{F}, \forall \sigma \in \mathcal{F}\}. \quad (A2)$$

CPTP represents the set of all completely positive channels on $\mathcal{D}(\mathcal{H})$. Any set of free operations is a subset of RNG. Different sets of free operations lead to different resource theories while the resource theories with RNG might have a universal property [30].
After defining the three ingredients in QRT, we need to characterize the resource by providing a resource measure. The measure of the resource is a functional $M$ mapping from $\mathcal{D}(\mathcal{H})$ to non-negative real numbers. A valid measure $M$ should satisfy the following two conditions. First, it is 0 for the set of free states:

$$M(\sigma) = 0, \forall \sigma \in \mathcal{F}. \quad (A3)$$

In some QRTs this requirement is more strict. $M$ is 0 if and only if the state is free state:

$$M(\sigma) = 0 \iff \sigma \in \mathcal{F}. \quad (A4)$$

Second, any proper resource measure $M$ cannot increase under the action of free operations. This is the monotone condition. Then, for any free operation $\Lambda$,

$$M(\Lambda(\rho)) \leq M(\rho), \forall \rho. \quad (A5)$$

There are some additional requirements for the measure in different QRTs. These extra requirements could vary a lot for different QRTs. In the resource theory of coherence, it is reasonable that coherence cannot increase under mixing from a physical point of view. This leads to the convexity condition of the measure:

$$M\left(\sum_n c_n \rho_n\right) \leq \sum_n c_n M(\rho_n), \sum_n c_n = 1. \quad (A6)$$

Here, we introduce two kinds of measures. The first is divergence measure, that is, the measure of the resource bases on the divergence between the state and the set of free states $\mathcal{F}$. Divergence is a functional mapping two quantum states into a non-negative real number:

$$D : \mathcal{D}(\mathcal{H}) \times \mathcal{D}(\mathcal{H}) \to \mathbb{R}^+, \quad (A7)$$

requiring $D(\rho, \sigma) = 0 \iff \rho = \sigma$. The divergence of a state $\rho$ and the set of free states $\mathcal{F}$ is defined as

$$D(\rho, \mathcal{F}) = \inf_{\sigma \in \mathcal{F}} D(\rho, \sigma). \quad (A8)$$

To get a well-defined divergence measure, we require $\mathcal{F}$ to be a convex set, i.e., for any $\rho, \sigma \in \mathcal{F}$, $t \rho + (1-t) \sigma \in \mathcal{F}, \forall 0 \leq t \leq 1$. Normally, we take $D$ as the K-L divergence or relative entropy $S$: $S(\rho||\sigma) = \text{tr}(\rho \log \sigma) - \text{tr}(\sigma \log \sigma)$. The relative entropy of the resource is $M(\rho) = S(\rho||\mathcal{F})$. We can prove that this measure meets the requirements of Eq. (A4), Eq. (A5) and Eq. (A6).

It can be verified that divergence measure satisfies Eq. (A4). Due to the contractive property of relative entropy $S$, i.e., for any CPTP channel $\mathcal{E}$,

$$S(\mathcal{E}(\rho)||\mathcal{E}(\sigma)) \leq S(\rho||\sigma), \quad (A9)$$

the monotone condition Eq. (A5) can be fulfilled:

$$M(\Lambda(\rho)) = \inf_{\sigma \in \mathcal{F}} S(\Lambda(\rho)||\sigma)$$

$$\leq \inf_{\sigma \in \mathcal{F}} S(\Lambda(\rho)||\Lambda(\sigma))$$

$$\leq \inf_{\sigma \in \mathcal{F}} S(\rho||\sigma)$$

$$= M(\rho). \quad (A10)$$

Moreover, relative entropy $S$ is jointly convex, then

$$M\left(\sum_n c_n \rho_n\right) \leq D\left(\sum_n c_n \rho_n, \sum_n c_n \sigma_n^*\right)$$

$$\leq \sum_n c_n D(\rho_n, \sigma_n^*)$$

$$= \sum_n c_n M(\rho_n), \quad (A11)$$

where $\sigma_n^*$ is the closest quantum state to $\rho_n$ in $\mathcal{F}$. Equation. (A11) leads to the convexity of the measure.
Another widely used resource measure is robustness of the resource, that is how hard to make the state become a free state with mixing another state. For any state $\rho \in D(\mathcal{H})$, the robustness of the resource is

\[
R(\rho) = \min_{\sigma \in D(\mathcal{H})} s,
\]

\[
s.t. \frac{\rho + s\sigma}{1+s} \in \mathcal{F}, s \geq 0.
\]

Obviously robustness measure satisfies Eq. (A4). We can prove it also satisfies monotone condition. For any quantum state $\rho$, we set $r = R(\rho)$. According to the definition of robustness measure, $\exists \sigma \in D(\mathcal{H}), \frac{\rho + r\sigma}{1+r} \in \mathcal{F}$. Then for any free operation $\Lambda$,

\[
\Lambda\left(\frac{\rho + r\sigma}{1+r}\right) = \frac{\Lambda(\rho) + r\Lambda(\sigma)}{1+r} \in \mathcal{F}.
\]

From Eq. (A12) we can see $R(\Lambda(\rho)) \leq r = R(\rho)$. That is the condition of Eq. (A5).

Moreover, the robustness measure meets the requirement of Eq. (A6) if the set of free states $\mathcal{F}$ is convex. Denote $r_i = R(\rho_i)$ for a set of states $\{\rho_1, \rho_2, \ldots, \rho_n\}$, then $\exists \{\sigma_1, \sigma_2, \ldots, \sigma_n\}, \frac{\rho_i + r_i\sigma_i}{1 + r_i} \in \mathcal{F}, \forall i$. For any convex combination of $\{\rho_1, \rho_2, \ldots, \rho_n\}$: $\rho = \sum_i c_i \rho_i, \sum_i c_i = 1, c_i \geq 0$, set $r = \sum_i c_i r_i$, $\sigma = \sum_i c_i r_i \sigma_i$,

\[
\frac{\rho + r\sigma}{1+r} = \frac{\sum_i c_i (\rho_i + r_i \sigma_i)}{1+r} \in \mathcal{F}.
\]

This accounts for the convexity of the robustness measure.

2. Resource Theory of Coherence

Here we apply the framework of resource theory to characterize coherence. In the resource theory of coherence [20], the set of free states and free operations are named incoherent states and incoherent operations. We first set the computational basis of Hilbert space $\mathcal{H}$, denoted as $\{|i\rangle, i \in [d]\}$. Then we define the incoherent states to be the states only with diagonal terms in computational basis, i.e.,

\[
\mathcal{I} = \left\{ \sum_i c_i |i\rangle \langle i|, \sum_i c_i = 1, c_i \geq 0 \right\}.
\]

The set of incoherent operations has different choices. The largest set of incoherent operations (RNG) is called maximally incoherent operations (MIO) in the resource theory of coherence. Any operation in MIO cannot generates coherence from incoherent states. It can be proved that [13]

\[
\text{MIO} = \{ O \in \text{CPTP} | O \circ \Delta = \Delta \circ O \circ \Delta \},
\]

where $\Delta$ is dephasing operation on $D(\mathcal{H})$. A smaller set of incoherent operations is dephasing-covariant incoherent operations (DIO):

\[
\text{DIO} = \{ O \in \text{CPTP} | O \circ \Delta = O \circ \Delta \}.
\]

DIO is the set of operations commuting with the dephasing operation. Dephasing operation can be viewed as the resource destroying map [13] in the resource theory of coherence. From it we can define another set of operations named coherence non-activating operations (CNAO):

\[
\text{CNAO} = \{ O \in \text{CPTP} | \Delta \circ O = O \circ \Delta \}.
\]

CNAO is not a subset of MIO so it cannot be specified as the set of free operations.

Provided the incoherent states and incoherent operations, we need to find coherence measure. We can verify that the set of incoherent states is convex. Then applying the conclusion in Subsection A1, we define the relative entropy of coherence and robustness of coherence. The relative entropy of coherence $C_{rel}$ is

\[
C_{rel}(\rho) = S(\rho | \mathcal{I}) = \inf_{\sigma \in \mathcal{I}} S(\rho | \sigma) = S(\rho | \Delta(\rho)) = S(\Delta(\rho)) - S(\rho),
\]
where \( S(\rho) = -\text{tr}(\rho \log \rho) \) is the von-Neumann entropy. The robustness of coherence \( R \) is
\[
R(\rho) = \min_{s \in D(\mathcal{H})} s,
\]
\[
s.t. \quad \frac{\rho + ss'}{1 + s} \in \mathcal{I}, s \geq 0.
\]  
(A20)

With these two measures we can study the interconversion of states under different sets of incoherent operations and explore the quantum advantage shown in coherence.

3. Channel Resource Theory

Previous two subsections discuss the state resource theory. We can also establish a resource theory framework for the channels. Channel is the completely positive and trace preserving map on \( \mathcal{D}(\mathcal{H}) \). In a channel resource theory, the resource is a functional mapping a channel instead of a state to the non-negative number. We need to specify the set of free channels and the set of free super-channels as the analogy of the set of free states and the set of free operations. The super-channel transforms one channel to another just like the operation acting on the states. We can choose RNG super-channels as the free super-channels when we establish a channel resource theory to avoid providing a specific representation of super-channels.

Similar with the state resource theory, we need to provide the resource measure. One approach to discussing resource measure is transforming the channel to its Choi-state representation, then characterize the resource of Choi-state with a specific representation of super-channels.

A concrete example of channel resource theory is entanglement breaking (EB) channel [10]. A channel \( \mathcal{N} \) is entanglement breaking if and only if its Choi-state \( \Phi_\mathcal{N} \) is separable. At the same time, any EB channel has the form:
\[
\mathcal{N} = \sum_i \sigma_i \text{tr}(M_i \rho),
\]  
(A22)

where \( \sigma_i \) is a density operator and \( \{ M_i \} \) form a POVM. That is, any EB channel can be realized by measuring the state with a POVM, and then prepare a state based on the measurement result. The set of EB channels is convex. We can take this set to be free channels and quantify the resource of a channel relative to this set.

Appendix B: Partial Replacement of Quantum Operation

Here, we discuss the scenario of replacing part of a quantum operation with classical processing in detail. For simplicity, we only consider the case of qcCRO. First, we focus on the case where the evolution of a subsystem originating from a large evolution on the overall system can be viewed as a CPTP map.

Consider a system composed of two subsystems, \( \mathcal{H}_a \) and \( \mathcal{H}_b \), as shown in Figure 6, where the quantum state of the whole system, \( \mathcal{H}_{ab} = \mathcal{H}_a \otimes \mathcal{H}_b \), is a product state. The quantum evolution on the whole system is given by a unitary operation \( U \). If we focus on the subsystem \( \mathcal{H}_b \) and ignore subsystem \( \mathcal{H}_a \), the equivalent action on \( \mathcal{H}_b \) is given by
\[
O(\rho_b) = \text{tr}_a(U \rho_a \otimes \rho_b U^\dagger).
\]  
(B1)
If we know the initial state of $\rho_a$ and the unitary $U$, we can obtain the form of $O$. Here, we set specific values for $\rho_a$ and $U$ so that $O$ is a qcCRO. Then, $O$ can be moved after the computational basis measurement and replaced by classical processing $O_c$.

Before classical replacement, the final state on system $\mathcal{H}_{ab}$ is

$$\rho'_{ab} = \sum_i \text{tr}_b[U(\rho_a \otimes \rho_b)U^\dagger |i\rangle \langle i| \otimes |i\rangle \langle i|],$$

which exhibits correlation between subsystem $\mathcal{H}_a$ and $\mathcal{H}_b$. After classical replacement, the interaction between $\mathcal{H}_a$ and $\mathcal{H}_b$ is eliminated, so there is no correlation between two subsystems in the end. In this sense, we conclude that replacing a CRO only guarantees the subsystem it acts on is unchanged while the other systems and the correlation between the subsystem and the rest might change.

In Figure 7, we show an example replacing a CRO while maintaining the state of the whole system after classical replacement. The two subsystems $\mathcal{H}_a$ and $\mathcal{H}_b$ are both qubits with computational basis as $Z$ basis. The initial state on the whole system is $\rho_a \otimes \rho_b$ and the interaction is a control unitary from $\mathcal{H}_a$ to $\mathcal{H}_b$. The action on $\mathcal{H}_a$ after ignoring $\mathcal{H}_b$ is a dephasing channel, for any state $\rho_a \in D(\mathcal{H}_a)$,

$$O(\rho_a) = (1 - p)\rho_a + p\rho_a Z,$$

where $p$ is dephasing rate depending on the interaction and initial state on $\mathcal{H}_b$. Here, $O$ is a qcCRO and can be moved after computational basis measurement. The corresponding classical processing $O_c$ is identity, mapping a classical bit to itself. To ensure the state on $\mathcal{H}_{ab}$ does not change before and after classical replacement, we add a classical-quantum control from subsystem $a$ to subsystem $b$. In this way, not only the subsystem the CRO acts on is unchanged, the whole system is unchanged as well.

It is worth noting that Figure 7 shows an example where we can perform classical replacement without knowing the concrete form of a CRO. In this example, the dephasing rate $p$ is unknown if the initial state on subsystem $\mathcal{H}_b$ is
uncharacterized. Nevertheless, for any value of $p$, $O$ is related to a fixed classical processing so that the replacement can be implemented.

In the discussion above, we restrict the initial state to be a product state so that the evolution on the whole system reduces to a CPTP map when we look at the subsystem. However, if the initial state is entangled, the evolution on a subsystem may not be able to be expressed as a CPTP map, as shown in Figure 8. The classical replacement in this case is different from replacing a quantum operation, which is another interesting problem.

![Figure 8](image)

**FIG. 8.** Given a quantum system $\mathcal{H}_a \otimes \mathcal{H}_b$ with an entangled initial state $\rho_{ab}$, the interaction between two subsystems is given by $U$. Due to the quantum correlation of the initial state, $U$ cannot reduce to a CPTP map when focusing on subsystem $b$.

The situation where a quantum evolution cannot reduce to a CPTP map on a subsystem comes from the violation of the Markovian assumption. Another example of non-Markovian evolution on the subsystem is shown in Figure 9. With a product initial state $\rho_a \otimes \rho_b$ on system $\mathcal{H}_a \otimes \mathcal{H}_b$, two unitary evolutions on the whole system, $U_1$ and $U_2$, can reduce to two qcCROs on subsystem $b$, $O_1$ and $O_2$, respectively. However, the concatenation of two unitary evolutions, $U_2 \circ U_1$, cannot reduce to $O_2 \circ O_1$. An interesting question for future research is that when $U_2 \circ U_1$ can reduce to a qcCRO in such a case. A positive example is when $U_1$ and $U_2$ are both CNOT gates. Here, $U_2 \circ U_1$ cannot reduce to $O_2 \circ O_1$ but can reduce to an identity operation, which is a qcCRO.

![Figure 9](image)

**FIG. 9.** Given a quantum system $\mathcal{H}_a \otimes \mathcal{H}_b$ with a product initial state, $\rho_a \otimes \rho_b$, the unitary evolutions, $U_1$ and $U_2$, can both reduce to qcCROs $O_1$ and $O_2$ on subsystem $b$ respectively. The concatenation of two unitary evolutions, $U_2 U_1$, cannot reduce to $O_2 \circ O_1$, as in general the evolution on subsystem $b$ is non-Markovian. Interestingly, in some cases $U_2 U_1$ can reduce to a qcCRO, $O$, while the general case remains an open problem.
Appendix C: Proof of Theorem 1

Here, we provide a proof of Theorem 1. We prove two lemmas to verify the equivalence of the second and the third criteria, and the equivalence of the first and the second, respectively. The notations are the same in the main-text.

1. Proof of the Equivalence of the Second and the Third Criteria

We start with proving Lemma 3 as shown below.

**Lemma 3.** \( \exists O' \in \text{CPTP}, \Delta \circ O = \Delta \circ O' \circ \Delta \Rightarrow \Delta \circ O = \Delta \circ O \circ \Delta. \)

*Proof.* If operation \( O \) satisfies \( \exists O' \in \text{CPTP}, \Delta \circ O = \Delta \circ O' \circ \Delta \), then \( \forall i, j \in [d], \)
\[
\Delta \circ O(i|i\rangle\ranglej) = \Delta \circ O'(i|i\rangle\ranglej) \quad \Leftrightarrow \quad \sum_{kl} \delta_{kl} O_{ik, jl} |k\rangle\langle k| = \sum_{kl} \delta_{kl} O'_{ik, jl} |k\rangle\langle k|
\]
\[
\Leftrightarrow \quad \sum_{k} O_{ik, jk} |k\rangle\langle k| = \delta_{ij} \sum_{k} O'_{ik, ik} |k\rangle\langle k|,
\]
where \( O_{ik,jl} = \text{tr}(|i\rangle\langle i| O(|j\rangle\ranglej)) \). Set \( i = j \) we get for any \( i,k, O'_{ik,ik} = O_{ik,ik}. \) If \( i \neq j, O_{ik,jk} = 0. \) Thus, for any \( i,j, \)
\[
\Delta \circ O(i|i\rangle\ranglej) = \delta_{ij} \sum_{k} O_{ik, ik} |k\rangle\langle k| = \Delta \circ O \circ \Delta(i|i\rangle\ranglej).
\] (C2)
That means \( \Delta \circ O = \Delta \circ O \circ \Delta. \) \( \square \)

From Lemma 3 we obtain that for qcCRO, the third criterion implies the second. The opposite is also true as the second criterion means that there exists \( O' = O, \Delta \circ O = \Delta \circ O' \circ \Delta. \) As a consequence, the two criteria are equivalent for qcCRO. The proof is similar for cqCRO and qqCRO, which is omitted here.

2. Proof of the Equivalence of the First and the Second Criteria

For the next proof, we mainly prove the equivalence of the first and second criteria for qcCRO, while briefly discuss the cases of other two CROs. We begin with defining an extension of qcCRO and qcCROext as qcCROext2 and propose Lemma 4.

Given any input state \( \rho \in \mathcal{D}(H) \), any quantum channel \( O \), we perform the PVM, \( M_1 = \{P_n\} \), on \( O(\rho) \) to get the measurement result \( s \). We call \( O \) a qcCROext2 if we can get the same result for any same input by measuring \( \rho \) with PVM, \( M_2 = \{Q_m\} \) followed with a classical processing \( O_c \) as shown in Figure 10. The two measurements \( M_1 \) and \( M_2 \) are predetermined and the coarse-graining dephasing operations for them are denoted as \( \Delta' = \text{tr}(P_n)P_n \) and \( \Delta'' = \text{tr}(Q_m)Q_m \), respectively.

In general, the two measurements \( M_1, M_2 \) are different and \( \Delta'' \) does not commute with \( \Delta' \). The case of qcCROext2 reduces to qcCROext when \( M_1 = M_2, \) and further reduces to qcCRO if \( M_1 \) and \( M_2 \) are both computational basis measurement. Similar with qcCRO and CQROext, we provide a necessary and sufficient criterion for qcCROext2 as shown in Lemma 4. Then we can express the set of qcCROext2 as
\[
\text{qcCROext2} = \{O \in \text{CPTP} | \Delta' \circ O = \Delta' \circ O \circ \Delta''\}.
\] (C3)

**Lemma 4.** For any CPTP map \( O \in \text{qcCROext2} \) \( \Leftrightarrow \Delta' \circ O = \Delta' \circ O \circ \Delta''. \)

*Proof.* We first find the corresponding classical processing for each operation satisfying Eq. (C3), and then prove that any operation outside Eq. (C3) cannot.

For any operation \( O \) in Eq. (C3), the measurement result of \( M_1 \) on \( O(\rho) \) is equivalent to the measurement result of \( \Delta' \circ O \circ \Delta''(\rho) \) as shown in Figure 10. Here, we utilise the measurement result does not change if inserting a dephasing operation before measurement and \( \Delta' \circ O = \Delta' \circ O \circ \Delta''. \) As
\[
\Delta' \circ O \circ \Delta''(\rho) = \sum_{n,m} \text{tr}(Q_m \rho) \text{tr}(P_n O(Q_m)) P_n,
\] (C4)
the measurement result \( s \) of \( M_1 \) on \( O(\rho) \) takes the value \( n \) with probability \( \sum_m \text{tr}(Q_m \rho) \text{tr}(P_n O(Q_m)) \). It is equivalent to measure \( \rho \) with \( M_2 \) to get \( m \) with probability \( \text{tr}(Q_m \rho) \), and then transfer \( m \) to \( n \) with probability \( \text{tr}(P_n O(Q_m)) \).
We can view \( \text{tr}(Q_m \rho) \) as an initial probability distribution and \( \text{tr}(P_n O(Q_m)) \) as a stochastic matrix of the classical processing. Then we find the corresponding classical processing for any operation satisfying Eq. (C3) and complete the proof of the first step.

For the second step, we show that for any operation \( O \) outside Eq. (C3), there exists two different states, \( \rho_1 \) and \( \rho_2 \) satisfying \( \Delta''(\rho_1) = \Delta''(\rho_2) \), while \( M_1 \) on \( O(\rho_1) \) and \( O(\rho_2) \) can output different probability distributions. Note that \( \Delta''(\rho_1) = \Delta''(\rho_2) \) implies PVM \( M_2 \) cannot distinguish the two. Then the classical processing only outputs the same results for these two inputs. That means the quantum operation and classical processing cannot output the same result for some state, then we prove \( O \) cannot be replaced by classical processing.

If an operation \( O \) does not satisfy Eq. (C3), then \( \Delta' \circ O \neq \Delta' \circ O \circ \Delta'' \). That means there exists a state \( \sigma \) satisfying

\[
\Delta' \circ O(\sigma) \neq \Delta' \circ O(\Delta''(\sigma)).
\]

Now we take \( \rho_1 = \sigma, \rho_2 = \Delta''(\sigma) \). Obviously, \( \Delta''(\rho_1) = \Delta''(\rho_2) \). As \( \Delta'(O(\rho_1)) \neq \Delta'(O(\rho_2)) \), the measurement results of \( M_1 \) on \( O(\rho_1) \) and \( O(\rho_2) \) are different. Then we complete the second step of the proof.

The equivalence of the first and the second criteria for qcCRO is proved as a special case of Lemma 4. For cqCRO and qqCRO, we simply discuss the scenario when the state preparation and measurement are performed under computational basis without loss of generality. For any operation \( O \) satisfying \( O \circ \Delta = \Delta \circ O \circ \Delta \), and classical input \( i \), the output state is \( \sum_k O_{ik,ik} |k\rangle \langle k| \). That is first transforming \( i \) into \( k \) with probability \( O_{ik,ik} \) classically, and then preparing the state. Reversely, the state preparation is implemented under computational basis, so a cqCRO can only output incoherent states for classical inputs. That means a cqCRO is a MIO and satisfies Eq. (4), implying the equivalence of the first and second criteria for cqCRO.

Similar for a qqCRO, \( O \), we can find the corresponding classical processing with stochastic matrix \( O_{ik,ik} \). If an operation \( O \) does not satisfy \( O = \Delta \circ O \circ \Delta \), then \( \exists \sigma, O(\sigma) \neq \Delta \circ O \circ \Delta(\sigma) \). If \( O(\sigma) \neq \Delta \circ O(\sigma) \), then \( O(\sigma) \) is a coherent state while state preparation can only output incoherent states, which means \( O \) cannot be replaced. If \( O(\sigma) = \Delta \circ O(\sigma) \), then \( \Delta \circ O(\sigma) \neq \Delta \circ O \circ \Delta(\sigma) \), implying \( O \) is not a qqCRO. As any qqCRO is a qcCRO, in this case \( O \) cannot be replaced neither. Here, we conclude that qqCRO is fully characterized by Eq. (5) and complete the whole proof.

### Appendix D: Proof of Theorems and Lemmas in Section IV

#### 1. Proof of Lemma 1

*Proof.* The equivalence between Eq. (18) and Eq. (19) can be directly verified by the definition. The equivalence of Eq. (18) and Eq. (20) comes from the one-to-one correspondence of a channel and its Choi-state. We further prove the equivalence between Eq. (20) and Eq. (21), Eq. (20) and Eq. (22) respectively to prove Lemma 1, which comes from the idempotence of the dephasing operation.
Notice that $\Delta \circ \Delta = \Delta$, we have
\[
\Delta \circ \frac{\mathcal{N} + s \mathcal{M}}{1 + s} = \Delta \circ \frac{\Delta \circ \mathcal{N} + s \Delta \circ \mathcal{M}}{1 + s}.
\] (D1)

Assuming the minimization of Eq. (20) gives $s_1$ and that of Eq. (21) gives $s_2$, that means
\[
\Delta \circ \frac{\mathcal{N} + s_1 \mathcal{M}}{1 + s_1} = \Delta \circ \frac{\Delta \circ \mathcal{N} + s_1 \Delta \circ \mathcal{M}}{1 + s_1} \circ \Delta.
\] (D2)

Combining Eq. (D1) with Eq. (D2) we get
\[
\Delta \circ \frac{\Delta \circ \mathcal{N} + s_1 \Delta \circ \mathcal{M}}{1 + s_1} = \Delta \circ \frac{\Delta \circ \mathcal{N} + s_1 \Delta \circ \mathcal{M}}{1 + s_1} \circ \Delta.
\] (D3)

From the minimization of robustness we know $s_2 \leq s_1$. Reversely, combining Eq. (D1) with Eq. (D3) while substituting $s_1$ with $s_2$, we can prove $s_1 \leq s_2$. Then $s_1 = s_2$ implying the equivalence between Eq. (20) and Eq. (21). Similarly, we assume the minimization of Eq. (22) gives $s_3$ and from $\Delta \circ \Delta = \Delta$ we have
\[
\Delta \circ \frac{\mathcal{N} + s \mathcal{M}}{1 + s} = \Delta \circ \frac{\Delta \circ \mathcal{N} + s \mathcal{M}}{1 + s}.
\] (D4)

Combining Eq. (D2) with Eq. (D4) we get
\[
\Delta \circ \frac{\Delta \circ \mathcal{N} + s_1 \mathcal{M}}{1 + s_1} = \Delta \circ \frac{\Delta \circ \mathcal{N} + s_1 \mathcal{M}}{1 + s_1} \circ \Delta.
\] (D5)

Then we get $s_3 \leq s_1$. Following the similar approach we can also prove $s_1 \leq s_3$ implying $s_1 = s_3$. Thus, we prove the equivalence between Eq. (20) and Eq. (22), and complete the whole proof. $\Box$

2. **Proof of Lemma 2**

*Proof.* First, we prove the monotonicity of robustness of irreplaceability. For any free super operation $\Lambda \in \mathcal{F}$, suppose $\mathcal{M}$ is the channel achieving the minimal value of $s$, s.t., $\frac{\mathcal{N} + s \mathcal{M}}{1 + s} \in \text{qcCRO}$. Then
\[
\Lambda \left( \frac{\mathcal{N} + s \mathcal{M}}{1 + s} \right) = \frac{\Lambda(\mathcal{N}) + s \Lambda(\mathcal{M})}{1 + s} \in \text{qcCRO}.
\] (D6)

As $\Lambda(\mathcal{M})$ is also a CPTP channel, from the definition we see that $R(\Lambda(\mathcal{N})) \leq s = R(\mathcal{N})$.

The convexity of robustness comes from the convexity of qcCRO. Given a set of channels $\{\mathcal{N}_i\}$ where $i$ belongs to an index set $\mathcal{I}$, assuming $s_i = R(\mathcal{N}_i)$ and $\mathcal{M}_i$ achieves the minimal value, that is,
\[
\mathcal{M}_i' = \frac{\mathcal{N}_i + s_i \mathcal{M}_i}{1 + s_i} \in \text{qcCRO}.
\] (D7)

We set $\mathcal{M} = \sum_i p_i \mathcal{M}_i' / s = \sum_i p_i s_i$. Due to the convexity of qcCRO,
\[
\sum_i p_i \mathcal{N}_i + s \mathcal{M} = \sum_i p_i (1 + s_i) \mathcal{M}_i' \in \text{qcCRO}.
\] (D8)

From the definition of the robustness we get $R(\sum_i p_i \mathcal{N}_i) \leq s = \sum_i p_i R(\mathcal{N}_i)$. $\Box$

3. **Proof of Theorem 2**

*Proof.* The proof idea is transforming the minimization problem in robustness of irreplaceability to a dual problem. The dual problem is a maximization problem, where the objective function corresponds to the performance of a channel in the non-local game. For further elaboration, we clarify the notations. The Choi-state $\mathbb{I} \otimes \mathcal{N}(|\Phi^+ \rangle \langle \Phi^+|)$ of a channel $\mathcal{N}$ is a bipartite state and we label the two parties with 0 and 1, respectively. That is, $\mathbb{I} \otimes \mathcal{N}(|\Phi^+ \rangle \langle \Phi^+|) \in \mathcal{H}_0 \otimes \mathcal{H}_1$. The
computational basis of each party $\mathcal{H}_k$ is denoted as $\{|i_k\rangle, i \in [d]\}$, $k = 0, 1$. With respect to the computation basis, we define two partial dephasing operations, such that $\forall \rho \in \mathcal{D}(\mathcal{H}_0 \otimes \mathcal{H}_1)$,

$$\Delta_0 = \sum_{i=0}^{d-1} |i_0\rangle\langle i_0| \otimes \text{tr}_0(|i_0\rangle\langle i_0| \rho),$$  \hspace{1cm} (D9)$$

$$\Delta_1 = \sum_{j_1=0}^{d-1} \text{tr}_1(|j_1\rangle\langle j_1| \rho) \otimes |j_1\rangle\langle j_1|. $$ \hspace{1cm} (D10)

For the joint system $\mathcal{H}_0 \otimes \mathcal{H}_1$, we take the computational basis as $\{|i_0,j_1\rangle, i_0, j_1 \in [d]\}$ and represent the dephasing operation on $\mathcal{H}_0 \otimes \mathcal{H}_1$ as

$$\Delta_{01} = \sum_{i_0,j_1=0}^{d-1} \text{tr}(|i_0,j_1\rangle\langle i_0,j_1| \rho) |i_0,j_1\rangle\langle i_0,j_1|. $$ \hspace{1cm} (D11)

Now we take the second equivalent definition of robustness of irreplaceability, i.e., Eq. (19), and express the minimization problem as a conic optimization problem,

$$R(\mathcal{N}) = \min_{s,A} s, \hspace{1cm} \text{s.t. } A - \Delta_1(\Phi_{\mathcal{N}}) \in \text{cone}(\mathbf{S}),$$

$$A \in \text{cone}(\mathbf{F}'), \hspace{1cm} \text{tr}_1(A) = (1+s)\mathbb{I}_0,$$ \hspace{1cm} (D12)

where $\mathbf{S} = \{\Delta_1(\Phi_{\mathcal{N}}), \mathcal{N} \in \text{CPTP}\}, \mathbf{F}' = \{\Delta_1(\Phi_{\mathcal{N}}), \mathcal{M} \in \text{qcCRO}\}$, and cone() represents their unnormalized versions, which forms a convex cone. Note that $\Phi_{\Delta\mathcal{N}} = \Delta_1(\Phi_{\mathcal{N}})$. We use $\mathbb{I}_0 \in \mathcal{L}(\mathcal{H}_0)$ to denote the identity operator, where $\mathcal{L}(\mathcal{H}_0)$ represents the set of linear operators on $\mathcal{H}_0$. The Lagrangian function of this problem is

$$L(s, A, W, X, Y) = s - \langle W, A - \Delta_1(\Phi_{\mathcal{N}}) \rangle - \langle X, A \rangle - \langle Y, (1+s)\mathbb{I}_0 - \text{tr}_1(A) \rangle.$$ \hspace{1cm} (D13)

Here, $\langle \cdot, \cdot \rangle$ represents Hilbert-Schmidt inner product, with

$$\langle A, B \rangle = \text{tr}(A^\dagger B), \forall A, B \in \mathcal{L}(\mathcal{H}),$$ \hspace{1cm} (D14)

and $W \in (\text{cone}(\mathbf{S}))^*, X \in (\text{cone}(\mathbf{F}'))^*, Y \in \mathcal{L}(\mathcal{H}_0)$. The dual problem is

$$\max_{W,X,Y} \min_{s,A} \frac{L(s, A, W, X, Y)}{s,A} \hspace{1cm} \text{s.t. } \langle W, B \rangle \geq 0, \forall B \in \mathbf{S},$$

$$\langle X, C \rangle \geq 0, \forall C \in \mathbf{F'}.$$ \hspace{1cm} (D15)

We can change the form of Lagrangian function:

$$L(s, A, W, X, Y) = s(1 - \langle Y, \mathbb{I}_0 \rangle) - \langle W + X - Y \otimes \mathbb{I}_1, A \rangle + \langle W, \Delta_1(\Phi_{\mathcal{N}}) \rangle - \langle Y, \mathbb{I}_0 \rangle.$$ \hspace{1cm} (D16)

Note that $\langle Y, \text{tr}_1(A) \rangle = \langle Y \otimes \mathbb{I}_1, A \rangle$. Then the dual problem is equivalent to:

$$\max_{W,X,Y} \langle W, \Delta_1(\Phi_{\mathcal{N}}) \rangle - \langle Y, \mathbb{I}_0 \rangle,$$ \hspace{1cm} s.t. 1 - $\langle Y, \mathbb{I}_0 \rangle = 0$, $W + X - Y \otimes \mathbb{I}_1 = 0$, $\langle W, B \rangle \geq 0, \forall B \in \mathbf{S}$, $\langle X, C \rangle \geq 0, \forall C \in \mathbf{F'}$. \hspace{1cm} (D17)

The above can also be written as

$$\max_{W,X,Y} \text{tr}(W \Delta_1(\Phi_{\mathcal{N}})) - 1,$$ \hspace{1cm} s.t. $\text{tr}(Y) = 1$, $W = Y \otimes \mathbb{I}_1 - X$, $\text{tr}(WB) \geq 0, \forall B \in \mathbf{S}$, $\text{tr}(WC) \leq \text{tr}(Y \otimes \mathbb{I}_1 C) = \text{tr}(Y) = 1, \forall C \in \mathbf{F'}$. \hspace{1cm} (D18)
By taking $X = \frac{\ln 2}{27}$, $Y = \frac{\ln 2}{2}$, $W = \frac{\ln 2}{23}$, we can check that Slater’s conditions are fulfilled and hence the problem satisfies the strong duality [31]. Note that the principal problem is minimizing a convex function $s$ and $(Y, (1+s)I_0 - tr_1(A))$ is an affine function relative to $s$ and $A$. Also, we can see that the objective function is not related to $X$ and $Y$, hence we can rewrite the optimization as

$$R(N) = \max_{W \in (\text{cone}(S))^*} \text{tr}(W \Delta_1(\Phi_N)) - 1,$$

s.t. $\text{tr}(WB) \geq 0, \forall B \in S$, $\text{tr}(WC) \leq 1, \forall C \in F'$. \hspace{1cm} (D19)

Note that for any qCRO $M$, $\Delta_1(\Phi_M) \in F'$, there always exists $W = \Delta_1(\Phi_M)$ satisfying the restriction and achieving the maximization. The corresponding robustness is 0 which is the value for any qCRO.

On the other side, the performance of channel $N$ in the non-local game is

$$p(N, \{\alpha_{ij}\}, \{\sigma_i\}, \{|j\rangle\}) = \sum_{i,j} \alpha_{ij} \text{tr}(N(\sigma_i)|j\rangle\langle j|)$$

$$= d \sum_{i,j} \alpha_{ij} \text{tr}(\Phi_N \sigma_i^T \otimes |j\rangle\langle j|)$$

$$= d \sum_{i,j} \alpha_{ij} \text{tr}(\Delta_1(\Phi_N) \sigma_i^T \otimes |j\rangle\langle j|)$$

$$= \text{tr}(W \Delta_1(\Phi_N)),$$

where $W = d \sum_{i,j} \alpha_{ij} \sigma_i^T \otimes |j\rangle\langle j| \in (\text{cone}(S))^*$. Therefore, by varying the strategy $\alpha_{ij}$ and $\sigma_i$, we can obtain any $W$ satisfying $\text{tr}(WB) \geq 0, \forall B \in S$ and $\text{tr}(WC) \leq 1, \forall C \in F'$. Comparing Eq. (26), Eq. (D19) and Eq. (D20), we relate the best advantage a quantum channel can provide to the robustness measure and complete the proof,

$$\max_{\{\alpha_{ij}\}, \{\sigma_i\}} \frac{\text{tr}(W \Delta_1(\Phi_N))}{\text{max}_{M \in \text{qCRO}} \text{tr}(W \Delta_1(\Phi_M))} = 1 + R(N).$$ \hspace{1cm} (D21)

\[\square\]

**Appendix E: Relative Entropy of Irreplaceability**

The relative entropy of irreplaceability is defined as follows.

**Definition 3 (Relative Entropy of Irreplaceability).** Given a channel $N \in \text{CPTP}$, the relative entropy of irreplaceability of $N$ is

$$C_{rel}(N) = \min_{M \in \text{qCRO}} D(\Phi_{\Delta_0N} | \Phi_{\Delta_0M}),$$ \hspace{1cm} (E1)

where $D(\rho| \sigma) = \text{tr}(\rho \log \rho) - \text{tr}(\rho \log \sigma)$ is the relative entropy from $\rho$ to $\sigma$.

The set of qCRO and $F' = \{\Phi_{\Delta_0M} | M \in \text{qCRO}\}$ are convex, hence Eq. (E1) is well-defined. This measure is different from the normal definition of the relative entropy measure, usually defined as the minimum relative entropy of a channel to the free channels, i.e., $\min_{M \in \text{qCRO}} D(\Phi_N | \Phi_M)$. Here, we adopt the Choi state of $\Delta \circ N$ instead of $\mathcal{N}$ as we concern about whether an operation followed with a measurement can be replaced by classical operations or not. The irreplaceability of a channel $\mathcal{N}$ should be the same as $\Delta \circ \mathcal{N}$ as no measurement can distinguish them. Also, we have proven that $R(N) = R(\Delta \circ \mathcal{N})$, showing that it is reasonable to consider the distance between $\Delta \circ \mathcal{N}$ and $\Delta \circ M$ instead of $\mathcal{N}$ and $M$.

The relative entropy of irreplaceability has an analytic expression, as given by the following lemma. The dephasing operations $\Delta_0, \Delta_1, \Delta_{01}$ are given by Eq. (D9), (D10) and (D11), respectively.

**Lemma 5.** The relative entropy of irreplaceability of a channel $\mathcal{N}$ equals to

$$C_{rel}(\mathcal{N}) = S(\Delta_{01}(\Phi_{\mathcal{N}})) - S(\Delta_1(\Phi_{\mathcal{N}})),$$ \hspace{1cm} (E2)

where $S(\rho) = -\text{tr}(\rho \log \rho)$ is the von-Neumann entropy.
Proof. For any qcCRO, \( \mathcal{M} \),

\[
\Phi_{\Delta \circ \mathcal{M}} = \Delta_1(\Phi_{\mathcal{M}}) = I \otimes (\Delta \circ \mathcal{M})(|\Phi^+\rangle \langle \Phi^+|)
\]

\[
= \frac{1}{d} \sum_{i,j} |i\rangle \langle i| \otimes (\Delta \circ \mathcal{M} \circ \Delta) |ij\rangle \langle jj|
\]

\[
= \frac{1}{d} \sum_{i} |i\rangle \langle i| \otimes \mathcal{M}(|i\rangle \langle i|)
\]

\[
= \Delta_0 \left( \frac{1}{d} \sum_{i} |i\rangle \langle i| \otimes \mathcal{M}(|i\rangle \langle i|) \right)
\]

\[
= \Delta_0(1 \otimes (\Delta \circ \mathcal{M})(|\Phi^+\rangle \langle \Phi^+|))
\]

\[
= \Delta_0(\Phi_{\mathcal{M}}).
\]

Then,

\[
D(\Delta_1(\Phi_N)||\Delta_1(\Phi_{\mathcal{M}})) = -S(\Delta_1(\Phi_N)) - tr(\Delta_1(\Phi_N) \log \Delta_1(\Phi_{\mathcal{M}}))
\]

\[
= -S(\Delta_1(\Phi_N)) - tr(\Delta_1(\Phi_N) \log \Delta_{01}(\Phi_{\mathcal{M}}))
\]

\[
= -S(\Delta_1(\Phi_N)) - tr(\Delta_1(\Phi_N) \Delta_{01}(\log \Phi_{\mathcal{M}}))
\]

\[
= -S(\Delta_1(\Phi_N)) - tr(\Delta_{01}(\Phi_{\mathcal{M}}))
\]

\[
= S(\Delta_{01}(\Phi_N)) - S(\Delta_1(\Phi_N)) + D(\Delta_{01}(\Phi_N)||\Delta_{01}(\Phi_{\mathcal{M}}))
\]

\[
\geq S(\Delta_{01}(\Phi_N)) - S(\Delta_1(\Phi_N)).
\]

In the last line, the state \( \Phi_{\mathcal{M}} = \Delta_{01}(N) \) saturates the equality. This completes the proof. \(\square\)

From the proof of Lemma 5, we can see that for any qcCRO, \( \mathcal{M} \), the partially dephased state \( \Delta_1(\Phi_{\mathcal{M}}) \) is an incoherent state on \( H_0 \otimes H_1 \). Correspondingly, Eq. (E2) is the relative entropy of coherence of \( \Delta_1(\Phi_N) \).

In the discussion below, we take a subset of RNG super operations to be the set of free super operations \( \mathcal{F} \),

\[
\mathcal{F} = \{ \Lambda \in \text{RNG}|\Delta \circ \Lambda(N) = \Lambda(\Delta \circ N), \forall N \in \text{CPTP} \}.
\]

The set \( \mathcal{F} \) is not empty as the identity super operation, \( \Lambda(N) = N \), belongs to \( \mathcal{F} \). It is also a convex set as \( \Lambda_1, \Lambda_2 \in \mathcal{F} \Rightarrow p\Lambda_1 + (1 - p)\Lambda_2 \in \mathcal{F}, \forall 0 \leq p \leq 1 \). Now we show that the relative entropy of irreplaceability has the properties of non-increasing under mixing and monotonicity under \( \mathcal{F} \), as shown by the following lemma.

Lemma 6. The relative entropy of irreplaceability in Eq. (E1), \( C_{\text{rel}}(N) \), has the following properties:

1. **Non-increasing under mixing**: Given an index set \( \mathcal{T} \), \( \forall N_{i} \in CPTP, \forall \{p_i\}_{i \in \mathcal{T}} \) such that \( p_i \geq 0, \sum_{i \in \mathcal{T}} p_i = 1 \),

\[
C_{\text{rel}}\left( \sum_{i} p_i N_i \right) \leq \sum_{i} p_i C_{\text{rel}}(N_i).
\]

2. **Monotonicity under a free super operation**: \( \forall N \in CPTP, \forall \Lambda \in \mathcal{F} \),

\[
C_{\text{rel}}(\Lambda(N)) \leq C_{\text{rel}}(N).
\]

Proof.

\[
C_{\text{rel}}\left( \sum_{i} p_i N_i \right) = \min_{\mathcal{M} \in \text{qcCRO}} D\left( \sum_{i} p_i \Delta_1(\Phi_{N_i})||\Delta_1(\Phi_{\mathcal{M}}) \right)
\]

\[
\leq D\left( \sum_{i} p_i \Delta_1(\Phi_{N_i})\| \sum_{i} p_i \Delta_1(\Phi_{\mathcal{M}_i}) \right)
\]

\[
\leq \sum_{i} p_i D(\Delta_1(\Phi_{N_i})||\Delta_1(\Phi_{M_i}))
\]

\[
= \sum_{i} p_i C_{\text{rel}}(N_i).
\]
where $\mathcal{M}_i$ is the CPTP channel minimizing the value $D(\Delta_i(N_i) \| \Delta_i(M_i))$. The inequality of the third line comes from the joint convexity of relative entropy. For any free super operation $\Lambda \in \mathcal{F}$,

$$C_{\text{rel}}(\Lambda(N)) = \min_{\Lambda \in \text{CPTP}} D(\Phi_{\Delta \odot \Lambda(N)} \| \Phi_{\Delta \odot \Lambda(M)})$$

$$\leq \min_{\Lambda \in \text{CPTP}} D(\Phi_{\Delta \odot \Lambda(N)} \| \Phi_{\Delta \odot \Lambda(M)})$$

$$= \min_{\Lambda \in \text{CPTP}} D(\Phi_{\Delta \odot \Lambda(N)} \| \Phi_{\Delta \odot \Lambda(M)})$$

$$\leq \min_{\Lambda \in \text{CPTP}} D(\Phi_{\Delta \odot \Lambda(N)} \| \Phi_{\Delta \odot \Lambda(M)})$$

$$= C_{\text{rel}}(N).$$

The inequality of the fourth line comes from data processing inequality.

\[ \square \]

[1] C. H. Bennett and G. Brassard, in Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing (IEEE Press, New York, 1984), pp. 175–179, URL https://doi.org/10.1103/PhysRevLett.67.661.

[2] A. K. Ekert, Phys. Rev. Lett. 67, 661 (1991), URL http://link.aps.org/doi/10.1103/PhysRevLett.67.661.

[3] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, et al., Nature 574, 505 (2019), ISSN 1476-4687, URL https://doi.org/10.1038/s41586-019-1666-5.

[4] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, et al., Science 370, 1460 (2020), https://www.science.org/doi/pdf/10.1126/science.abe8770, URL https://www.science.org/doi/10.1126/science.abe8770.

[5] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, et al., Phys. Rev. Lett. 127, 180501 (2021), URL https://link.aps.org/doi/10.1103/PhysRevLett.127.180501.

[6] H.-K. Lo and H. F. Chau, Science 283, 2050 (1999), ISSN 0036-8075, https://science.sciencemag.org/content/283/5410/2050.full.pdf, URL https://science.sciencemag.org/lookup/doi/10.1126/science.283.5410.2050.

[7] P. W. Shor and J. Preskill, Phys. Rev. Lett. 85, 441 (2000), URL https://link.aps.org/doi/10.1103/PhysRevLett.85.441.

[8] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, Rev. Mod. Phys. 92, 025002 (2020), URL https://link.aps.org/doi/10.1103/RevModPhys.92.025002.

[9] M. A. Nielsen and I. Chuang, Quantum computation and quantum information (2002).

[10] M. Horodecki, P. W. Shor, and M. B. Ruskai, Rev. Math. Phys. 15, 629 (2003), https://doi.org/10.1142/S0129055X03001709.

[11] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, et al., Nat. Rev. Phys 3, 625 (2021), ISSN 2522-5820, https://doi.org/10.1038/s42254-021-00348-9.

[12] H.-Y. Huang, R. Kueng, and J. Preskill, Nat. Phys. 16, 1050 (2020), ISSN 1745-2481, URL https://doi.org/10.1038/s41567-020-0932-7.

[13] Z.-W. Liu, X. Hu, and S. Lloyd, Phys. Rev. Lett. 118, 060502 (2017), URL https://link.aps.org/doi/10.1103/PhysRevLett.118.060502.

[14] G. Gour, Phys. Rev. A 95, 062314 (2017), URL https://link.aps.org/doi/10.1103/PhysRevA.95.062314.

[15] D. Rosset, F. Buscemi, and Y.-C. Liang, Phys. Rev. X 8, 021033 (2018), URL https://link.aps.org/doi/10.1103/PhysRevX.8.021033.

[16] and others, J. of Math 58, 5 (1936).

[17] J. E. Savage, J. ACM 19, 660 (1972).

[18] E. S. Santos, Proc. Am. Math. 22, 704 (1969), ISSN 00029939, 10886826, URL http://www.jstor.org/stable/2037463.

[19] J. Gill, SIAM J. Comput. 6, 675 (1977), https://doi.org/10.1137/0206049, URL https://doi.org/10.1137/0206049.

[20] T. Baumgratz, M. Cramer, and M. B. Plenio, Phys. Rev. Lett. 113, 140401 (2014), URL https://link.aps.org/doi/10.1103/PhysRevLett.113.140401.

[21] E. Chitambar and G. Gour, Phys. Rev. Lett. 117, 030401 (2016), URL https://link.aps.org/doi/10.1103/PhysRevLett.117.030401.

[22] E. Chitambar and G. Gour, Rev. Mod. Phys. 91, 025001 (2019), URL https://link.aps.org/doi/10.1103/RevModPhys.91.025001.

[23] Y. Liu and X. Yuan, Phys. Rev. Research 2, 012035 (2020), URL https://link.aps.org/doi/10.1103/PhysRevResearch.2.012035.

[24] G. Gour and M. M. Wilde, Phys. Rev. Research 3, 023096 (2021), URL https://link.aps.org/doi/10.1103/PhysRevResearch.3.023096.

[25] R. Takagi and B. Regula, Phys. Rev. X 9, 031053 (2019), URL https://link.aps.org/doi/10.1103/PhysRevX.9.031053.
[26] X. Yuan, Y. Liu, Q. Zhao, B. Regula, J. Thompson, and M. Gu, NPJ Quantum Inf. 7, 108 (2021), ISSN 2056-6387, URL https://doi.org/10.1038/s41534-021-00444-9.

[27] R. Uola, T. Kraft, J. Shang, X.-D. Yu, and O. Gühne, Phys. Rev. Lett. 122, 130404 (2019), URL https://link.aps.org/doi/10.1103/PhysRevLett.122.130404.

[28] J. Eisert, Phys. Rev. Lett. 127, 020501 (2021), URL https://link.aps.org/doi/10.1103/PhysRevLett.127.020501.

[29] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. 78, 2275 (1997), URL https://link.aps.org/doi/10.1103/PhysRevLett.78.2275.

[30] F. G. S. L. Brandão and G. Gour, Phys. Rev. Lett. 115, 070503 (2015), URL https://link.aps.org/doi/10.1103/PhysRevLett.115.070503.

[31] S. Boyd, S. P. Boyd, and L. Vandenberghe, Convex optimization (Cambridge university press, 2004).