Universal limitation of quantum information recovery: symmetry versus coherence

Hiroyasu Tajima

1. Department of Communication Engineering and Informatics,
University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo, 182-8585, Japan and
2. JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan

Keiji Saito

Department of Physics, Keio University, 3-14-1 Hiyoshi, Yokohama, 223-8522, Japan

Quantum information is scrambled via chaotic time evolution in many-body systems. The recovery of initial information embedded locally in the system from the scrambled quantum state is a fundamental concern in many contexts. From a dynamical perspective, information recovery can measure dynamical instability in quantum chaos, fault-tolerant quantum computing, and the black hole information paradox. This article considers general aspects of quantum information recovery when the scrambling dynamics have conservation laws due to Lie group symmetries. Here, we establish fundamental limitations on the information recovery from scrambling dynamics with arbitrary Lie group symmetries. We show universal relations between information recovery, symmetry, and quantum coherence, which apply to many physical situations. The relations predict that the behavior of the Hayden-Preskill black hole model changes qualitatively under the assumption of the energy conservation law. Consequently, we can rigorously prove that under the energy conservation law, the error of the information recovery from a small black hole remains unignorably large until it completely evaporates. Moreover, even when the black hole is very large, the recovery of information thrown into the black hole is not completed until most of the black hole evaporates. The relations also provide a unified view of the symmetry restrictions on quantum information processing, such as the approximate Eastin-Knill theorem and the Wigner-Araki-Yanase theorem for unitary gates.

I. INTRODUCTION

Information locally embedded in many-body isolated systems generically diffuses over the entire system due to its chaotic dynamical motion. Recently, a significant amount of research has been conducted on the question of how to accurately recover the initial information from the scrambled state using a fixed protocol without any partial knowledge of the input information. In classical systems, a chaotic motion with a positive Lyapunov exponent scrambles the phase space without limitations. Hence, the information recovery is practically hopeless due to the sensitivity against a small perturbation to the recovery protocol [1, 2]. However, in quantum systems, the scrambling nature becomes much more moderate than that in the classical case due to the existence of the Planck cell arising from Heisenberg’s uncertainty relation. Hence, information recovery in the quantum regime seems feasible. Fundamental questions here are how and how accurately one can recover the information. The quantum information recovery problems have provided a lot of surprises at a fundamental level. Additionally, they have brought practical importance in several fields, such as fault-tolerant quantum computation.

The quantum information recovery problem historically dates back to the information paradox of the black hole [3, 4]. In the 1970s, Hawking raises a question on the information loss of the thrown information into a black hole. In the classical picture, information leakage from a classical black hole is unlikely due to the no-hair theorem [5]. However, quantum black holes can release the quantum information via the Hawking radiation [3, 4, 7–11]. Hayden and Preskill developed a model based on the quantum information theory and have found a remarkable result, i.e., arbitrary k-qubit quantum data thrown into the black hole can be almost perfectly recovered by collecting only a few more than k-qubit information from the Hawking radiation [7]. In other words, quantum black holes act as informative mirrors. This finding is highly counterintuitive compared to the clas-
tical dynamical nature, and hence it has triggered a lot of studies \cite{8,11}. Furthermore, today the technique and concept of information recovery go beyond the black hole physics and commonly appear in various problems considering the recovery of quantum correlations, such as quantum chaos \cite{12}, fault tolerance of quantum information \cite{13} and measurement-induced phase transition \cite{14}. With recent development in experimental techniques related to quantum information, predictions of quantum information recovery are becoming verifiable. Various experimental setups on information recovery have been proposed, and several have been actually realized in the laboratory \cite{15–19}.

Since the seminal work of Hayden and Preskill, the typical theoretical framework for information recovery adopted thus far is to use the Haar random unitary without any conservation laws to represent scrambling dynamics, exception of several works in specific situations \cite{20–22}. However, we should note that in real many-body systems, information scrambling can occur in dynamics with conservation laws, such as energy and momentum conservations. Even for a black hole, which is considered the most chaotic system in the universe, the energy conservation law needs to be satisfied. Using the out-of-time-order correlation, an estimator of the degree of scrambling \cite{23,24}, one can show that isolated quantum many-body systems with nonlocal interactions demonstrate even fast scrambling phenomenon \cite{25–27}. Conservation laws also can play a critical role in the dynamics in typical isolated many-body quantum systems such as isolated cold atomic systems. With this backgrounds, it is vital to determine the universal effects of symmetries on the information recovery for the in-depth understanding of its quantum nature and also further applications.

In this work, we address this question by developing the techniques in resource theory of asymmetry \cite{28–38}. Consequently, we present several general limitations on information recovery when the scrambling dynamics possess Lie group symmetries. The limitations quantitatively state that when the scrambling has symmetries, significant inevitable errors occur in the recovery, and only large amounts of quantum fluctuation of the conserved quantities can mitigate these errors to a certain extent.

Since our technique does not require assumptions other than unitarity and symmetry of the scrambling dynamics, the established limitations can be applied to many important situations. An interesting application exists within black hole physics. In the information recovery from black holes, our results indicate that quantum coherence is required in the initial state of the black hole for accurate information recovery. Consequently, we demonstrate that when considering Hayden-Preskill’s black hole model with the energy conservation law, that law drastically limits the success rate of information recovery. Our results are summarized in two messages. First, when the size of both black hole and thrown object are comparable, the error of information recovery remains large until the black hole completely evaporates. Second, even when the black hole is much larger than the thrown object, information escapes very slowly, and a significant error in recovery remains until the black hole has almost evaporated. Furthermore, our theorems can explain several limitations on quantum information processing with symmetry \cite{34,35,39–47}. Examples include the approximate Eastin-Knill theorem on covariant codes \cite{39–43}. Our study demonstrates that limitations discussed independently so far can be derived from a single general theorem in a unified way.

This paper is organized as follows. In the section II we formulate a general setup of quantum information recovery from unitary dynamics with symmetry. In the section III we present the main results, i.e., the fundamental limitations on quantum information recovery. In the section IV we apply the main results to the Hayden-Preskill black hole model. In the section V we apply the results to the quantum information processing with symmetry. In the section VI we give a numerical check of the main results. Finally, in the Appendix, we provide basic tips of resource theory of asymmetry, and present the proof of the main results.

II. SETUP

A setup on the information recovery is introduced in a general form. As discussed later, the setup described here is directly applicable to various situations including black hole scrambling \cite{6–11}, error correcting codes \cite{39–43} and the implementation of quantum computation gates \cite{34,35,44–47}.

We consider four finite-level quantum systems $A$, $B$, $R_A$ and $R_B$, represented schematically in Fig. 2. The part $A$ is the system of interest with a mixed state $\rho_A$ as an initial state. Then, we make a purification between the system $A$ and $R_A$, the wave function of which is described as $|\psi_{AR_A}\rangle$. We assume that the initial state of the composite system $BR_B$ is pure state $|\phi_{BR_B}\rangle$, which is an entangled state. Through entanglement, the systems $R_A$ and $R_B$ have partial quantum information of the system $A$ and $B$, respectively. For this initial state, the unitary operation $U$ is applied on the systems $A$ and $B$, which scrambles the quantum information of the initial state. A main task in the information recovery problem is to recover the initial state $|\psi_{AR_A}\rangle$ with aid of partial information of the scrambled state. To this end, we suppose that the composite system $AB$ is either naturally or artificially divided into an accessible part $A'$ and the other part $B'$ after the unitary operation, where the Hilbert space of $AB$ and $A'B'$ are the same (see Fig. 2again). We then apply a recovery operation $R$ which is a completely positive and trace preserving (CPTP) map acting from $A'R_B$ to $A$ without touching $R_A$. Through this recovery operation, we try to recover the initial state $|\psi_{AR_A}\rangle$ as accurate as possible using the quantum information contained in the subsystems $A'$ and $R_B$. Following the
We now introduce two key quantities to describe the limitation of information recovery. While the conservation law for the total system is assumed, local conserved quantities can fluctuate. The first key quantity we focus on is the dynamical fluctuation associated with the quantum operation $\mathcal{E}$, i.e., a fluctuation of the change between the initial value of $X_A$ and the value of $X_{A'}$ after the quantum operation. The change of the values of the local conserved quantity depends on the initial state $\rho_A$. We characterize such fluctuation arising from the choice of the initial state, considering that the initial reduced density operator for the system $A$ can be decomposed as $\rho_A = \sum_j p_j \rho_j$ with weight $p_j$ satisfying $\sum_j p_j = 1$. Such a decomposition is not unique. While the linearity on the CPTP map guarantees that the decomposition reproduces the same output state on $A'$, i.e., $\mathcal{E}(\rho_A) = \sum_j p_j \mathcal{E}(\rho_j)$, each path from the density operator $\rho_j$ shows a variation on the change of local conserved quantities in general. Taking account of this property, we define the following quantity $A$ to quantify the dynamical fluctuation on the local conserved quantity for a given initial density operator:

$$A := \max_{\{p_j, \rho_j\}} \sum_j p_j |\Delta_j|,$$

$$\Delta_j := \left(\langle X_A \rangle_{\rho_j} - \langle X_{A'} \rangle_{\mathcal{E}(\rho_j)}\right) - \left(\langle X_A \rangle_{\rho_A} - \langle X_{A'} \rangle_{\mathcal{E}(\rho_A)}\right),$$

where $\langle \ldots \rangle_\rho := \text{Tr}(\ldots \rho)$, and the set $\{p_j, \rho_j\}$ covers all decompositions $\rho_A = \sum_j p_j \rho_j$. Note that the quantity $A$ is a function of the state $\rho_A$ and the CPTP map. When the systems $A$ and $B$ are identical to $A'$ and $B'$, respectively, and the unitary operator is decoupled between the systems as $U = U_A \otimes U_B$, the dynamical fluctuation is trivially zero. A finite value of the dynamical fluctuation is generated for a finite interaction between the systems. This is reflected from the fact that the global symmetry does not completely restrict the behaviour of the subsystem.

Another key quantity is quantum coherence. Following the standard argument in the resource theory of asymmetry, we employ the SLD-quantum Fisher information $\mathcal{F}_\rho (X)$ for the state family $\{e^{-iXt} \rho e^{iXt}\}_{t \in \mathbb{R}}$ to quantify the quantum coherence on $\rho$:

$$\mathcal{F}_\rho (X) := 4 \lim_{\epsilon \to 0} \frac{\text{Tr}(e^{-iX\epsilon} \rho e^{iX\epsilon}, \rho)^2}{\epsilon^2}.$$  

The quantum Fisher information is a good indicator of the amount of quantum coherence in $\rho$ with the basis of the eigenvectors of $X$. It is known that this quantity is directly connected to the amount of quantum fluctuation (see Appendix A) [59, 60]. We consider the quantum coherence contained inside the system $B$ as discussed below.
III. MAIN RESULTS

A. Fundamental limitation on information recovery

With the two key quantities introduced above, we establish two fundamental relations on the limitations of the information recovery. We note that the relations are obtained for general unitary operations with conservation laws, without assumptions such as the Haar random unitary.

The first relation on the limitation of the information recovery is described as follows [61]:

\[
\frac{A}{2(\sqrt{F} + 4\Delta_+)} \leq \delta, \tag{5}
\]

where \(F := F_{\rho BR_B}(X_B \otimes 1_{R_B})\) is the quantum coherence in the initial state of the system \(BR_B\). The quantity \(\Delta_+\) is a measure of possible change on the local conserved quantities, i.e., \(\Delta_+ := (D_{X_A} + D_{X_{A'}})/2\) where \(D_{X_A}\) and \(D_{X_{A'}}\) are the differences between the maximum and minimum eigenvalues of the operators \(X_A\) and \(X_{A'}\), respectively.

The inequality [5] shows a close relation between the recovery error (irreversibility), the dynamical fluctuation, and the quantum coherence. It shows that when the dynamical fluctuation is finite, perfect recovery is impossible. Moreover, high performance recovery is possible only when the quantum coherence sufficiently fills the initial state of \(BR_B\). Note that the dynamical fluctuation is generically finite, since the systems \(A\) and \(B\) interact with each other via the unitary operation. We show a specific example in Supplementary Materials Supp.II, where filling vast quantum coherence in \(BR_B\) actually makes the error \(\delta\) smaller than \(A/8\Delta_+\) and negligibly small.

The above inequality uses the quantum coherence \(F\) of the initial state of \(BR_B\). We can also establish another inequality with the quantum coherence of the final state, which is the second main relation [61]:

\[
\frac{A}{2(\sqrt{F_f} + \Delta_{\text{max}})} \leq \delta, \tag{6}
\]

where \(\Delta_{\text{max}} := \max_{\{p_j, \rho_j\}} \max_j |\Delta_j|\), and the set \(\{p_j, \rho_j\}\) covers all decompositions satisfying \(\rho = \sum_j p_j \rho_j\). The quantum coherence here is measured for the final state as \(F_f := F_{\sigma_{B'R_B}}(X_{B'} \otimes 1_{R_B})\), where the state \(\sigma_{B'R_B}\) is a purification of the final state of \(B'\) using the reference \(R_B\).

It is critical to comment on what happens if the symmetry of the violated is the symmetry, by defining the operator \(Z := (X_A + X_B) - U^\dagger (X_{A'} + X_B') U\) and its variance \(V_Z := V_{\rho_A \otimes \rho_B}(Z)\). Then, the dynamical fluctuation term in the relations [5] and [6] is replaced by a modified function which becomes small when the degree of violation is large (see Supplementary Material Supp.VII). For instance, the relation [5] is modified as the inequality \((A - V_Z)/[2(\sqrt{F} + 4\Delta_+ + 3V_Z)] \leq \delta\). When the violation of the symmetry is large, the numerator becomes negative, which implies that the inequalities reduce to trivial bounds. Hence, the meaningful limitations provided above exist due to the existence of symmetry.

B. Limitation on the information recovery without using \(R_B\)

Here we discuss the case without using the information of \(R_B\). The recovery operation \(R\) in this case maps the state on the system \(A'\) to \(A\). We then define the recovery error as

\[
\tilde{\delta} := \min_{R \in \mathcal{R}} D_{F}(\rho_{AR_A}, \text{id}_{R_A} \otimes R \circ \mathcal{E}(\rho_{AR_A})). \tag{7}
\]

Since \(\tilde{\delta} \geq \delta\), we can substitute \(\tilde{\delta}\) for \(\delta\) in [5] and [6] to get a limitation of recovery in the present setup. Moreover, in Supplementary Material Supp.IV, we can derive a tighter relation than this simple substitution as follows [61]:

\[
\frac{A}{2(\sqrt{F} + 4\Delta_+)} \leq \tilde{\delta}, \tag{8}
\]

where \(F_B := F_{\rho_B}(X_B)\). Note that \(F_B \leq F\) holds in general.

C. Mechanism of how conservation laws hinder quantum information recovery

Let us explain in an intuitive manner why conservation laws prevent quantum information recovery. We focus here on the case of the perfect recovery, i.e., \(\delta = 0\). In this case, after applying the recovery map \(R\), the state of the system \(AR_A\) is equal to the initial state \(|\psi_{AR_A}\rangle\).

(Fig. 3) Given that the state \(|\psi_{AR_A}\rangle\) is pure, the final state of the system \(B'\) is uncorrelated with \(AR_A\):

\[
\psi_{AR_A} \otimes \rho_{B'} = (R \otimes \text{id}_{B'R_B}) \circ (U \otimes \text{id}_{R_A R_B}) (|\psi_{AR_A}\rangle \otimes \phi_{BR_B}), \tag{9}
\]

where \(U(\ldots) := U\ldots U^\dagger\), \(\psi_{AR_A} := |\psi_{AR_A}\rangle\langle \psi_{AR_A}|\) and \(\phi_{BR_B} := |\phi_{BR_B}\rangle\langle \phi_{BR_B}|\). Therefore, no matter what measurement is made on the final state of system \(R_A\), no change will occur in the system \(B'\) according to the result of that measurement. Because \(R_A\) does not interact with any other system during the whole process, performing a measurement on the final state of \(R_A\) and performing the same measurement on the initial state of \(R_A\) will have exactly the same result. This is confirmed by performing a measurement \(\{M_i, R_A\}\) on \(R_A\) on both hand sides of [9]:

\[
\rho_{j,AR_A} \otimes \rho_{B'} = (R \otimes \text{id}_{B'R_B}) \circ (U \otimes \text{id}_{R_A R_B}) (\rho_{j,AR_A} \otimes \phi_{BR_B}), \tag{10}
\]

where \(\rho_{j,AR_A}\) is replaced by a modified function which becomes small when the degree of violation is large (see Supplementary Material Supp.VII). For instance, the relation [5] is modified as the inequality \((A - V_Z)/[2(\sqrt{F} + 4\Delta_+ + 3V_Z)] \leq \delta\). When the violation of the symmetry is large, the numerator becomes negative, which implies that the inequalities reduce to trivial bounds. Hence, the meaningful limitations provided above exist due to the existence of symmetry.
where \( \rho_{j,AR} := M_j R_A \psi_{AR} M_j^{\dagger} / q_j \) and \( q_j := \text{Tr}[M_j^{\dagger} M_j \psi_{AR}] \). Because the recovery map \( \mathcal{R} \) is a CPTP map from \( A' R_B \) to \( A \), we can remove it from (10) through the partial tracing of \( AR \) on both sides of the equation. We then obtain

\[
\rho_{B'} = \text{Tr}_{AR}[U (\rho_{j,A} \otimes \rho_B) U^{\dagger}].
\]

Here \( \rho_{j,A} := \text{Tr}_{R_A}[\rho_{j,AR}] \) and \( \rho_B := \text{Tr}_{R_B}[\phi_{BR_B}] \). Noting that \( \rho_{j,A} \) is the resultant state on \( A \) obtained when the measurement outcome of \( \{ M_j \} \) on \( R_A \) is \( j \), we can see that no matter what measurement is applied to the initial state of \( R_A \), the final state of \( B' \) will be independent of the results of the measurement.

However, when there is a conservation law and a local conserved quantity is changed by the local dynamics (i.e., when \( A > 0 \) is valid), the above never holds. This is because by performing a measurement on the initial state of \( R_A \), the state of \( A \) changes according to the results of the measurement, and the expectation value of the conserved quantity in \( B' \) also changes. A simple example to understand this argument is when the conserved quantity \( X \) is energy and the initial state \( |\psi_{AR_A}\rangle \) of \( AR_A \) is \( |00\rangle_{AR_A} + |11\rangle_{AR_A} / \sqrt{2} \) such that the strict inequalities \( \langle X_A \rangle_{00} - \langle X_{A'} \rangle_{\mathcal{E}(00)} > 0 \) and \( \langle X_A \rangle_{11} - \langle X_{A'} \rangle_{\mathcal{E}(11)} < 0 \) hold. This example indicates the change in expectation value of energy from \( A \) to \( A' \) is positive if the initial state of \( A \) is \( |0\rangle_A \) and negative if \( |1\rangle_A \). Because energy conservation holds, the change in energy from \( B \) to \( B' \) is negative (positive) if the state of \( A \) is \( |0\rangle_A \) (\( |1\rangle_A \)). In other words, the expectation value of \( X_{B'} \) is different depending on whether the state of \( A \) is \( |0\rangle_A \) or \( |1\rangle_A \).

IV. APPLICATION TO THE HAYDEN-PRESKILL MODEL WITH A CONSERVATION LAW

Our results are applicable to the black hole information recovery problems with a conservation law.

Here, we briefly review the Hayden-Preskill (HP) model \( \uparrow \) (Fig. 4). The HP model is a quantum mechanical model where Alice trashes her diary into a black hole \( B \), and Bob tries to recover the contents of the diary through Hawking radiation, assuming that the dynamics of the black hole is unitary. The diary \( A \) contains \( k \)-qubit quantum information, and is initially maximally entangled with another system \( R_A \). The black hole is as-

FIG. 3. Schematic of a fully successful quantum information recovery. In this case, the error \( \delta \) is equal to zero, and the final state of \( AR_A \) is equal to the initial state \( |\psi_{AR_A}\rangle \). Moreover, the final state of \( AR_A \) is completely uncorrelated with \( B' \), because pure states do not have any correlation with other systems. Therefore, when \( \delta = 0 \), any measurement on \( R_A \) cannot affect the state of \( B' \). However, when a conservation law exists, and when \( A > 0 \) holds, the unitary \( U \) correlates \( AR_A \) and \( B' \), and hence measurements on \( R_A \) affect the final state of \( B' \). Therefore, under the conservation law, the perfect information recovery from the scrambling processes satisfying \( A > 0 \) cannot be attained.

FIG. 4. Schematic diagram of the HP black hole model, which is almost a special case of our setup illustrated in Fig. 2.
assumed to contain $N$-qubit quantum information, where $N := S_{BH}$ is interpreted as the Bekenstein-Hawking entropy. After throwing the diary into the black hole, the HP model assumes a Haar random unitary operation that scrambles the quantum information [7, 9, 27]. Another assumption is that the black hole scrambles the quantum information [7, 9, 27]. This assumption is that the black hole dynamics $U$ satisfies the following inequality for arbitrary eigenstates $|i,a⟩$ and $|j,b⟩$ of $X_A$ and $X_B$, unless the sum of the eigenvalues of $|i,a⟩$ and $|j,b⟩$ is too close to the maximum or minimum eigenvalues of $X_A + X_B$:

$$V_{ρ_B|i,a,j,b,U}(X_{α′}) \leq \frac{1 + ε}{4} \min\{l, γ(N + k)\}$$ (13)

where $ε$ is a negligibly small number describing the error of the equidistribution on the expectation value and $ρ_B := Tr_{ARB}|ϕ_{ARB}⟩⟨ϕ_{ARB}|$ of the black hole state $B$ satisfies $V_{ρ_B}(X_B) ≤ D_{X_B}/4(=N/4)$. We remark that this assumption is satisfied when $X$ is energy, $B$ is a natural thermodynamic system, and $ρ_B$ is a non-zero temperature Gibbs state including the maximally mixed state.

Under the above conditions, we now use the result [9]. In particular, when $ρ_A$ commutes with $X_A$, we can evaluate $A, F_f$, and $Δ_{max}$ in (6) as follows (for details, see Supplementary Material Supp.III):

$$A ≥ γM(1 - ε),$$ (14)

$$\sqrt{F_f} ≤ 2(1 + ε)\sqrt{γ(N + k^2)},$$ (15)

$$Δ_{max} ≤ γk(1 + ε),$$ (16)

where $γ := (1 - l/(N + k))$, and $M := ⟨|X_A - ⟨X_A⟩|⟩_{ρ_A}$ is the mean deviation of $X_A$ in $ρ_A$. Due to (14)-(16), when $N + k > l$, we can convert (8) into the following form:

$$\frac{1 - ε}{1 + 2ε(2γk + 2√γ(N + k^2))} ≤ δ$$ (17)

To interpret this inequality, we set $k = \sqrt{N}$ and $M = k/2$ (we can take such an $M$ by considering a relevant $ρ_A$ and its decomposition, e.g., $ρ_A = (ρ_A^{max} + ρ_A^{min})/2$, where $ρ_A^{max}$ is the maximally mixed state of the eigenspace of $X_A$ corresponding to eigenvalue $x$). We then obtain the following lower bound of the recovery error:

$$\frac{\text{const.}}{1 + 2\sqrt{γ}} ≤ δ,$$ (18)

where const. is a real number larger than 0.24. This inequality rigorously restricts recovery of quantum information from the black hole. Since this inequality depends only on the ratio between the remaining part of

![Schematic diagram of the assumption of how the expectation value of the conserved quantity $X$ is distributed. In this diagram, we refer to the expectation values of $X$ in $A$ as $x_a$ ($α = A, B, A'$, and $B'$). We assume that the expectation value is given through the equidistribution. Precisely, we assume that after the unitary time evolution $U$, the expectation values of the conserved quantity $X$ are divided among $A'$ and $B'$ in proportion to the corresponding number of qubits.](image-url)
Recovery error \( \delta \) is negligibly small compared with the black hole energy conservation applies, even when 90 percent of the neous, whereas a non-negorable error remains when the conservation law, the recovery of information is instanta-

The restriction \( \text{(18)} \) is important in two respects. First, this result differs at a qualitative level from the original prediction \( \text{(12)} \), which is given in the absence of energy conservation (see Fig. 6). The graphs \( \text{(a)} \) in Fig. 6 plot the lower bound \( \text{(18)} \) of the recovery error when the energy conservation law holds, and the upper bound \( \text{(12)} \) of the error, which holds in the absence of any conservation law. The graphs deal with the case \( N=1024, k=32 \) and \( N=10^{84}, k=10^{42} \), respectively. The horizontal axis corresponds to the ratio \( l/(N+k) \) quantifying how much the black hole has evaporated after the diary \( A \) is thrown in. No evap-
oration has yet occurred when the horizontal axis is 0; when the ratio is 1, the black hole has evaporated entirely. As seen from the graphs, imposition of the energy conservation law changes the behavior of the recovery error considerably, even when \( N \gg k \). When no conservation laws apply, an almost-complete information recovery is possible when the amount of Hawking radiation emitted is equal to that of the object thrown in. In contrast, under the energy conservation law, there is an inevitable error that depends only on the fraction of the black hole evaporated. The error remains even when the evaporation of the black hole is quite advanced. Here, we use a tighter version of the bound \( \text{(6)} \) for which the right-hand side is twice that of the original \( \text{(6)} \). (see the footnote \[61\]).

The black hole \( B' \) and the total amount of qubits, i.e. \( \gamma = 1 - 1/(N + k) \), we see that the recovery error remains non-negorable, even after a considerable evaporation of the black hole.

The restriction \( \text{(18)} \) is important in two respects. First, this result differs at a qualitative level from the original prediction \( \text{(12)} \), which is given in the absence of energy conservation (see Fig. 6). The graphs \( \text{(a)} \) in Fig. 6 plot the lower bound \( \text{(18)} \) of the recovery error when the energy conservation law holds, and the upper bound \( \text{(12)} \) of the error, which holds in the absence of any conservation law. The graphs deal with the case \( k=32 \) and \( N=1024 \). As seen from the graphs, when there is no conservation law, the recovery of information is instantaneous, whereas a non-ignorable error remains when the energy conservation applies, even when 90 percent of the black hole evaporates.

Second, this result works even when “Alice’s diary” \( A \) is negligibly small compared with the black hole \( B \). Note that the qubit number \( N \) of the black hole is considered to correspond to the Bekenstein-Hawking entropy of the black hole, and thus it is a very large number. For this reason, \( k = \sqrt{N} \) is usually compatible with \( k \) being much smaller than \( N \). For example, the Bekenstein-Hawking entropy of Sagittarius A (the black hole at the center of the Milky Way) is approximately equal to \( 10^{85} \). In this case, \( \sqrt{N} = 10^{42.5} \), and thus \( \text{(18)} \) is valid when \( k/N = 10^{-42.5} \ll 1 \). Even in this scenario, the recovery error remains non-negorable until the 90 percent of the black hole evaporates (see the graph \( \text{(b)} \) in Fig. 6). Therefore, the energy conservation law provides a clear limit on the escape of information from a black hole, even if the object thrown into the black hole is much smaller than the black hole itself.

The above bound restricts general scenarios, including those for which the black hole is much larger than the diary. When the size of the black hole is comparable to the size of the diary, \( \text{(6)} \) provides another restriction. Since \( \sqrt{F_{\text{f}}} \) is always smaller than the square of the qubit number of \( B' \), we obtain

\[
\sqrt{F_{\text{f}}} \leq \gamma(N + k).
\]

Combining \( \text{(14)} \), \( \text{(16)} \), and \( \text{(19)} \), we obtain

\[
\frac{1 - \epsilon}{1 + \epsilon} \times \frac{M}{2(N + 2k)} \leq \delta.
\]

Similar to the derivation of \( \text{(18)} \), we set \( M = k/2 \) and
obtain the following lower bound of the recovery error:

$$\frac{\text{const.}}{1 + N/2k} \leq \delta,$$  \hspace{1cm} (21)

where the constant on the left-hand side is larger than 1/9. Unlike the inequality \[18\], the inequality \[21\] becomes trivial when \(N \gg k\). However, when \(N\) is comparable to \(k\), the inequality \[21\] gives a non-negligible lower bound for the error \(\delta\) that is independent of \(l\). The lower bound is valid whenever \(l < N + k\) holds. Therefore, when the ratio \(N/k\) is not so large, the recovery error cannot be small until \(l = N + k\) holds. In other words, the recovery of the quantum information associated with Alice’s diary does not finish until the black hole completely evaporates.

Our results \[18\] and \[21\] show that the behavior of the quantum information recovery under the energy conservation law is qualitatively different from the original analysis in the HP model without energy conservation \[12\], as evident in Fig. 7.

V. APPLICATIONS TO QUANTUM INFORMATION PROCESSING WITH SYMMETRY

Our formulae \[5\] and \[6\] are applicable to various phenomena other than scrambling. Below, we apply our bounds to quantum error correction (QEC) and implementation of unitary gates as examples of application.

A. Application 1: quantum error correcting codes with symmetry

In QEC, we encode quantum information in a logical system \(A\) into a physical system \(A'\) which is a composite system of \(N\) subsystems \(\{A'_i\}^N_{i=1}\) by an encoding channel \(C\), which is a CPTP map. After the encoding, noise occurs on the physical system \(A'\), which is described by a CPTP-map \(N\). Finally, we recover the initial state by performing a recovery CPTP map \(R\) from \(A'\) to \(A\). Then, the recovery error is defined as

$$\delta_C := \min_{(R^{A'\to A})} \max_{\rho_{\text{id}}} D_F(\rho_{AR_A}, R \circ N \circ C(\rho_{AR_A})).$$  \hspace{1cm} (22)

Here we focus on the case where the channel \(C\) transversal with respect to a unitary representation \(\{U_{A,t}\} \in \mathbb{R}\), i.e.

$$C \circ U_{t}^{A'}(...) = U_{t}^{A'} \circ C(...), \hspace{1cm} \forall t \in \mathbb{R},$$  \hspace{1cm} (23)

where \(U_{t}^{A'}(...) = e^{i X_{A'} t}(...) e^{-i X_{A'} t} \ (\alpha = A, A')\) and \(X_{A'}\) is described as \(X_{A'} := \sum_j X_{A'_j}\) with operators \(\{X_{A'_j}\}^N_{j=1}\) on \(A'_j\).

The limitations of the transversal codes is a critical issue \[39\]-\[43\]. It is shown that the code \(C\) cannot make \(\delta_C = 0\) for local noise by the Eastin-Knill theorem \[39\]. Recently, the Eastin-Knill theorem were extended to the cases where \(\delta_C\) is finite \[40-43\]. These approximate Eastin-Knill theorems show that the size \(N\) of the physical system must be inversely proportional to \(\delta_C\).

From \[6\], we can derive a variant of the approximate Eastin-Knill theorem as a corollary (see Supplementary Material Supp.V):

$$\frac{D_{X_A}}{4D_{\text{max}}(N + D_{X_A}/(4D_{\text{max}}))} \leq \delta_C.$$  \hspace{1cm} (24)

Here \(D_{\text{max}} := \max_{\rho} D_{X_A}\). Our bounds \[7\] and \[8\] are also applicable to cases where \(N\) is non-local, and more general covariant codes with general Lie group symmetries (see Supplementary Materials Supp.VI).

B. Application 2: Implementation of unitary dynamics

The last application is on the implementation of the unitary dynamics on the subsystem \(A\) through the unitary time-evolution of the isolated total system \[34\]-\[35\]. This subject has a long history in the context of the limitation on the quantum computation imposed by conservation laws \[34\]-\[35\]-\[37\]-\[37\], which is considered as an extension of the Wigner-Araki-Yanase theorem \[52\]-\[54\] to quantum computation. Suppose that we try to approximately realize a desired unitary dynamics \(U_A\) on a system \(A\) as a result of the interaction between another system \(B\). We assume that the interaction satisfies a conservation law: \([U, X_A + X_B] = 0\). We then define the implementation error \(\delta_U\) as:

$$\delta_U := \max_{\rho_{\text{id}}} D_F(\rho_{AR_A}, \text{id}_{AR_A} \otimes U_A \circ \mathcal{E}(\rho_{AR_A})).$$  \hspace{1cm} (25)

Here \(U_A(\ldots) := U_A^{t}(\ldots) U_A\). The quantum operation \(\mathcal{E}\) is the CPTP-map where \(A' = A\). Then, by definition, the inequality \(\delta_U \geq \max_{\rho_{\text{id}}} \delta \geq \max_{\rho_{AR_A}} \delta\) holds. Therefore, we can directly apply \[5\] and \[6\] to this problem. In particular, we obtain the following inequality from \[8\]:

$$\frac{A}{2(\sqrt{2} + 4\Delta)} \leq \delta_U.$$  \hspace{1cm} (26)

This inequality gives a trade-off between the implementation error and the coherence cost of implementation of unitary gates. The physical message is that the implementation of the desired unitary operator requires the quantum coherence inversely proportional to the square of the implementation error. We remark that several similar bounds for the coherence cost were already given in Refs. \[34\]-\[35\]. However, we stress that \[26\] is given as a corollary of a more general relation \[5\]. Moreover, as we pointed out several times, our results can be extended to the cases of general Lie group symmetries. In Supplementary Materials Supp.VI, we show a generalized version of \[26\] for such cases.
VI. NUMERICAL CHECK OF THE MAIN INEQUALITY

So far, we have applied our main result to information scrambling and quantum information processing. Our bound works regardless of the size of systems $A$ and $B$ and is especially tight when system $B$ is large. For example, as shown in the previous section, the bounds obtained from become optimal when $F$ is very large.

We next give a numerical check of situations in which the system $B$ is small to show that the bound also works rigorously. For this purpose, we prepare a concrete model. Let us consider four qubits, $A$, $R_A$, $B_1$, and $R_{B_2}$. We also take a natural number $b$ and a $b + 2$-dimensional system $B_2$. For $A$, $B_1$, and $B_2$, we define Hermitian operators $X_A$, $X_{B_1}$, and $X_{B_2}$ as follows:

\[
X_A := |1\rangle\langle 1|_A, \quad X_{B_1} := |1\rangle\langle 1|_{B_1}, \quad X_{B_2} := \sum_{x=1}^{b+2} 2x|x\rangle\langle x|_{B_2},
\]

where $\{|x\rangle_A\}_{x=0,1}$, $\{|x\rangle_{B_1}\}_{x=0,1}$, $\{|x\rangle_{B_2}\}_{x=0,...,d+2}$ are orthogonal basis on $A$, $B_1$, and $B_2$, respectively.

For the above system, we prepare the following initial states:

\[
|\psi_{AR_A}\rangle := \frac{|00\rangle_{AR_A} + |11\rangle_{AR_A}}{\sqrt{2}}, \quad |\phi_{B_1R_{B_2}}\rangle := \frac{|00\rangle_{B_1R_{B_2}} + |11\rangle_{B_1R_{B_2}}}{\sqrt{2}},
\]

\[
|\phi_{B_2}\rangle := \frac{1}{\sqrt{b}} \sum_{x=1}^{b} x|\rangle_{B_2}.
\]

We prepare the following unitary $U_{AB_1B_2}$ on $AB_1B_2$:

\[
U_{AB_1B_2} := \sum_{1 \leq k \leq b+1} (|11\rangle\langle 00|_{AB_1} \otimes |k - 1\rangle\langle k| 
+ |10\rangle\langle 01|_{AB_1} \otimes |k\rangle\langle k| 
+ |00\rangle\langle 00|_{AB_1} \otimes |k\rangle\langle k| 
+ |01\rangle\langle 01|_{AB_1} \otimes |k - 1\rangle\langle k| 
+ |10\rangle\langle 10|_{AB_1} \otimes |0\rangle\langle 0| 
+ |00\rangle\langle 00|_{AB_1} \otimes |0\rangle\langle 0| 
+ |10\rangle\langle 10|_{AB_1} \otimes |0\rangle\langle 0| 
+ |11\rangle\langle 11|_{AB_1} \otimes |b+1\rangle\langle b+1|).
\]

After a unitary operation $U$, we perform a CPTP map on $AR_{B_2}$ to recover the initial state $|\psi_{AR_A}\rangle$ on $AR_A$. Following our framework, the minimum recovery error is defined as follows:

\[
\delta = \min_R D_F(\psi_{AR_A}, \id_{R_A} \otimes R \circ \mathcal{E}_{A \rightarrow AR_{B_1}}(\psi_{AR_A})).
\]

Here $\mathcal{E}_{A \rightarrow AR_{B_1}}(\ldots) := \Tr_{B_1B_2}[U_{AB_1B_2}(\ldots \otimes \phi_{B_1R_{B_2}} \otimes \phi_{B_2})U_{AB_1B_2}^\dagger]$. Since $U_{AB_1B_2}$ conserves $X_A + X_{B_1} + X_{B_2}$ and $\rho_A := \Tr_{R_{B_2}}[\psi_{AR_A}] = \frac{|00\rangle\langle 00| + |11\rangle\langle 11|}{2}$ our bound is applicable to the above model. In this case, each term on the right-hand side of is evaluated as follows:

\[
A \geq 1, \quad \Delta_{\max} \leq 1, \quad F_f = \frac{b^2 + 14}{3}.
\]

Thus, for each $b = 2, ..., 7$, the inequality predicts that there is no recovery that can make the recovery error smaller than the black dot in Fig. In Fig. we confirm the prediction by numerical calculation. We randomly generate a CPTP-map from $AR_{B_2}$ to $A$ as a recovery map $10^8$ times for each $b$, and plot the smallest recovery error obtained by the random recoveries (blue point in Fig. ). Fig. clearly shows that all of the random recoveries satisfy the bound.

Furthermore, to ensure that a sufficiently large number of trials is obtained, we give the following specific upper bound of the minimum recovery error by extending the Åberg protocol to the model:

\[
\delta \leq \sqrt{\frac{2}{7}}.
\]

Combined with (37), (38) gives the following upper bound for each $b$ (purple points in Fig.):

\[
\delta \leq \sqrt{\frac{2}{3F_f - 14}}.
\]

Clearly, the best of the randomly generated recoveries performs better than this upper bound. Therefore, our number of trials is considered sufficient. In addition, the lower bound and the upper bound are off by a factor of at most 2 or 3 (Fig. ). This suggests that even for a small system, the evaluation of the recovery error by the bound is acceptable.
\begin{center}
\includegraphics[width=\textwidth]{fig9}
\end{center}

FIG. 9. Comparisons between lower bound, analytical upper bound, and numerical upper bound from plots of recovery error versus $F_f$. The data points correspond to $b = 3, 4, 5, 6, 7$. Here, we use a tighter version of the bound (6) for which the right-hand side is twice that of the original (see the footnote [61]).

VII. SUMMARY

In summary, we have clarified fundamental limitations to information recovery from dynamics with general Lie group symmetry. As demonstrated in Appendix [C], all results in this paper are given as corollaries of (6). It is remarkable that one single inequality (6) provides a unifying limit for black holes, quantum error correcting codes and unitary gates. In particular, the HP model with the energy conservation, some of the information thrown into the black hole cannot escape to the end. We also remark that our prediction may be validated in laboratory experiments that mimic the HP model with symmetry [15–19]. Moreover, an intriguing topic to consider is the relationship between our relations and a recent argument on the weak violation of global symmetries in quantum gravity [64–66]. The effect of symmetry on the OTOC decay is another interesting future direction from our results.
Coherence cost

G. Gour, R. W. Spekkens, New perspectives on covari-
systems $E$ and $E'$ satisfying $SE = S'E'$, Hermite operators $X_E$ and $X_{E'}$ on $E$ and $E'$, a unitary operation $U$ on $SE$ satisfying $U(X_S + X_E)U^\dagger = X_{S'} + X_{E'}$, and a symmetric state $\mu_E$ on $E$ satisfying $[\mu_E, X_E] = 0$ such that
\begin{equation}
C(\ldots) = \text{Tr}_{E'}[U(\ldots \otimes \mu_E)U^\dagger].
\end{equation}

The SLD-quantum Fisher information for the family $\{e^{-iXt}\rho e^{iXt}\}_{t \in \mathbb{R}}$, described as $F_{\rho_S}(X_S)$, is a standard resource measure in the resource theory of asymmetry\textsuperscript{37, 38}. It is also known as a standard measure of quantum fluctuation, since it is related to the variance $V_{\rho_S}(X_S) := \langle X_S^2 \rangle_{\rho_S} - \langle X_S \rangle_{\rho_S}^2$ as follows\textsuperscript{35, 59, 60}:
\begin{equation}
F_{\rho_S}(X_S) = 4 \min_{\{q_i, \phi_i\}} \sum_i q_i V_{\phi_i}(X_S) \quad (A3)
= 4 \min_{|\Psi_{SR}\rangle, X_R} V_{\Psi_{SR}}(X_S + X_R) \quad (A4)
\end{equation}
where $\{q_i, \phi_i\}$ runs over the ensembles satisfying $\rho = \sum_i q_i \phi_i$ and each $\phi_i$ is pure, and $\{|\Psi_{SR}\rangle, X_R\}$ runs over purifications of $\rho_S$ and Hermitian operators on $R$. The equality of (A3) shows that $F_{\rho}(X)$ is the minimum average of the fluctuation arising from quantum superposition. Note that it also means that if $\rho$ is pure, $F_{\rho}(X) = 4V_{\rho}(X)$ holds. The $|\Psi_{SR}\rangle$ and $X_R$ achieving the minimum of $V_{\Psi_{SR}}(X_S + X_R)$ in (A4) are $|\Psi_{SR}\rangle := \sum_i \sqrt{q_i}|l_S\rangle|l_R\rangle$ and
\begin{equation}
X_R := \sum_i 2\frac{\sqrt{r_i r_j}}{r_i + r_j} \langle l_S|X_S|l_S\rangle|l_R\rangle\langle l_R|,
\end{equation}
where $\{r_i\}$ and $\{|l_S\rangle\}$ denote the eigenvalues and eigenvectors of $\rho_S$\textsuperscript{35}.

\section*{Appendix B: Note on entanglement fidelity and average gate fidelity}

In this subsection, we show that the recovery error $\delta$ can approximate the average of the recovery error which is averaged thorough pure states on the entire Hilbert space of $A$ or on its subspace using special initial states as $|\psi_{AR}\rangle$\textsuperscript{63}.

For explanation, let us introduce the average fidelity and the entanglement fidelity. For a CPTP map $C$ from a quantum state $Q$ to $Q$, these two quantities are defined as follows:
\begin{equation}
F_{\text{avg}}^{(2)}(C) := \int d\psi_Q F(|\psi_Q\rangle, C(|\psi_Q\rangle)^2, \quad (B1)
\end{equation}
\begin{equation}
F_{\text{ent}}^{(2)}(C) := F(|\psi_{QR}\rangle, 1_{RQ} \otimes C(|\psi_{QR}\rangle)^2, \quad (B2)
\end{equation}
where $|\psi_{QR}\rangle$ is a maximally entangled state between $Q$ and $R_Q$, and the integral is taken with the uniform (Haar) measure on the state space of $Q$. For these two quantities, the following relation is known\textsuperscript{63}:
\begin{equation}
F_{\text{avg}}^{(2)}(C) = \frac{d_Q F_{\text{ent}}^{(2)}(C) + 1}{d_Q + 1}. \quad (B3)
\end{equation}
Then, due to (B3), when we set \( \delta \) define the following average recovery error:

\[
\delta_{\text{avg},S}^{(2)} := \min_{\mathcal{R}} \int_{S} d\psi_A D_F(\psi_A), R(\text{Tr}_B U(\psi_A \otimes \phi_{B\rho_B})U^\dagger))^2. \tag{B4}
\]

Then, due to (B3), when we set \( |\psi_{AR_A,s}\rangle = \sum_i |i\rangle_A |i\rangle_R \) where \( \{|i\rangle_A\} \) is an arbitrary orthonormal basis of \( S \) and \( d_S \) is the dimension of \( S \), the recovery error \( \delta_S := \delta(|\psi_{AR_A,s}\rangle, |\phi_{BR_B}\rangle, U) \) satisfies the following relation:

\[
\delta_{\text{avg},S}^{(2)} = \frac{d_S}{d_S + 1} \delta_S^2. \tag{B5}
\]

Therefore, when we use a maximally entangled state between a subspace of \( A \) and \( R_B \) as \( |\psi_{AR_A}\rangle \), the recovery error \( \delta \) for the \( |\psi_{AR_A}\rangle \) approximates the average of recovery error which is averaged through all pure states of the subspace of \( A \).

**Appendix C: Relations between main results and applications in this paper**

Next, we show the relation between the main results and applications in this paper (Fig. 10). We derive (6) from two lemmas which we give in the next two subsections. All of the physical results in this paper including (5) and (21) are given as corollaries of (6). In that sense, (6) is a universal restriction on information recovery from dynamics with Lie group symmetry. In addition to what is described in the main text, various results can be given in a similar way. For instance, we can derive the Wigner–Araki–Yanase theorem for unitary gates from (6). We also derive another restriction on HP model with symmetry from (5).

We remark that there exist several variations and generalizations of the results in Fig. 10. For instance, in Appendix G, we derive tighter variations of (5) and (6). We also extend (5) and (6) to general Lie group symmetries in Supplementary Material Supp.VI.

**Appendix D: Trade-off relation between irreversibility and back-reaction**

In the derivation of (5) and (6), we use the following lemma:

**Lemma 1** In the setup of Section 2, let us consider an arbitrary decomposition of the initial state of \( A \) as \( \rho_A = \sum_j p_j \rho_j \). We also refer to the final states of \( B' \) for the cases where the initial states of \( A \) are \( \rho_j \) and \( \rho_A \) as \( \rho_{j,B'} \) and \( \rho_{B'} \), respectively. Namely, \( \rho_{j,B'} := \text{Tr}_A[|U(\rho_j \otimes \rho_B)U^\dagger|] \) and \( \rho_{B'} := \text{Tr}_A[|U(\rho_A \otimes \rho_B)U^\dagger|] \) where \( \rho_B := \text{Tr}_{R_B}[\rho_{BR_B}] \). Then, there exists a state \( \sigma_{B'} \) such that

\[
\sum_j p_j D_F(\rho_{j,B'}, \sigma_{B'})^2 \leq \delta^2. \tag{D1}
\]

Moreover, the following inequality holds:

\[
\sum_j p_j D_F(\rho_{j,B'}, \rho_{B'})^2 \leq 4\delta^2. \tag{D2}
\]

Lemma (1) holds even when \( U(X_A + X_B)U^\dagger \neq X_A + X_B \). The proof of this lemma is given in Appendix H. Roughly speaking, this lemma means that when the recovery error \( \delta \) is small (i.e. the realized CPTP map \( E \) is approximately reversible), then the final state of \( B' \) becomes almost independent of the initial state of \( A \).

This lemma is a generalized version of (16) in Ref. 34 and Lemma 3 in Ref. 35. The original lemmas are given for the implementation error of unitary gates, and used for lower bounds of resource costs to implement desired unitary gates in the resource theory of asymmetry 34, 35 and in the general resource theory 57.

**Appendix E: mean-variance-distance trade-off relation**

For an arbitrary Hermite operator \( X \) and arbitrary states \( \rho \) and \( \sigma \), there is a trade-off relation between the difference of expectation values \( \Delta := \langle X \rangle_\rho - \langle X \rangle_\sigma \), the variances \( V_\rho(X) \) and \( V_\sigma(X) \), and the distance between \( \rho \) and \( \sigma \) [68]:

\[
|\Delta| \leq D_F(\rho, \sigma)(\sqrt{V_\rho(X)} + \sqrt{V_\sigma(X)} + |\Delta|), \tag{E1}
\]

This is an improved version of the original inequality (15) in Ref. 32. In the original inequality, the purified distance \( D_F(\rho, \sigma) \) is replaced by the Bures distance \( L(\rho, \sigma) := \sqrt{2(1 - F(\rho, \sigma))} \). These inequalities mean that if two states have different expectation values and are close to each other, then at least one of the two states exhibits large fluctuation.

**Appendix F: Properties of variance and expectation value of the conserved quantity \( X \)**

We use several properties of variance and expectation value of the conserved quantity \( X \). In our setup described in Section 11, we have assumed that the unitary dynamics \( U \) satisfies the conservation law of \( X \): \( U(X_A + X_B)U^\dagger = X_A + X_B \). Under this assumption, for arbitrary states

\[
\rho_B := \text{Tr}_{R_B}[\rho_{BR_B}].
\]
Lemma 1 (Eq. (D2)):
\[ \sum_j p_j D_F(\rho_{j,B}^f, \rho_{j,B}^{f'})^2 \leq 4B^2 \]

MVD trade-off (Eq. (E1)):
\[ |\Delta| \leq D_F(\rho, \sigma) \left\{ \sqrt{\rho_j(X)} + \sqrt{\sigma_j(X)} + |\Delta| \right\} \]

Main 1 (Eq. (6)):
\[ \frac{A}{2(\sqrt{F_j} + \Delta_{\text{max}})} \leq \delta \]

Application to BH 1 (Eq. (18)):
\[ \frac{\text{const.}}{1 + 2\sqrt{\delta}} \leq \delta \]

Main 2 (Eq. (5)):
\[ \frac{A}{2(\sqrt{F_j} + 4\Delta_+)} \leq \delta \]

Application to BH 1 (Eq. (21)):
\[ \frac{\text{const.}}{1 + N/2k} \leq \delta \]

Limitation for recovery without \( R_B \) (Eq. (8)):
\[ \frac{A}{2(\sqrt{F_j} + 4\Delta_+)} \leq \delta \]

Wigner-Araki-Yanase type bound for unitary gates (Eq. (26))
\[ \frac{A}{2(\sqrt{F_j} + 4\Delta_+)} \leq \delta_U \]

FIG. 10. Schematic depicting the relationship between the main results and applications.

\[ \xi_A \text{ and } \xi_B \text{ on } A \text{ and } B, \text{ the following relations hold:} \]
\[ \langle X_A \rangle_{\xi_A} - \langle X_A' \rangle_{\xi_A'} - \langle X_B \rangle_{\xi_B} \]
\[ \sqrt{V_{\xi_B}(X_B)} \leq \sqrt{V_{\xi_A}(X_A')} + \sqrt{V_{\xi_B}(X_B)} \]
\[ \leq \sqrt{V_{\xi_B}(X_B)} + \Delta_+, \quad (F2) \]
\[ \sqrt{V_{\xi_B}(X_B)} \leq \sqrt{V_{\xi_A}(X_A')} + \sqrt{V_{\xi_B}(X_B)} \]
\[ \leq \sqrt{V_{\xi_B}(X_B)} + \Delta_+ , \quad (F3) \]

where \( \xi_A' := \mathcal{E}(\xi_A) = \text{Tr}_B[U(\xi_A \otimes \xi_B)U^\dagger] \) and \( \xi_B' := \text{Tr}_A[U(\xi_A \otimes \xi_B)U^\dagger] \). We show these two relations in Appendix G.

Appendix G: Derivation of the limitations of information recovery error (case of single conserved quantity)

Now, we derive (6) as follows:
\[ A \leq \sum_j p_j |\langle X_B \rangle_{\rho_{j,B}^f} - \langle X_B \rangle_{\rho_{j,B}^{f'}}| \]
\[ \leq \sum_j p_j D_F(\rho_{j,B}^f, \rho_{j,B}^{f'}) \]
\[ \times \left( \sqrt{V_{\rho_{j,B}^f}(X_B)} + \sqrt{V_{\rho_{j,B}^{f'}}(X_B)} + |\Delta_j| \right) \]
\[ \leq \sum_j p_j D_F(\rho_{j,B}^f, \rho_{j,B}^{f'})^2 \sum_j p_j V_{\rho_{j,B}^{f'}}(X_B) \]
\[ + 2\delta \left( \sqrt{V_{\rho_{j,B}^{f'}}(X_B)} + \Delta_{\text{max}} \right) \]
\[ \leq 2\delta \left( 2\sqrt{V_{\rho_{j,B}^{f'}}(X_B)} + \Delta_{\text{max}} \right) \]
\[ \leq 2\delta \left( \sqrt{F_j} + \Delta_{\text{max}} \right). \quad (G2) \]

Here we use (G1) in (a), (E1) in (b), the Cauchy-Schwartz inequality, Lemma 1 and \( |\Delta_j| \leq \Delta_{\text{max}} \) in (c), Lemma 1, and the concavity of the variance in (d), and \( F_j = 4V_{\rho_{j,B}^{f'}}(X_B) \) in (e).

We also derive (5) from (6):
\[ A \leq 2\delta \left( \sqrt{F_j} + \Delta_{\text{max}} \right) \]
\[ \leq 2\delta \left( 2\sqrt{V_{\rho_{j,B}^{f'}}(X_B)} + \Delta_{\text{max}} \right) \]
\[ \leq 2\delta \left( 2\sqrt{V_{\rho_{j,B}^{f'}}(X_B)} + 4\Delta_+ \right) \]
\[ \leq 2\delta \left( \sqrt{F_j} + 4\Delta_+ \right). \quad (G3) \]

Here we use \( F_j = 4V_{\rho_{j,B}^{f'}}(X_B) \) in (a), (F2) in (b), and \( F = 4\rho_B(\rho_B) = 4V_B(X_B) \) in (e).

Combining the above three methods, we can derive our main results [5] and [6]. We first decompose \( \rho_A = \sum_j p_j \rho_j \) such that \( A = \sum_j p_j |\Delta_j| \). Then, due to (F1), we obtain
\[ |\Delta_j| = |\langle X_B \rangle_{\rho_{j,B}^f} - \langle X_B \rangle_{\rho_{j,B}^{f'}}|. \quad (G1) \]
1. The case where $\rho_A = (\rho_0 + \rho_1)/2$ is possible

Let us consider the case that $|\psi_{AR_A}\rangle$ satisfies $\rho_A = (\rho_0 + \rho_1)/2$ with some proper density matrices $\rho_0$ and $\rho_1$. In this case, we can define a variation of $A$ as follows:

$$A_2 := \max_{\rho_0,\rho_1} \frac{1}{2} \sum_{j=0}^{1} |\Delta_j|.$$  

(G4)

where $\{\rho_0,\rho_1\}$ runs over $\rho_A = \frac{\rho_0 + \rho_1}{2}$. For $A_2$, we can obtain the following relations:

$$\frac{A_2}{\sqrt{F} + 4\Delta_+} \leq \delta,$$  

(G5)

and

$$\frac{A_2}{\sqrt{F_f} + \Delta_{\max}} \leq \delta.$$  

(G6)

Let us prove (G5) and (G6). We can derive (G5) from (G6) in the same manner as the derivation of (5) from (6). Therefore, we only have to prove (G5). From (D1), we obtain

$$\sum_{j=0}^{1} \frac{1}{2} D_F(\rho_{0,B'}^f, \sigma_{B'})^2 \leq \delta^2.$$  

(G7)

Therefore, using $(x+y)^2/4 \leq (x^2+y^2)/2$ and the triangle inequality for $D_F$, we obtain

$$D_F(\rho_{0,B'}^f, \rho_{1,B'}^f) \leq 2\delta.$$  

(G8)

Let us take a decomposition $\rho_A = \sum_{j=0}^{1} \frac{1}{2} |\Delta_j|$. Then, due to (F1), we obtain the following relation for both $j = 0$ and $j = 1$:

$$|\Delta_j| = |\langle X_{B'} \rangle_{\rho_{0,B'}^f} - \langle X_{B'} \rangle_{\rho_{1,B'}^f}|$$

$$= |\langle X_{B'} \rangle_{\rho_{0,B'}^f} - \langle X_{B'} \rangle_{\rho_{1,B'}^f} + \rho_{1,B'}^f|$$

$$= \frac{|\langle X_{B'} \rangle_{\rho_{0,B'}^f} - \langle X_{B'} \rangle_{\rho_{1,B'}^f}|}{2}$$  

(G9)

Then, we derive (G5) as follows:

$$A_2 \leq \sum_{j=0}^{1} \frac{1}{4} |\langle X_{B'} \rangle_{\rho_{0,B'}^f} - \langle X_{B'} \rangle_{\rho_{1,B'}^f}|$$

$$\leq \frac{1}{2} (\langle X_{B'} \rangle_{\rho_{0,B'}^f} - \langle X_{B'} \rangle_{\rho_{1,B'}^f})$$

$$\leq \frac{1}{2} \sqrt{F_f} (\rho_{0,B'}^f, \rho_{1,B'}^f)$$

$$\times (\sqrt{V_{\rho_{0,B'}^f}(X_{B'})} + \sqrt{V_{\rho_{1,B'}^f}(X_{B'})} + \Delta_{0,1})$$

$$\leq \delta (\sqrt{V_{\rho_{0,B'}^f}(X_{B'})} + \sqrt{V_{\rho_{1,B'}^f}(X_{B'})} + \Delta_{\max})$$

$$\leq \delta (2\sqrt{V_{\rho_{0,B'}^f}(X_{B'})} + \Delta_{\max})$$

$$\leq \delta (2\sqrt{V_{\rho_{0,B'}^f}(X_{B'})} + \Delta_{\max})$$

(G10)

Here, we use (G9) in (a), (E1) in (b), (G8) and $\Delta_{0,1} := |\langle X_{B'} \rangle_{\rho_{0,B'}^f} - \langle X_{B'} \rangle_{\rho_{1,B'}^f}| \leq \Delta_{\max}$ in (e), $\sqrt{x} + \sqrt{y} \leq 2\sqrt{(x+y)/2}$ in (d), the concavity of the variance in (e), and $F_f = 4V_{\rho_{0,B'}^f}(X_{B'})$ in (f).

Appendix H: Derivation of Lemma 1

In this section, we prove Lemma 1.

Proof of Lemma 1. We denote by $R_*$ the best recovery operation which achieves $\delta$ and take its Stinespring representation $(V, |\eta_C\rangle)$ (Fig. 11). Here, $V$ is a unitary operation on $A' R_B C$, and $|\eta_C\rangle$ is a pure state on $C$. Since $R_*$ is a CPTP-map from $A' R_B$ to $A$, we can take another system $C'$ satisfying $A' R_B C = AC'$. We refer to the initial and final state of the total system as $|\psi_{tot}\rangle$ and $|\psi_{tot}'\rangle$.

Then, these two states are expressed as follows:

$$|\psi_{tot}\rangle := |\psi_{AR_A}\rangle \otimes |\varphi_{RB}\rangle \otimes |\eta_C\rangle,$$  

(H1)

$$|\psi_{tot}'\rangle := (1_{R_A} \otimes V \otimes 1_{B'}) (1_{R_A} \otimes U \otimes 1_{R_B C}) |\psi_{tot}\rangle.$$  

(H2)

Due to the definitions of $\delta$ and $R_*$, for $\psi_{AR_A} := \text{Tr}_{B'C'}[\psi_{tot'}]$, we obtain

$$D_F(\psi_{AR_A}^f, |\psi_{AR_A}\rangle) = \delta.$$  

(H3)

Therefore, due to the Uhlmann theorem and the fact that $|\psi_{AR_A}\rangle$ is pure, there exists a pure state $|\phi_{B'C'}\rangle$ such that

$$D_F(\psi_{tot'}^f, |\psi_{AR_A}\rangle \otimes |\phi_{B'C'}\rangle) = \delta.$$  

(H4)
Due to (H8), (H10) and the monotonicity of $D_F$, we obtain
\[ D_F(\psi_{B'C'}^f, \phi_{B'C'}^f) \leq \delta. \] (H5)
where $\psi_{B'C'}^f := \text{Tr}_{AR} \rho_{(1)}$. Let us define $\sigma_{B'}$ as $\sigma_{B'} := \text{Tr}_{C'}[\phi_{B'C'}^f]$. Then, due to $\text{Tr}_{C'}[\psi_{B'C'}^f] = \rho_{B'}^f$ and (H5),
\[ D_F(\rho_{B'}^f, \sigma_{B'}) \leq \delta. \] (H6)
Here, we assume that there are states $\{\psi_{j,B'C'}^f\}$ on $B'C'$ such that
\[ \psi_{j,B'C'}^f = \sum_j p_j \tilde{\psi}_{j,B'C'}^f, \] (H7)
\[ \text{Tr}_{C'}[\tilde{\psi}_{j,B'C'}^f] = \rho_{j,B'}^f. \] (H8)
Below, we firstly prove (D1) and (D2) under the assumption of the existence of $\{\psi_{j,B'C'}^f\}$. We demonstrate the existence of $\{\psi_{j,B'C'}^f\}$ at the end of the proof.

Combining (H5) and (H7), we obtain
\[ D_F(\sum_j p_j \tilde{\psi}_{j,B'C'}^f, \phi_{B'C'}^f) \leq \delta. \] (H9)

From $D_F(\rho, \sigma) = \sqrt{1 - F(\rho, \sigma)^2}$ and $F(\rho, |\phi\rangle^2 = \langle \phi | \rho | \phi \rangle$, we obtain
\[ 1 - \delta^2 \leq \sum_j p_j \langle \phi_{B'C'}^f | \tilde{\psi}_{j,B'C'}^f | \phi_{B'C'}^f \rangle^2. \]
\[ = 1 - \sum_j p_j D_F(\tilde{\psi}_{j,B'C'}^f, \phi_{B'C'}^f)^2. \] (H10)

Due to (H8), (H10) and the monotonicity of $D_F$, we obtain the (D1):
\[ \sum_j p_j D_F(\rho_{j,B'}, \sigma_{B'}) \leq \delta. \] (H11)
Since the root mean square is greater than the average, we also obtain
\[ \sum_j p_j D_F(\rho_{j,B'}, \sigma_{B'}) \leq \delta. \] (H12)

Since the purified distance $D_F$ is not increased by the partial trace, we obtain
\[ D_F(\psi_{B'C'}^f, \phi_{B'C'}^f) \leq \delta. \] (H5)

Here we use (H6) in (a) and (H11) and (H12) in (b).

Finally, we show the existence of $\{\psi_{j,B'C'}^f\}$ satisfying (H7) and (H8). We firstly take a partial isometry $W_{RA}$ from $R_A$ to $R'_{A1}R'_{A2}$ such that
\[ 1_A \otimes W_{RA} \rho_{AR} = \sum_j \sqrt{p_j} |\psi_{j,AR_{A1}'}, \rangle \otimes |j_{R_{A2}'}\rangle, \] (H14)
\[ 1_A \otimes W_{RA}^\dagger W_{RA} \rho_{AR} = \rho_{AR}. \] (H15)
Here $\{ |j_{R_{A2}'}\rangle\}$ are orthonormal and $|\psi_{j,AR_{A1}'}\rangle$ is a purification of $\rho_{j}$. We abbreviate $R'_{A1}R'_{A2}$ as $R'_{A}$. The existence of $W_{RA}$ is guaranteed as follows. We firstly note that there exists a “minimal” purification $|\psi_{AR_{A1}'}\rangle$ of $\rho_{A}$, for which we can take isometries $W(1)$ from $R_A$ to $R_A$ and $W(2)$ from $R_A$ to $R_A$ such that $|70|
\[ (1_A \otimes W(1)) |\psi_{AR_{A1}'}\rangle = |\psi_{AR_{A1}'}\rangle, \] (H16)
\[ (1_A \otimes W(2)) |\psi_{AR_{A1}'}\rangle = \sum_j \sqrt{p_j} |\psi_{j,AR_{A1}'}\rangle \otimes |j_{R_{A2}'}\rangle. \] (H17)

The desired $W_{RA}$ is defined as $W_{RA} := W(2)W(1)^\dagger$. Since $W(2)$ and $W(1)$ are isometry, $W_{RA}$ is a partial isometry. Using $W(2)W(2) = W(1)W(1) = 1_{R_A}$, we obtain (H15) as follows:
\[ 1_A \otimes W_{RA}^\dagger W_{RA} |\psi_{AR_{A1}'}\rangle = 1_A \otimes W(1)W(2)^\dagger W(2)W(1)^\dagger |\psi_{AR_{A1}'}\rangle = 1_A \otimes W(1)W(2)^\dagger W(2)W(1)^\dagger W(1)|\psi_{AR_{A1}'}\rangle = |\psi_{AR_{A1}'}\rangle. \] (H18)
Since the partial isometry $W_{R_A}$ works only on $R_A$, we obtain
\[
(W_{R_A} \otimes 1_{AB'C'})(1_{R_A} \otimes V \otimes 1_{B})(1_{R_A} \otimes U \otimes 1_{R_B}C)
= (1_{R_A} \otimes V \otimes 1_{B})(1_{R_A} \otimes U \otimes 1_{R_B}C)(W_{R_A} \otimes 1_{AB'R_BC})
\]
(H19)

Therefore, for $|\tilde{\psi}_{tot}'\rangle := (W_{R_A} \otimes 1_{AB'C'})|\psi_{tot}'\rangle$,
\[
|\tilde{\psi}_{tot}'\rangle = (1_{R_A} \otimes V \otimes 1_{B})(1_{R_A} \otimes U \otimes 1_{R_B}C)
\sum_j \sqrt{p_j} |\tilde{\psi}_{j,AR'}\rangle \otimes |j_{R_a}'\rangle \otimes |\phi_{BRB}\rangle \otimes |\eta_C\rangle
= \sum_j \sqrt{p_j} |\tilde{\psi}_{j,AR]\otimes B'C'}\rangle \otimes |j_{R_a}'\rangle,
\]
(H20)

where $|\tilde{\psi}_{j,AR]\otimes B'C'}\rangle := (1_{R_a}' \otimes V \otimes 1_{B})(1_{R_a}' \otimes U \otimes 1_{R_B}C)|\tilde{\psi}_{j,AR'}\rangle \otimes |\phi_{BRB}\rangle \otimes |\eta_C\rangle$.

Now, we define the desired $\tilde{\psi}_{j,B'C'}$ as $\tilde{\psi}_{j,B'C'} := Tr_{AR'}[|\tilde{\psi}_{j,AR]\otimes B'C'}\rangle$.

Then, since $\{j_{R_a}'\}$ are orthonormal, for $\psi_{B'C'} := Tr_{AR'}[|\tilde{\psi}_{tot}'\rangle$,
\[
\tilde{\psi}_{B'C'} = \sum_j p_j \tilde{\psi}_{j,B'C'}
\]
(H21)

We establish $\tilde{\psi}_{B'C'} = \psi_{B'C'}$ as follows:

\[
\tilde{\psi}_{B'C'} = Tr_{AR'}[|\tilde{\psi}_{tot}'\rangle
= Tr_{AR'}[(W_{R_A} \otimes 1_{AB'C'}|\psi_{tot}'\rangle W_{R_A}^\dagger \otimes 1_{AB'C'})
= Tr_{AR'}[(W_{R_A}^\dagger W_{R_A} \otimes 1_{AB'C'})|\psi_{tot}'\rangle
= Tr_{AR'}([1_{R_A} \otimes V \otimes 1_{B})(1_{R_A} \otimes U \otimes 1_{R_B}C)
\sum_j \sqrt{p_j} |\tilde{\psi}_{j,AR'}\rangle \otimes |j_{R_a}'\rangle \otimes |\phi_{BRB}C\rangle \otimes |\eta_C\rangle
= \sum_j \sqrt{p_j} \psi_{j,AR}\otimes B'C'}\rangle \otimes |j_{R_a}'\rangle
= \psi_{B'C'}
\]
(H22)

Here we use (H5) in (a). Combining (H21) and (H22), we obtain (H7).

Similarly, we can obtain (H8) as follows:

\[
\text{Tr}C[|\tilde{\psi}_{j,B'C'}\rangle = Tr_{AR'}C[|\tilde{\psi}_{j,AR']\otimes B'C'}\rangle
= Tr_{AR'}C[(1_{R_a}' \otimes V \otimes 1_{B})(1_{R_a}' \otimes U \otimes 1_{R_B}C)|\tilde{\psi}_{j,AR'}\rangle \otimes |\phi_{BRB}\rangle \otimes |\eta_C\rangle
= Tr_{AC'}[(V \otimes 1_{B'})\otimes (U \otimes 1_{R_B}C)|\tilde{\psi}_{j,AR'}\rangle \otimes |\phi_{BRB}\rangle \otimes |\eta_C\rangle
= \rho_{j,B'}.
\]
(H23)

Therefore, $\{\tilde{\psi}_{j,B'C'}\}$ actually satisfy (H7) and (H8).

**Appendix I: Derivation of the properties of the variance and expectation values of the conserved quantity $X$**

In this section, we prove (F1)–(F3). We first derive (F1). We evaluate the difference between the left-hand side and the right-hand side of (F1) as follows:

\[
\langle (X_A + X_B)^2 \rangle_U \otimes \xi_B \otimes U^\dagger
= Tr[(X_A + X_B)^2 \xi_B \otimes U^\dagger]
= Tr[X_A + X_B]^2 \xi_B \otimes U^\dagger
= \langle (X_A + X_B)^2 \xi_B \otimes U^\dagger \rangle
= 0
\]
(11)

Here we use $U(X_A + X_B)U^\dagger = X_A + X_B$ in (a).

We next show (F2). Note that

\[
\langle (X_A + X_B)^2 \xi_B \otimes U^\dagger \rangle
= Tr[(X_A + X_B)^2 \xi_B \otimes U^\dagger]
= Tr[(X_A + X_B)^2 \xi_B \otimes U^\dagger]
= \langle (X_A + X_B)^2 \xi_B \otimes U^\dagger \rangle
\]

Combining this and $\langle X_A + X_B \xi_B \otimes U \rangle = (X_A + X_B)U \xi_B \otimes U^\dagger$, which is easily obtained from (F1), we obtain

\[
\langle X_A + X_B \xi_B \otimes U \rangle = V_{\xi_B}(X_A + X_B) = V_{\xi_B}(X_A + X_B).
\]
(13)

From (13), the lower bound for $V_{\xi_B}(X_A) + V_{\xi_B}(X_B)$ is

\[
V_{\xi_B}(X_A) + V_{\xi_B}(X_B)
= V_{\xi_B}(X_A + X_B)
= U_{\xi_B}(X_A + X_B)
= U_{\xi_B}(X_A + X_B)
= V_{\xi_B}'(X_A) + V_{\xi_B}'(X_B) + 2\text{cov}(U_{\xi_B}(X_A + X_B))
\ge V_{\xi_B}'(X_A) + V_{\xi_B}'(X_B) - 2\sqrt{V_{\xi_B}'(X_A) V_{\xi_B}'(X_B)}
= \left(\sqrt{V_{\xi_B}'(X_A) - V_{\xi_B}'(X_B)}\right)^2,
\]
(14)

where $\text{cov}(X : Y) := \langle(X - \langle X \rangle_Y)(Y - \langle Y \rangle_Y)\rangle / 2$ and $\langle X, Y \rangle := XY + YX$. Taking the square root of both sides and applying $\sqrt{X + \sqrt{Y}} \ge \sqrt{X + Y}$ to the left-hand side, we obtain

\[
\sqrt{V_{\xi_B}'(X_A) + V_{\xi_B}(X_B)} \le \sqrt{V_{\xi_B}'(X_A)} + \sqrt{V_{\xi_B}(X_A)} + \sqrt{V_{\xi_B}(X_B)} + \Delta_a,
\]
(15)

We can derive (F3) in the same way as (F2).
Supplementary information for “Universal limitation of quantum information recovery: symmetry versus coherence”

Hiroayasu Tajima$^{1,2}$ and Keiji Saito$^3$

$^1$Graduate School of Informatics and Engineering, The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan
$^2$JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan
$^3$Department of Physics, Keio University, 3-14-1 Hiyoshi, Yokohama, 223-8522, Japan

The supplementary material is organized as follows. In Sec. [Supp.1] we introduce several useful tips about the resource theory of asymmetry. The tips is a generalized version of tips in Appendix [A]. In Sec. [Supp.2] we give a concrete example that quantum coherence alleviates the recovery error. In Sec. [Supp.3] we introduce several additional assumptions with Setup 1. When we use such additional assumptions, we mention them. Note that Setup 1 does not refer to this setup as “Setup 1.” In each section of this Supplementary Material, we use several different additional assumptions as the main text. We also use abbreviations for density operators of pure states like $\delta$. Here we use the purified distance $D_F(\rho, \sigma) := \sqrt{1 - F^2(\rho, \sigma)} = \sqrt{1 - \text{Tr}[\sqrt{\sqrt{\rho} \sqrt{\sigma} \sqrt{\rho} \sqrt{\sigma}}]}$ and abbreviations $\psi_{AR_A} := |\psi_{AR_A}\rangle \langle \psi_{AR_A}|$, $\phi_{BR_B} := |\phi_{BR_B}\rangle \langle \phi_{BR_B}|$ and $I := (\phi_{BR_B}, U)$. Without special notice, we abbreviates $\delta(\psi_{AR_A}, I)$ as $\delta$ as the main text. We also use abbreviations for density operators of pure states like $\eta = |\eta\rangle \langle \eta|$. Hereafter, we refer to this setup as “Setup 1.” In each section of this Supplementary Material, we use several different additional assumptions with Setup 1. When we use such additional assumptions, we mention them. Note that Setup 1 does not contain the conservation law of $X$. When we assume the conservation law of $X$, i.e. $U(X_A + X_B)U^\dagger = X_{A'} + X_{B'}$ for Hermite operators $X_\alpha$ on $\alpha (\alpha = A, B, A', B')$, we say “Setup 1 with the conservation law of $X$.”

![Schematic diagram of Setup 1.](Fig_S1.png)

**FIG. S.1.** Schematic diagram of Setup 1.

**Appendix Supp.1: Tips for resource theory of asymmetry for the case of general symmetry**

In this section, we give a very basic information about the resource theory of asymmetry (RToA) [4][7] for the case of general symmetry.

We firstly introduce covariant operations that are free operations in RToA. Let us consider a CPTP map $\mathcal{E}$ from a system $A$ to another system $A'$ and unitary representations $\{U_{g,A}\}_{g \in G}$ on $A$ and $\{U_{g,A'}\}_{g \in G}$ on $A'$ of a group $G$. The CPTP $\mathcal{E}$ is said to be covariant with respect to $\{U_{g,A}\}_{g \in G}$ and $\{V_{g,A'}\}_{g \in G}$, when the following relation holds:

$$V_{g,A'} \circ \mathcal{E}(\ldots) = \mathcal{E} \circ U_{g,A}(\ldots), \ \forall g \in G,$$

(1.1)
where $U_{g,A}(\ldots) := U_{g,A}(\ldots)U_{g,A}^\dagger$ and $V_{g,A'}(\ldots) := V_{g,A'}(\ldots)V_{g,A'}^\dagger$. Similarly, a unitary operation $U_A$ on $A$ is said to be invariant with respect to $\{U_{g,A}\}_{g \in G}$ and $\{V_{g,A}\}_{g \in G}$, when the following relation holds:

$$V_{g,A} \circ U(\ldots) = U \circ U_{g,A}(\ldots), \quad \forall g \in G,$$

where $U(\ldots) := U(\ldots)U^\dagger$.

Next, we introduce symmetric states that are free states of resource theory of asymmetry. A state $\rho$ on $A$ is said to be a symmetric state when it satisfies the following relation:

$$\rho = U_{g,A}(\rho), \quad \forall g \in G.$$

When a CPTP-map $\mathcal{E}$ is covariant, it can be realized by invariant unitary and symmetric state $\mathcal{E}[6][7]$. To be concrete, when a CPTP map $\mathcal{E}: A \to A'$ is covariant with respect to $\{U_{g,A}\}_{g \in G}$ and $\{U_{g,A'}\}_{g \in G}$, there exist another system $B$, unitary representations $\{U_{g,B}\}_{g \in G}$ and $\{V_{g,B'}\}_{g \in G}$ on $B$ and $B'$ ($AB = A'B'$), a unitary $U_{AB}$ which is invariant with respect to $\{U_{g,A} \otimes U_{g,B}\}_{g \in G}$ and $\{V_{g,A'} \otimes V_{g,B'}\}_{g \in G}$, and a symmetric state $\mu_B$ with respect to $\{U_{g,B}\}_{g \in G}$ such that

$$\mathcal{E}(\ldots) = \text{Tr}_{B'}[U_{AB}(\ldots \otimes \mu_B)U_{AB}^\dagger].$$

**Appendix Supp.II: An example of the error suppression by quantum coherence in information recovery**

In this section, we give a concrete example that large $\mathcal{F}$ actually enables the recovery error $\delta$ to be smaller than $\mathcal{A}/8\Delta_+$. We consider Setup 1 with the conservation law of $X$, i.e., $U(X_A + X_B)U^\dagger = X_{A'} + X_{B'}$. We set $A$ to be a qubit system and $B$ to be a $6M + 1$-level system, where $M$ is a natural number that we can choose freely. We also set $R$ and $R_B$ as copies of $A$ and $B$, respectively. We take $X_A$ and $X_B$ as follows:

$$X_A := |1\rangle_A \langle 1|_A,$$

$$X_B := \sum_{k=-3M}^{3M} k|k\rangle_B \langle k|_B.$$

where $\{|k\rangle_A\}_{k=0}^1$ and $\{|k\rangle_B\}_{k=-3M}^{3M}$ are orthonormal basis of $A$ and $B$.

Under this setup, we consider the case where $A = A'$, $B = B'$, $X_A = X_{A'}$ and $X_B = X_{B'}$. In this case, due to [II.1] and $X_A = X_{A'}$, the equality $\Delta_+ = 1$ holds. Therefore, (II.1) becomes the following inequality:

$$\frac{\mathcal{A}}{2(\sqrt{\mathcal{F}} + 4)} \leq \delta.$$  

Therefore, when $\mathcal{F} = 0$, the error $\delta$ can not be smaller than $\mathcal{A}/8$. Here, we show that when $\mathcal{F}$ is large enough, the error $\delta$ actually becomes smaller than $\mathcal{A}/8$. Let us take $|\psi_{AR_A}\rangle$, $|\phi_{BR_B}\rangle$ and $U$ as

$$|\psi_{AR_A}\rangle = \frac{|0\rangle_A|0\rangle_{R_A} + |1\rangle_A|1\rangle_{R_A}}{\sqrt{2}},$$

$$|\phi_{BR_B}\rangle = \frac{\sum_{k=-M}^{M} |k\rangle_B |k\rangle_{R_B}}{\sqrt{2M + 1}},$$

$$U = \sum_{-2M \leq k \leq 2M} |1\rangle_A |0\rangle_A \otimes |k - 1\rangle_B \langle k\rangle_B + \sum_{-2M - 1 \leq k \leq 2M - 1} |0\rangle_A |1\rangle_A \otimes |k + 1\rangle_B \langle k\rangle_B

+ \sum_{k \leq -2M - 2M < k} |0\rangle_A |0\rangle_A \otimes |k\rangle_B \langle k\rangle_B + \sum_{-2M - 1 < k \leq 2M - 1} |1\rangle_A |1\rangle_A \otimes |k\rangle_B \langle k\rangle_B.$$  

Then, $U$ is a unitary satisfying $U(X_A + X_B) = X_A + X_B$, and the CPTP-map $\mathcal{E}$ implemented by $(U, |\phi_{BR_B}\rangle)$ is expressed as

$$\mathcal{E}(\ldots) = |1\rangle_A |0\rangle_A (\ldots) |0\rangle_A |1\rangle_A + |0\rangle_A |1\rangle_A (\ldots) |1\rangle_A |0\rangle_A.$$  

Due to [II.7] and $\rho_A := \text{Tr}_{R_A}[|\psi_{AR_A}\rangle] = \frac{|0\rangle_A |0\rangle_A + |1\rangle_A |1\rangle_A}{2}$, the quantity $\mathcal{A}$ is equal to 1/2. Here, let us define a recovery CPTP-map $\mathcal{R}_V$ as

$$\mathcal{R}_V(\ldots) := \text{Tr}_{R_B}[V_{AR_A}(\ldots) V_{AR_A}^\dagger].$$
where $V_{AR_B}$ is a unitary operator on $AR_B$ defined as

$$V_{AR_B} := \sum_{-3M+1 \leq k \leq 3M} |0\rangle\langle 1_A \otimes (k-1)\rangle_{R_B} + \sum_{-3M \leq k \leq 3M-1} |1\rangle\langle 0_A \otimes (k+1)\rangle_{R_B}$$

$$+ |0\rangle\langle 1_A \otimes (3M)\rangle_{R_B} + |1\rangle\langle 0_A \otimes (-3M)\rangle_{R_B}.$$  \hfill (II.9)

(Note that the recovery $V_{AR_B}$ is not required to satisfy the conservation law). Then, after $V_{AR_B}$, the total system is in

$$(V_{AR_B} \otimes 1_{BR_A})(U_{AB} \otimes 1_{RA,B})(|\psi_{AR_A}\rangle \otimes |\phi_{BR_B}\rangle)$$

$$= \frac{1}{\sqrt{2(2M+1)}} \sum_{k=-M}^{M} (|0\rangle_A |0\rangle_{R_A}|k-1\rangle_B |k-1\rangle_{R_B} + |1\rangle_A |1\rangle_{R_A}|k+1\rangle_B |k+1\rangle_{R_B})$$

$$= \frac{\sqrt{2M-1}}{\sqrt{2M+1}} |\psi_{AR_A}\rangle \otimes |\tilde{\phi}_{BR_B}\rangle + \frac{1}{\sqrt{2M+1}} |00\rangle_{AR_A} |M, -M\rangle_{BR_B} + |M - 1, -M - 1\rangle_{BR_B}$$

$$+ \frac{1}{\sqrt{2M+1}} |11\rangle_{AR_A} |M+1, M+1\rangle_{BR_B},$$  \hfill (II.10)

where $|\tilde{\phi}_{BR_B}\rangle := \frac{1}{\sqrt{2M+1}} \sum_{k=-M}^{M} |k, k\rangle_{BR_B}$. By partial trace of $BR_B$, we obtain the final state of $AR_A$ as follows:

$$\psi^f_{AR_A} = \frac{2M-1}{2M+1} \psi_{AR_A} + \frac{1}{2M+1} |00\rangle_{AR_A} + \frac{1}{2M+1} |11\rangle_{AR_A}.$$  \hfill (II.11)

Therefore,

$$D_F(\psi^f_{AR_A}, \psi_{AR_A})^2 = 1 - \langle \psi_{AR_A} | \psi^f_{AR_A} \rangle | \psi_{AR_A} \rangle = \frac{2}{2M+1}.$$  \hfill (II.12)

Thus, we obtain

$$\delta \leq \sqrt{\frac{2}{2M+1}}.$$  \hfill (II.13)

Hence, when $M$ is large enough, we can make $\delta$ strictly smaller than $\mathcal{A}/8 = 1/16$. Since $\mathcal{F} = 4V_{\rho_B}(X_B)$ ($\rho_B := Tr_{R_B}(|\phi_{BR_B}\rangle)$), large $M$ means large $\mathcal{F}$. Therefore, when $\mathcal{F}$ is large enough, we can make $\delta$ smaller than $1/16$.

Appendix Supp.III: Tips for the application to Hayden-Preskill model with a conservation law

1. Derivation of (14), (16), and (19) in the main text

In this subsection, we give the detailed description of the scrambling of the expectation values and derivation of (14), (19), and (19) in the main text.
For the readers’ convenience, we firstly review the Hayden-Preskill model with the conservation law of $X$ which is introduced in the section [IV] in the main text. (Fig. 2) The model is a specialized version of Setup 1 with the conservation law of $X$. The specialized points are as follows: 1. $A$, $B$, $A'$ and $B'$ are $k$, $N$, $l$- and $k + N - l$-qubit systems, respectively. 2. We assume that the difference between minimum and the maximum eigenvalue of $X$ is $\epsilon_M$. We refer to the initial state of $X$ as $|\phi\rangle$. Under the above setup, when the conserved quantities are scrambled in the sense of the expectation values, we can derive (14), (16), and (19). Below, we show the derivation. For simplicity of the description, we will use the following expression for real numbers $x$ and $y$:

$$x \approx_{\epsilon} y \iff |x - y| \leq \epsilon.$$  

We also express the expectation values of $X\alpha\ (\alpha = A, B$ and $A')$ as follows:

$$x_A(\rho_A) := \langle X A \rangle_{\rho_A},$$  

$$x_B(\rho_B) := \langle X B \rangle_{\rho_B},$$  

$$x_{A'}(\rho_A, \rho_B, U) := \langle X A' \rangle_{\rho_A},$$  

We show (14), (16), and (19) as the following theorem:

**Theorem 1** Let us take a real positive number $\epsilon$ which is lower than $1/(N + k)^2$, and the set of $(|\psi_{AR\alpha}\rangle, |\phi_{BR\alpha}\rangle, U)$. We refer to the initial state of $A$ as $|\rho_A := Tr_{R\alpha}[\psi_{AR\alpha}]\rangle$, and assume that $|\rho_A, X A\rangle = 0$ and $M := M_{\rho_A}(X A) > 0$ holds. We also assume that $(|\phi_{BR\alpha}\rangle, U)$ satisfies the following relation for an arbitrary state $\rho$ on the support of $\rho_A$:

$$x_{A'}(\rho, \rho_B, U) \approx_{\frac{\epsilon}{2} M_{\gamma}} (x_A(\rho) + x_B(\rho_B)) \times \frac{l}{N + k},$$  

where $\gamma := \left(1 - \frac{l}{N + k}\right)$. Then, the following three inequalities hold:

$$A \geq M_{\gamma}(1 - \epsilon)$$  

(III.6)  

$$\sqrt{F_f} \leq \gamma(N + k)$$  

(III.7)  

$$\Delta_{\text{max}} \leq \gamma k(1 + \epsilon)$$  

(III.8)

And when $\rho_A$ can be decomposed as $\rho_A = \frac{\rho_{0,A} + \rho_{1,A}}{2}$ such that $\rho_{0,A}$ and $\rho_{1,A}$ are eigenstates of $X A\alpha$, the following inequality also holds:

$$A_2 \geq M_{\gamma}(1 - \epsilon)$$  

(III.9)

**Proof:** We firstly point out (III.7) is easily derived by noting that $F_f = 4V_{\rho_B}(X B')$ and that the number of qubits in $B'$ is $N + k - l$, which is equal to $(N + k)\gamma$.

To show (III.6), (III.8) and (III.9), let us take an arbitrary decomposition $\rho_A = \sum_j p_j \rho_{j,A}$, and evaluate $|\Delta_j|$ for the decomposition as follows:

$$|\Delta_j| = \left| x_A(\rho_{j,A}) - x_{A'}(\rho_{j,A}, \rho_B, U) \right| - \left| \sum_j p_j x_A(\rho_{j,A}) - \sum_j p_j x_{A'}(\rho_{j,A}, \rho_B, U) \right|$$

$$\approx_{\frac{\epsilon}{2} M_{\gamma}} \left| x_A(\rho_{j,A}) - (x_A(\rho_{j,A}) + x_B(\rho_B)) \frac{l}{N + k} - \sum_j p_j x_A(\rho_{j,A}) + \sum_j p_j (x_A(\rho_{j,A}) + x_B(\rho_B)) \frac{l}{N + k} \right|$$

$$= \left| x_A(\rho_{j,A}) - \sum_j p_j x_A(\rho_{j,A}) \right| \gamma.$$  

(III.10)

To derive (III.6) from the above evaluation, let us choose a decomposition $\rho_A = \sum_j p_j \rho_{j,A}$ where each $\rho_{j,A}$ is in eigenspace of $X A\alpha$. We can choose such a decomposition due to $|\rho_A, X A\rangle = 0$. Then, $\sum_j p_j |x_A(\rho_{j,A}) - \sum_j p_j x_A(\rho_{j,A})| = M$ holds. Applying (III.10) to this decomposition, we obtain (III.6):

$$A \geq \sum_j p_j |\Delta_j| \geq M_{\gamma} - M_{\gamma}\epsilon.$$  

(III.11)
And when $\rho_A$ can be decomposed as $\rho_A = \frac{\rho_0 + \rho_1}{2}$ where $\rho_0$ and $\rho_1$ are eigenstates of $X_A$, applying (III.10) to the decomposition $\rho_A = \frac{\rho_0 + \rho_1}{2}$, we obtain (III.9) because of $\sum_{j=1}^{N} \frac{1}{2} \left| x_A(\rho_{j,A}) - \sum_{j=0,1} \frac{1}{2} x_A(\rho_{j,A}) \right| = M$:

$$A_2 \geq \frac{1}{2} \sum_{j=0,1} |\Delta_j| \geq M\gamma - M\gamma e.$$ 

Similarly, we can derive (III.8) as follows.

$$\Delta_{\max} = \max_{\{p_j, p_j,A\}} |\Delta_j|$$

$$\leq \max_{\{p_j, p_j,A\}} \left( x_A(\rho_{j,A}) - \sum_j p_j x_A(\rho_{j,A}) \right) \gamma + M\gamma e,$$

$$\leq \max_{\{p_j, p_j,A\}} \left( x_A(\rho_{j,A}) - \sum_j p_j x_A(\rho_{j,A}) \right) \gamma(1 + \epsilon),$$

$$\leq D_{\max} \gamma(1 + \epsilon),$$

$$\leq k\gamma(1 + \epsilon).$$ (III.13)

where $\{p_j, p_j,A\}$ runs over all possible decompositions of $\rho_A$.

2. Derivation of (15) in the main text

In this subsection, we derive (15) based on the assumptions introduced in the main text. We show that the assumptions actually hold in the Haar random unitary with the conservation law of $X$ in the next subsection. We show (15) as the following theorem:

**Theorem 2** Let us take a real positive number $\epsilon$ which is lower than $1/(N + k)^2$, and a real non-negative number $x$. We define projections $P^A_{\epsilon B}$, $P^A_x$ and $P^B_x$ on $AB$ and $B$ as follows:

$$P^A_{\epsilon B} := \sum_{\lambda_{min}(X_A + X_B) + x \leq x_A(i) + x_B(j) \leq \lambda_{max}(X_A + X_B) - x} |i, a\rangle \langle i, a| \otimes |j, b\rangle \langle j, b|,$$

$$P^A_x := \sum_{\lambda_{min}(X_A) + x \leq x_A(i) \leq \lambda_{max}(X_A) - x} |i, a\rangle \langle i, a|,$$ (III.15)

$$P^B_x := \sum_{\lambda_{min}(X_B) + x \leq x_B(j) \leq \lambda_{max}(X_B) - x} |j, b\rangle \langle j, b|,$$ (III.16)

where $\lambda_{min}(X_A + X_B)$ and $\lambda_{max}(X_A + X_B)$ are the minimum and maximum eigenvalues of $X_A + X_B$, and $|i, a\rangle$ and $|j, b\rangle$ are eigenvectors of $X_A$ and $X_B$ whose eigenvalues are $x_A(i)$ and $x_B(j)$. We also take pure states $|\psi_{\epsilon A}\rangle$ and $|\phi_{\epsilon B}\rangle$ on $AR_A$ and $BR_B$. We assume that $\rho_A := \text{Tr}_{R_A} |\psi_{\epsilon A}\rangle \langle \psi_{\epsilon A}|$ satisfies $[X_A, \rho_A] = 0$ and $\rho_B := \text{Tr}_{R_B} |\phi_{\epsilon B}\rangle \langle \phi_{\epsilon B}|$ satisfies $[X_B, \rho_B] = 0$ and

$$V_{\epsilon B}(X_B) \leq N/4.$$ (III.17)

We also assume that $\rho_A$ and $\rho_B$ satisfy at least one of the following two inequalities:

$$\left\| \frac{P^A_{\epsilon B} \rho_A P^A_{\epsilon B}}{\text{Tr}(\rho_A) P^A_x} - \rho_A \right\|_1 \leq \frac{\epsilon^2}{3(N + k)^2},$$ (III.18)

$$\left\| \frac{P^B_{\epsilon B} \rho_B P^B_{\epsilon B}}{\text{Tr}(\rho_B) P^B_x} - \rho_B \right\|_1 \leq \frac{\epsilon^2}{3(N + k)^2}.$$ (III.19)

Let us take a unitary operation $U$ on $AB$ satisfying the conservation law of $X$. We assume that $U$ satisfies the following two relations for arbitrary $|i, a\rangle \otimes |j, b\rangle$ satisfying $P^A_{\epsilon B} |i, a\rangle \otimes |j, b\rangle = |i, a\rangle \otimes |j, b\rangle$:

$$V'_{\epsilon B'}(|i, a\rangle \otimes |j, b\rangle) \leq \frac{1 + \epsilon}{4} \min\{l, \gamma(N + k)\},$$ (III.20)

$$x_A(|i, 1\rangle, |j, b\rangle, U) \approx \frac{1}{N + k} (x_A(|i, a\rangle) + x_B(|j, b\rangle)).$$ (III.21)
Then, the following inequality hold:
\[ \sqrt{\mathcal{F}_f} \leq 2(1 + \epsilon)\sqrt{\gamma(N + k^2)}. \] (III.22)

**Proof of (III.22) (= in the main text):**  We use the following equation which is valid for arbitrary probability \{q_k\} and operator \(W\):
\[
V_{\sum_k q_k \sigma_k}(W) = \sum_k q_k V_{\sigma_k}(W) + V(q_k)\{\{W\}_\sigma\}. 
\] (III.23)
where \(V(q_k)\{\{W\}_\sigma\} := \sum_k q_k \langle W \rangle_{\sigma_k}^2 - (\sum_k q_k \langle W \rangle_{\sigma_k})^2\) is the variance of the expectation values \{\langle W \rangle_{\sigma_k}\} with the probability \{q_k\}. We prove (III.23) as follows:
\[
V_{\sum_k q_k \sigma_k}(W) = \sum_k q_k \langle W^2 \rangle_{\sigma_k} - (\langle W \rangle_{\sum_k q_k \sigma_k})^2 
= \sum_k q_k \langle W^2 \rangle_{\sigma_k} - \sum_k q_k \langle W \rangle_{\sigma_k}^2 + \sum_k q_k \langle W \rangle_{\sigma_k}^2 - (\langle W \rangle_{\sum_k q_k \sigma_k})^2 
= \sum_k q_k \langle W^2 \rangle_{\sigma_k} - (\langle W \rangle_{\sigma_k})^2 + \sum_k q_k \langle W \rangle_{\sigma_k}^2 - (\sum_k q_k \langle W \rangle_{\sigma_k})^2 
= \sum_k q_k V_{\sigma_k}(W) + V(q_k)\{\{W\}_\sigma\}. 
\] (III.24)

We also use the following relation for arbitrary states \(\rho\) and \(\tilde{\rho}\) and an arbitrary operator \(Y\):
\[
V_{\rho}(Y) \approx 3\|Y\|_\infty \|\rho - \tilde{\rho}\|_1 V_{\tilde{\rho}}(Y). 
\] (III.25)
We prove (III.25) as follows:
\[
|V_{\tilde{\rho}}(Y) - V_{\rho}(Y)| \leq |\langle Y^2 \rangle_{\tilde{\rho}} - \langle Y^2 \rangle_{\rho}| + |\langle Y \rangle_{\tilde{\rho}} - \langle Y \rangle_{\rho}| 
\leq \|Y\|_\infty^2 \|\rho - \tilde{\rho}\|_1 + 2 \max\{\|\langle Y \rangle_{\tilde{\rho}}\|, \|\langle Y \rangle_{\rho}\|\}\|Y\|_\infty \|\rho - \tilde{\rho}\|_1 
\leq 3\|Y\|_\infty \|\rho - \tilde{\rho}\|_1. 
\] (III.26)

Let us derive (III.22) First we consider the case when (III.19) holds. Hereafter we use the abbreviations \(\tilde{\rho}_B := \frac{\rho_B \rho_B}{\text{Tr}[\rho_B \rho_B^\dagger]}\) and \(\tilde{\rho}_B^\dagger := \text{Tr}_A[U(\rho_A \otimes \tilde{\rho}_B)U^\dagger]\). Then, we obtain
\[
\|\rho_B' - \tilde{\rho}_B'\|_1 \leq \|\rho_A \otimes \rho_B - \rho_A \otimes \tilde{\rho}_B\|_1 \leq \frac{\epsilon^2}{3(N + k^2)}. 
\] (III.27)
Combining (III.25), (III.27) and \(\|X_B\|_\infty = \gamma(N + k)\), we obtain
\[
V_{\rho_B'}(X_B') \approx \gamma^2 \epsilon^2 V_{\tilde{\rho}_B'}(X_B') 
\] (III.28)
Note that the state \(\rho_A \otimes \tilde{\rho}_B\) on \(AB\) can be written in the following form:
\[
\rho_A \otimes \tilde{\rho}_B = \sum_{i,a,j,b} r_{i,a,ARj,b,B} |i, a\rangle\langle i, a| \otimes |j, b\rangle\langle j, b|. 
\] (III.29)
Therefore, we can divide \(V_{\rho_B'}(X_B')\) into two parts:
\[
V_{\rho_B'}(X_B') = \sum_{i,a,j,b} r_{i,a,ARj,b,B} V_{\rho_B'}(X_B') + V_{\{r_{i,a,ARj,b,B}\}}(\{\langle X_B'\rangle_{\rho_B'}\}_{i,a,j,b}). 
\] (III.30)
Due to (III.21) and \(P_{AB} \rho_A \otimes \tilde{\rho}_B P_{AB} = \rho_A \otimes \tilde{\rho}_B, \langle X_B'\rangle_{\rho_B'} = \gamma(x_A(i) + x_B(j)) + \epsilon_{i,a,j,b}\) satisfies
\[
\langle X_B'\rangle_{\rho_B'} = \gamma(x_A(i) + x_B(j)) + \epsilon_{i,a,j,b} 
\] (III.31)
where $|e_{i,a,j,b}| \leq \epsilon$. Therefore, due to $\sqrt{V_{(A)}}(z_k + \varepsilon_k) \leq \sqrt{V_{(A)}}(\{z_k\}) + \sqrt{V_{(A)}}(\{\varepsilon_k\})$, we obtain

$$\sqrt{V_{\{r_{i,a}r_{j,b}\}}(\{(X_B')\rho_{B'}^{r_{i,a}r_{j,b}}\})} \leq \sqrt{V_{\{r_{i,a}r_{j,b}\}}(\{(X_B')\rho_{B'}^{r_{i,a}r_{j,b}}\})} + \epsilon$$

$$\leq \gamma \sqrt{V_{\{r_{i,a}\}}(\{(X_B')\rho_{B'}^{r_{i,a}r_{j,b}}\})} + \epsilon$$

$$\leq \gamma \sqrt{\frac{N+k^2}{2}} + 2 \epsilon$$

(III.32)

Here we used $V_{\{r_{i,a}\}}(\{(X_B')\rho_{B'}^{r_{i,a}}\}) \leq k^2/4$ and $V_{\{r_{j,b}\}}(\{(X_B')\rho_{B'}^{r_{i,a}r_{j,b}}\}) = V_{\{r_{j,b}\}}(X_B) \leq N/4 + \gamma^2 \epsilon^2$, which is given by (III.17) and (III.28), in the last line.

Combining (III.20), (III.28), (III.30) and (III.32), we obtain (III.22) as follows

$$\sqrt{\mathcal{F}_f} = 2 \sqrt{V_{\rho_B'(X_B')}} + 2 \gamma \epsilon$$

$$\leq 2 \left( \mathbf{\sum}_{i,a,j,b} r_{i,a}r_{j,b} V_{\rho_B'(r_{i,a}r_{j,b})}(X_B') + \sqrt{V_{\{r_{i,a}r_{j,b}\}}(\{(X_B')\rho_{B'}^{r_{i,a}r_{j,b}}\})} \right) + 2 \gamma \epsilon$$

$$\leq 2 \left( \mathbf{\sum}_{i,a,j,b} \frac{1}{N+k+1} \min \{ l, \gamma(N+k) \} + \gamma \sqrt{N+k^2} + 2 \epsilon \right)$$

$$\leq 2 \frac{1+\epsilon}{N+k+1} \sqrt{\gamma(N+k^2)}.$$ (III.33)

Here we use $\epsilon \leq 1/(N+k)^2$, $M \leq k$ and $\frac{1}{N+k+1} \leq \gamma \leq 1$ (since now we treat the case of $l < N+k$) in the last line.

Next, we consider the case when (III.18) holds. In this case we use the abbreviations $\hat{\rho}_A := \frac{P_A P_A P_A}{\text{Tr}(P_A P_A P_A)}$ and $\hat{\rho}_B' := \text{Tr}_A[U(\hat{\rho}_A \otimes \rho_B)U^\dagger]$. Then, we obtain

$$\|\rho_{B'}' - \hat{\rho}_B'\|_1 \leq \|\rho_A \otimes \rho_B - \hat{\rho}_A \otimes \rho_B\|_1 \leq \frac{\epsilon}{3(N+k)^2}.$$ (III.34)

Therefore, from (III.25) and $\|X_B\|_\infty = \gamma(N+k)$, we obtain

$$V_{\rho_{B'}'}(X_B') \approx \gamma \epsilon \ V_{\rho_{B'}'}(X_B').$$ (III.35)

Note that the state $\hat{\rho}_A \otimes \rho_B$ on $AB$ can be written in the following form:

$$\hat{\rho}_A \otimes \rho_B = \sum_{i,a,j,b} r'_{i,a}r'_{j,b} (i,a) \langle i,a \rangle \otimes (j,b) \langle j,b \rangle.$$ (III.36)

Therefore, we can divide $V_{\rho_{B'}'}(X_B')$ into two parts:

$$V_{\rho_{B'}'}(X_B') = \sum_{i,a,j,b} r'_{i,a}r'_{j,b} V_{\rho_{B'}'}(i,a) \langle i,a \rangle + V_{\rho_{B'}'}(j,b) \langle j,b \rangle.$$ (III.37)

Due to $P_A P_B \hat{\rho}_A \otimes \rho_B P_A P_B = \hat{\rho}_A \otimes \rho_B$, we can derive the following inequality in the same manner as the derivation of (III.32):

$$\sqrt{V_{\{r_{i,a}r'_{j,b}\}}(\{(X_B')\rho_{B'}^{r_{i,a}r'_{j,b}}\})} \leq \gamma \sqrt{\frac{N+k^2}{2}} + 2 \gamma \epsilon.$$ (III.38)

Combining (III.20), (III.35), (III.37) and (III.38), we obtain (III.22) again:

$$\sqrt{\mathcal{F}_f} \leq 2(1+\epsilon) \sqrt{\gamma(N+k^2)}.$$ (III.39)
3. Derivation of a tighter variation of \[18\]

Here, let us derive a tighter variation of \[18\] that we use in Figure 6 in the main text. Substituting (III.8), (III.9), (III.15) into (G6), we obtain

\[
\frac{1 - \epsilon}{1 + \epsilon} \frac{\gamma M}{\gamma k + 2\sqrt{\gamma(N + k^2)}} \leq \delta
\]  

(III.40)

To interpret the meaning of this inequality, we consider the case of \(k = \sqrt{N}\) and \(M = k/2\) (we can take such an \(M\) by considering a relevant \(\rho_A\) and its decomposition, e.g., \(\rho_A = (\rho_x^{\text{max}} + \rho_0^{\text{max}})/2\), where \(\rho_x^{\text{max}}\) is the maximally mixed state of the eigenspace of \(X_A\) whose eigenvalues is \(x\)). Then, we obtain

\[
\frac{\text{const.}}{1 + \frac{2\sqrt{2}}{\sqrt{\epsilon}}} \leq \delta,
\]  

(III.41)

where the const. is a real number larger than 0.49. This inequality is twice tighter than \[18\] in the main text.

4. Evaluation of expectation values and variance of conserved quantity in Haar random unitary with the conservation law

In this subsection, we show that when \(U\) is a typical Haar random unitary with the conservation law of \(X\), the assumptions (III.5), (III.17), (III.20) and (III.21) which are the assumptions used in the previous subsection actually hold. To prove them explicitly, we define the Haar random unitary with the conservation law of \(X\) in the black hole model. In this subsection, we assume that each operator \(X_i\) on each \(i\)-th qubit is the same, and that \(X_\alpha = \sum_{i \in \alpha} X_i\) \((\alpha = A, B, A' and B')\). We also assume that \(D_{X_i} = 1\) and the ground eigenvalue of each \(X_i\) is 0. We remark that under these assumptions, (III.17) clearly holds whenever \(\rho_B\) is a non-zero temperature Gibbs state \(e^{-\beta X_B}/\text{Tr}[e^{-\beta X_B}]\) for \(0 \leq \beta < \infty\).

Let us refer to the eigenspace of \(X_A + X_B\) whose eigenvalue is \(m\) as \(\mathcal{H}^{(m)}\). Then, the Hilbert space of \(AB\) is written as

\[
\mathcal{H}_{AB} = \oplus_{m=0}^{N+k} \mathcal{H}^{(m)}.
\]  

(III.42)

In this model, \(X_A + X_B = X_{A'} + X_{B'} = \sum_h X_h\) holds (\(X_h\) is the operator of \(X\) on the \(h\)-th qubit), and thus \(U\) satisfying (2) is also written as

\[
U = \oplus_{m=0}^{N+k} U^{(m)},
\]  

(III.43)

where \(U^{(m)}\) is a unitary operation on \(\mathcal{H}^{(m)}\). We refer to the unitary group of all unitary operations on \(\mathcal{H}^{(m)}\) as \(\mathcal{U}^{(m)}\), and refer to the Haar measure on \(\mathcal{U}^{(m)}\) as \(\mathcal{H}^{(m)}\). Then, we can define the product measure of the Haar measures \(\{\mathcal{H}^{(m)}\}_{m=0}^{N+k}\) as follows:

\[
\mathcal{H}^{\mathcal{M}^{(m)}} := \times_{m=0}^{N+k} \mathcal{H}^{(m)},
\]  

(III.44)

where \(\mathcal{M}^{(m)} := \{0, 1, \ldots, M + k\}\). The measure \(\mathcal{H}^{\mathcal{M}^{(m)}}\) is a probabilistic measure on the following unitary group on \(\mathcal{H}_{AB}\):

\[
\mathcal{U}^{\mathcal{M}^{(m)}} := \times_{m=0}^{N+k} \mathcal{U}^{(m)}.
\]  

(III.45)

Since every \(U \in \mathcal{U}^{\mathcal{M}^{(m)}}\) satisfies \(U(X_A + X_B)U^\dagger = X_{A'} + X_{B'}\), we refer to \(U\) chosen from \(\mathcal{U}^{\mathcal{M}^{(m)}}\) with the measure \(\mathcal{H}^{\mathcal{M}^{(m)}}\) as “the Haar random unitary with the conservation law of \(X\).”

Additionally, for the later convenience, we also define the following subspace of \(\mathcal{M}^{(m)}\):

\[
\mathcal{M}_s := \{s, s + 1, \ldots, N + k - s\},
\]  

(III.46)

and the following products of Haar measures and unitary groups

\[
\mathcal{H}^{\mathcal{M}_s} := \times_{m \in \mathcal{M}_s} \mathcal{H}^{(m)},
\]  

(III.47)

\[
\mathcal{U}^{\mathcal{M}_s} := \times_{m \in \mathcal{M}_s} \mathcal{U}^{(m)}.
\]  

(III.48)
In this subsection, hereafter we study the property of the Haar random unitaries with the conservation law of $X$. In particular, we show that the assumptions used in Theorems 1 and 2 actually hold in the above Haar random unitary model. For clarity, below we illustrate several propositions for specific parameter regions that are derived from the results obtained in this subsection. We stress that the results given in this subsection themselves are more general than the following propositions.

**Proposition 1:** When the parameters $N$, $k$ and $\epsilon$ satisfy $N \geq 10^3$, $k \leq 10N$ and $1/(N + k)^3 \leq \epsilon \leq 1/(N + k)^2$, and when the initial state $\rho_B$ of the black hole $B$ is the maximally mixed state, the following relation holds for an arbitrary state $\rho$ on $A$ satisfying $M \geq 1/(N + k)$:

$$\text{Prob}_{U \sim H^\text{Haar}_k} \left[ x_A'(\rho, \rho_B, U) \approx \frac{\sqrt{\frac{e}{2}}}{\sqrt{2}} (x_A(\rho) + x_B(\rho_B)) \times (1 - \gamma) \right] \geq 1 - e^{-(N+k)^2} \quad (\text{III.49})$$

Therefore, the relation (III.5), which is the assumption used in Theorem 1, holds with the probability larger than $1 - e^{-(N+k)^2}$ in this case.

**Proposition 2:** When the parameters $N$, $k$ and $\epsilon$ satisfy $N \geq 10^3$, $47 < k \leq 10N$, and $1/(N + k)^3 \leq \epsilon \leq 1/(N + k)^2$, and when $\rho_A$ satisfies $M \geq 1/(N + k)$ and its support is included by $H_{A}^{23 \leq x \leq k-23}$ ($H_{A}^{23 \leq x \leq k-23}$ is the subspace of $H_A$ spanned by eigenstates of $X_A$ whose eigenvalues are larger than 13 and lower than $k - 13$), the following relation holds for an arbitrary state $\rho_B$ on $B$:

$$\text{Prob}_{U \sim H^\text{Haar}_k} \left[ x_A'(\rho, \rho_B, U) \approx \frac{\sqrt{\frac{e}{2}}}{\sqrt{2}} (x_A(\rho) + x_B(\rho_B)) \times (1 - \gamma) \right] \geq 1 - e^{-(N+k)^2} \quad (\text{III.50})$$

Therefore, the relation (III.5), which is the assumption used in Theorem 1, holds with the probability larger than $1 - e^{-(N+k)^2}$ in this case.

**Remark on Proposition 2:** Note that the support of $\rho_{A,*} := (\rho_{k-24}^\text{max} + \rho_{24}^\text{max})/2$ ($\rho_x^\text{max}$ is the maximally mixed state of the eigenspace of $X_A$ corresponding to eigenvalue $x$) is included in $H_{A}^{23 \leq x \leq k-23}$. Therefore, the following inequality holds:

$$1 - \epsilon \frac{k - 24}{1 + \epsilon 2(N + 2k)} \leq \delta \quad (\text{III.51})$$

Hence, when $N \geq 10^3$, $10^3 \leq k \leq 10N$, regardless of the initial state of the black hole, the inequality (21) in the main text holds for a typical dynamics $U$ on $A B$:

$$\frac{\text{const} \times k}{1 + N/2k} \leq \delta \quad (\text{III.52})$$

where const. is a positive number larger than 1/9.

**Proposition 3:** When the parameters $N$, $k$, and $\epsilon$ satisfy $N \geq 10^3$ and $1/(N + k)^3 \leq \epsilon \leq 1/(N + k)^2$, the following two relations for arbitrary $|i, a|$ and $|j, b|$ of eigenstates of $X_A$ and $X_B$ whose eigenvalues are $i$ and $j$ satisfying $18 \leq i + j \leq N + k - 18$:

$$\text{Prob}_{U \sim H^\text{Haar}_k} \left[ V_{A'} |i, a, j, b, U \rangle \langle x_A' | \leq (1 + \epsilon) \frac{\min\{(1 - \gamma)(N + k), \gamma(N + k)\}}{4} \right] \geq 1 - 4 \exp\left(-10(N + k)\right), \quad (\text{III.53})$$

$$\text{Prob}_{U \sim H^\text{Haar}_k} \left[ x_A'(i, a) \langle j, b, U \rangle \approx \epsilon (x_A(i) + x_B(j)) \times (1 - \gamma) \right] \geq 1 - e^{-(N+k)^2} \quad (\text{III.54})$$

where $A' = A$, $B'$. Therefore, in the above parameter region of $N$, $k$ and $\epsilon$, (III.20) and (III.21) which are the assumptions used in Theorem 2 actually hold for $x = 18$ with very high probability.

**Remark on Proposition 3:** In the above parameter region, due to the inequalities (III.53) and (III.54), whenever the initial states $\rho_A$ and $\rho_B$ satisfy (III.17) and at least one of (III.19) and (III.18) for $x = 18$, the inequality (III.22) ($= (15)$ in the main text) holds. We give two examples of such $\rho_A$ and $\rho_B$ and show that (18) in the main text actually holds for the examples. The first example is that $\rho_A$ is $\rho_{A,x} := (\rho_{k-19}^\text{max} + \rho_{19}^\text{max})/2$ and $\rho_B$ is an arbitrary non-zero temperature Gibbs state $e^{-\beta X_B}/\text{Tr}[e^{-\beta X_B}]$ for $0 \leq \beta < \infty$. Due to $X_B = \sum_{i \in B} X_i$, the Gibbs state $e^{-\beta X_B}/\text{Tr}[e^{-\beta X_B}]$ for $0 \leq \beta < \infty$ clearly satisfies (III.17). Also, $\rho_{A,x}$ satisfies (III.18) for $x = 18$. 

Therefore, for these initial states the inequality (III.22) holds, and thus we can derive the following inequality from (6):

\[
\frac{1 - \epsilon}{1 + \epsilon} \frac{\gamma(\frac{1}{2} - 19)}{2(\gamma k + 2\sqrt{\gamma(N+k^2)})} \leq \delta \tag{III.55}
\]

By setting \(k = \sqrt{N}\) and \(N > 10^{10}\), the inequality (18) in the main text holds for typical dynamics \(U\):

\[
\frac{\text{const.}}{1 + \frac{2\sqrt{\gamma}}{\sqrt{N}}} \leq \delta, \tag{III.56}
\]

where const. is a positive constant larger than 0.24. We can make this constant larger than 0.48 by using (G6) instead of (6).

The second example is given for the case of \(k \leq 10N\) (Note that Proposition 3 does not require this restriction). The example is that \(\rho_A = \rho_{A, \epsilon} := (\rho_A^{\text{max}} + \rho_B^{\text{max}})/2\) and \(\rho_B\) is an enoughly-high temperature Gibbs state \(\frac{e^{-\beta X_B}}{\text{Tr}[e^{-\beta X_B}]}\) for \(0 \leq \beta \leq \log 2\). Again, due to \(X_B = \sum_{i \in B} x_i\), the Gibbs state satisfies (III.17). And \(\rho_B\) satisfies (III.19) for \(x = 18\). (Proof: Note the following inequalities:

\[
\left\| \frac{P_B \rho_B P_B}{\text{Tr}(\rho_B P_B^2)} - \rho_B \right\|_1 = \sum_{m=0}^{18} \frac{e^{-\beta m} + e^{-\beta(N-m)}}{(e^{-\beta} + 1)^N} \times N C_m, \tag{III.57}
\]

\[
\frac{\epsilon^2}{3(N+k)^2} \geq \frac{1}{3 \times 11^8 \times N^8}. \tag{III.58}
\]

Here we used \(k \leq 10N\) and \(1/(N+k)^3 \leq \epsilon\) in the second inequality. Therefore, to show that \(\rho_B\) satisfies (III.19) for \(x = 18\), we only have to show

\[
3 \times 11^8 \times N^8 \sum_{m=0}^{18} \frac{e^{-\beta m} + e^{-\beta(N-m)}}{(e^{-\beta} + 1)^N} \times N C_m \leq 1. \tag{III.59}
\]

Due to the left-hand side of the above inequality is monotonically decreasing with \(N\) for \(N \geq 10^3\), and is monotonically increasing with \(\beta\) for \(0 \leq \beta \leq \log 2\). Therefore, because of

\[
\left(3 \times 11^8 \times N^8 \sum_{m=0}^{18} \frac{e^{-\beta m} + e^{-\beta(N-m)}}{(e^{-\beta} + 1)^N} \times N C_m \right)_{\beta = \log 2, N = 10^3} \approx 2 \times 10^{-111}, \tag{III.60}
\]

\(\rho_B\) satisfies (III.19) for \(x = 18\). (Proof end) Hence, for these initial states the inequality (III.22) holds, and thus we can derive the following inequality from (6):

\[
\frac{1 - \epsilon}{1 + \epsilon} \frac{\gamma k}{2(\gamma k + 2\sqrt{\gamma(N+k^2)})} \leq \delta \tag{III.61}
\]

By setting \(k = \sqrt{N}\) and \(N > 10^3\), the inequality (18) in the main text holds for typical dynamics \(U\) again, where the positive constant const. is larger than 0.24. Again, we can make this constant larger than 0.48 by using (G6) instead of (6).

Now, we show our results giving the above propositions one by one. The first theorem guarantees that if we take average over all Haar random unitary, the expectation values of \(X_{A'}\) and \(X_{B'}\) are in proportion to the number of qubits in \(A'\) and \(B'\), respectively.

**Theorem 3** For the quantity \(x_\alpha (\alpha = A, B, A')\) in Theorem 2 and arbitrary \(\rho\) and \(\rho_B\) on \(A\) and \(B\), the following equality holds:

\[
x_{A'}(\rho, \rho_B, U) = (x_A(\rho) + x_B(\rho_B)) \frac{l}{N + k}, \tag{III.62}
\]

where \(f(U)\) is the average of the function \(f\) with the product Haar measure \(H^{M_{x_{A'}}}_{\times}\). Additionally, when the support of \(\rho \otimes \rho_B\) is included in the subspace \(H^{M_{x_{A'}}}_{\times} := \otimes_{m \in M_{x_{A'}}} H^{(m)}\), the following equality holds:

\[
x_{A'}(\rho, \rho_B, U)|_{H^{M_{x_{A'}}}_{\times}} = (x_A(\rho) + x_B(\rho_B)) \frac{l}{N + k}, \tag{III.63}
\]
where $\tilde{U}$ is a unitary which is described as $\tilde{U} = (\oplus_{m \in M_s} U^{(m)}) \oplus (\oplus_{n \notin M_s} I^{(m)})$ where $U^{(m)} \in U^{(m)}$, and $\int \langle \tilde{U} \rangle_{H_x^{M_s}}$ is the average of the function $f$ with the product Haar measure $H_x^{M_s}$.

**Proof:** We refer to the state of the $h$-th qubit in $A'B'$ after $U$ as $\rho^f_h$. The state $\rho^f_h$ satisfies

$$\rho^f_h = \text{Tr}_{-h}[U(\rho \otimes \rho_B)U^\dagger],$$

(III.64)

where $\text{Tr}_{-h}$ is the partial trace of the qubits other than the $h$-th qubit. Note that the following equality holds:

$$x_{A'}(\rho, \rho_B, U) = \sum_{h \in A'} \langle X_h \rangle_{\rho^f_h},$$

(III.65)

where $X_h$ is the operator of $X$ on the $h$-th qubit. Therefore, in order to show (III.62), we only have to show

$$\overline{\rho^f_h} = \overline{\rho^f_h} \forall h, h'.$$

(III.66)

To show (III.66), we note that the swap gate $S_{h, h'}$ between the $h$-th and the $h'$-th qubits can be written in the following form:

$$S_{h, h'} = \oplus_{m \in M_{all}} S_{h, h'}^{(m)},$$

(III.67)

where each $S_{h, h'}^{(m)}$ is a unitary on $H_{(m)}$. Therefore, for any $U \in U_x^{M_{all}}$, the unitary $S_{h, h'} U$ also satisfies $S_{h, h'} U \in U_x^{M_{all}}$. With using this fact and the invariance of the Haar measure, we can derive (III.66) as follows:

$$\overline{\rho^f_h} = \text{Tr}_{-h}[\int_{H_x^{M_{all}}} \mu U(\rho \otimes \rho_B)U^\dagger]$$

$$= \text{Tr}_{-h'}[\int_{H_x^{M_{all}}} \mu U(\rho \otimes \rho_B)U^\dagger S_{h, h'}^{(m)}]$$

$$= \text{Tr}_{-h'}[\int_{H_x^{M_{all}}} \mu S_{h, h'} U(\rho \otimes \rho_B)(S_{h, h'} U)^\dagger]$$

$$= \text{Tr}_{-h'}[\int_{H_x^{M_{all}}} \mu U^\dagger(\rho \otimes \rho_B)U^\dagger]$$

$$= \overline{\rho^f_h}$$

(III.68)

Therefore, we have obtained (III.62).

We can also derive (III.63) in a very similar way. For an arbitrary unitary $V \in U_x^{M_{all}}$, let us define $\tilde{V} \in U_x^{M_s}$ as follows:

$$\tilde{V} = (\oplus_{m \in M_s} V^{(m)}) \oplus (\oplus_{n \notin M_s} I^{(m)}),$$

(III.69)

where $\{V^{(m)}\}$ are defined as $V = \oplus_{m \in M_{all}} V^{(m)}$. Note that when $\rho \otimes \rho_B$ is included in $H_{M_s}$, we can substitute $\tilde{S}_{h, h'}$ and $\tilde{U}$ for $S_{h, h'}$ and $U$ in the above derivation of (III.62). By performing this substitution, we obtain (III.63).

In the next theorem, we show that under a natural assumption, the value of $x_{A'}(\rho, \rho_B, U)$ with a Haar random unitary $U$ is almost equal to its average with very high probability.

**Theorem 4** For the quantity $x_{A'}$ in Theorem 7, an arbitrary positive number $t$, an arbitrary non-negative integer $s$, and arbitrary states $\rho$ and $\rho_B$ on $A$ and $B$ which satisfy that the support of $\rho \otimes \rho_B$ is included in the subspace $H_{M_s} := \oplus_{m \in M_s} H_{(m)}$ for $M_s := \{s, s+1, \ldots, N + k - s\}$, the following relation holds:

$$\text{Prob}_{U \sim H_x^{M_{all}}}[|x_{A'}(\rho, \rho_B, U) - \overline{x_{A'}(\rho, \rho_B, U)}| > t] \leq 2 \exp\left(-\frac{(N+kC_s-2)t^2}{48l^2}\right),$$

(III.70)

Here $\text{Prob}_{U \sim H_x^{M_{all}}}[...]$ is the probability that the event (...) happens when $U$ is chosen from $U_x^{M_{all}}$ with the measure $H_x^{M_{all}}$. 
This theorem implies that when $U$ is a typical Haar random unitary with the conservation law of $X$, the assumption $\text{(III.5)}$ actually holds with very high probability. To be concrete, we can derive the following corollary from Theorem 4.

**Corollary 1** Let us take arbitrary states $\rho$ on $A$ and $\rho_B$ on $B$ and an arbitrary non-negative integer $s$. We define the following $\rho_B^{M_s}$ from $\rho_B$ and refer to the distance between $\rho_B^{M_s}$ and $\rho_B$ as $\epsilon_{B,M_s}$:

$$
\rho_B^{M_s} := \frac{P^{M_s} \rho_B P^{M_s}}{\text{Tr}[P^{M_s} \rho_B]},
$$

$$
\epsilon_{B,M_s} := \left\| \rho_B^{M_s} - \rho_B \right\|_1,
$$

where $P^{M_s}$ is the projection to the subspace $\mathcal{H}^{M_s}$. Then, the following relation holds:

$$
\text{Prob}_{U \sim \mathcal{H}^{M_s}_{\text{Haar}}} \left[ x_A'(\rho, \rho_B, U) \approx_{t+(2N+k)\epsilon_{B,M_s}} (x_A(\rho) + x_B(\rho_B)) \times (1 - \gamma) \right] \geq 1 - 2 \exp \left( \frac{-(N+kC_s - 2)^2}{48(N+k)^2} \right)
$$

(III.73)

**Proof of Corollary 1.** For an arbitrary unitary $U$ on $AB$ satisfying the conservation law, we obtain

$$
|x_A'(\rho, \rho_B, U) - x_A'(\rho, \rho_B^{M_s}, U)| = |\text{Tr}[X_A(U \rho \otimes \rho_B U^\dagger - U \rho \otimes \rho_B^{M_s} U^\dagger)]| \\
\leq \|X_A\|_{\infty} \epsilon_{B,M_s} \\
\leq (N + k) \times \epsilon_{B,M_s}.
$$

(III.74)

Therefore, $x_A'(\rho, \rho_B, U) \approx_{(N+k)\epsilon_{B,M_s}} x_A'(\rho, \rho_B^{M_s}, U)$ always holds. We also obtain

$$
|x_B(\rho_B) - x_B(\rho_B^{M_s})| = |\text{Tr}[X_B(\rho_B - \rho_B^{M_s})]| \\
\leq \|X_B\|_{\infty} \epsilon_{B,M_s} \\
\leq N \times \epsilon_{B,M_s}.
$$

(III.75)

Since the support of $\rho \otimes \rho_B^{M_s}$ is on $\mathcal{H}^{M_s}$, Theorem 4 and (III.62) imply

$$
\text{Prob}_{U \sim \mathcal{H}^{M_s}_{\text{Haar}}} \left[ x_A'(\rho, \rho_B^{M_s}, U) \approx_{t} (x_A(\rho) + x_B(\rho_B^{M_s})) \times (1 - \gamma) \right] \geq 1 - 2 \exp \left( \frac{-(N+kC_s - 2)^2}{48(N+k)^2} \right).
$$

(III.76)

Here we used $t^2 \leq (N + k)^2$. Combining (III.74)–(III.76), we obtain (III.73). □

**Derivation of Propositions 1 and 2:** As shown below, Theorem 4 and Corollary 1 imply Propositions 1 and 2, i.e., (III.49) and (III.50). In other words, under natural settings, for the Haar random unitary with the conservation law of $X$, the assumption (III.5) used in Theorem 1 holds with very high probability. First, we derive (III.49). We set the initial state $\rho_B$ on $B$ is the maximally mixed state and set the parameters $N$, $k$ and $\epsilon$ satisfying $N \geq 10^3$, $k \leq 10N$ and $1/(N+k)^2 \geq \epsilon \geq 1/(N+k)^3$. We also take an arbitrary state $\rho$ on $A$ satisfying $M \geq 1/(N+k)$. To derive (III.49) from (III.73), let us set $s = 23$ and $t = \frac{M\gamma}{2}$, and evaluate $\frac{(N+kC_s - 2)^2}{48(N+k)^2}$ and $\epsilon_{B,M_s}$. We remark that to obtain (III.49), we only have to derive the following inequalities:

$$
\epsilon_{B,M_s} \leq \frac{M\gamma}{8(N+k)},
$$

(III.77)

$$
\frac{(N+kC_{23} - 2)^2}{48(N+k)^2} - \log 2 \geq (N+k)^2.
$$

(III.78)

Let us derive (III.77). Due to the definition of $\epsilon_{B,M_s}$, $\epsilon_{B,M_s} = 2\sum_{m=0}^{22} \frac{N C_m}{2N - 1}$. Because of $M \geq \frac{1}{N+k}$, $\gamma \geq \frac{1}{N+k}$, $\epsilon \geq \frac{1}{(N+k)^2}$ and $k \leq 10N$, we obtain $\frac{M\gamma}{8(N+k)} \leq \frac{1}{8 \times 1.1^6 \times N^5}$. Therefore, to obtain (III.77), it is sufficient to show

$$
\sum_{m=0}^{22} \frac{N C_m}{2N - 1} \leq \frac{1}{8 \times 1.1^6 \times N^5}.
$$

(III.79)
for \( N \geq 10^3 \). Because \( 8 \times 11^6 \times N^6 \times 4^{22} = 8^{\frac{N+1}{2(N+k)}} \) is monotonically decreasing with \( N \) for \( N \geq 10^3 \), and because \((8 \times 11^6 \times N^6 \times 4^{22})|_{N=10^3} \) is approximately equal to \( 5 \times 10^{-230} \), we obtain (III.79), and thus the inequality (III.77) holds.

Next, we derive (III.78). Because of \( M \geq \frac{1}{N+k} \), \( \gamma \geq \frac{1}{N+k} \), \( \epsilon \geq \frac{1}{(N+k)^2} \), the inequality \( \frac{(N+kC_s-2)^2}{48(N+k)^2} \geq \frac{N+kC_s-2}{48\times16\times(N+k)^2} \) holds. Therefore, to show (III.78), it is sufficient to show

\[
1 \leq \frac{N+kC_s=23 - 2}{48 \times 16 \times (N + k)^2} - \frac{\log 2}{(N + k)^2}. 
\]

(III.80)

Because the right-hand side of the above is monotonically increasing with \( N + k \) for \( N + k \geq 1001 \), and because of \( \left(\frac{N + kC_s = 23 - 2}{48 \times 16 \times (N + k)^2} - \frac{\log 2}{(N + k)^2}\right)|_{N+k=1001} \approx 39 \), we obtain (III.80) and thus (III.78) holds. Therefore, we have shown (III.49).

Next, we show (III.50). Let us set the parameters \( N, k \) and \( \epsilon \) satisfying \( N \geq 10^3 \), \( 47 < k \leq 10N \) and \( 1/(N+k)^3 \leq \epsilon \leq 1/(N+k)^2 \). We also take an arbitrary state \( \rho_B \) on \( B \) and an arbitrary state \( \rho_A \) such that \( M \geq \frac{1}{N+k} \) and the support of \( \rho_A \) is included in \( \mathcal{H}_A^{23 \leq \sigma \leq k - 23} \). To show that (III.50) holds under this condition, we remark that when the support of \( \rho_A \) is included in \( \mathcal{H}_A^{23 \leq \sigma \leq k - 23} \), the support of \( \rho_A \otimes \rho_B \) is included in \( \mathcal{H}_{M=23} \). Therefore, to obtain (III.50) from Theorem 4, it is sufficient to set \( t = M\gamma\epsilon/4 \) and \( s = 23 \) and show the inequality

\[
\frac{(N+kC_s=23 - 2)t^2}{48^2} - \log 2 \geq (N + k)^2.
\]

(III.81)

Due to \( t \leq N + k \), this inequality is easily obtained from (III.78). Therefore, we obtain (III.50). □

Now, let us show Theorem 4. To show it, we introduce two definitions and a theorem.

**Definition 1** Let \( f \) be a real-valued function on a metric space \((X,d)\). When \( f \) satisfies the following relation for a real positive constant \( L \), then \( f \) is called \( L \)-Lipschitz:

\[
L = \sup_{x \neq y \in X} \frac{|f(x) - f(y)|}{d(x,y)}.
\]

(III.82)

**Definition 2** Let \( \mathcal{U}_X^M \) be a product of unitary groups \( \times_{i=1}^M \mathcal{U}(d_i) \), where each \( \mathcal{U}(d_i) \) is the unitary group of all unitary operations on a \( d_i \)-dimensional Hilbert space. For \( U = \oplus_{i=1}^M U_i \in \mathcal{U}_X^M \) and \( V = \oplus_{i=1}^M V_i \in \mathcal{U}_X^M \), the \( L_2 \)-sum \( D(U,V) \) of the Hilbert-Schmidt norms on \( \mathcal{U}_X^M \) is defined as:

\[
D(U,V) := \sqrt{\sum_{i=1}^M \|U_i - V_i\|_2^2}.
\]

(III.83)

where \( \|\cdot\|_2 := \sqrt{\text{Tr}[(\cdot)(\cdot)^\dagger]} \).

**Theorem 5** (Corollary 3.15 in Ref. [1]) Let \( \mathcal{U}_X^M \) be a product of unitary groups \( \times_{i=1}^M \mathcal{U}(d_i) \), where each \( \mathcal{U}(d_i) \) is a unitary group of all unitary operations on a \( d_i \)-dimensional Hilbert space. Let \( \mathcal{U}_X^M \) be equipped with the \( L_2 \)-sum of Hilbert-Schmidt norms, and \( H_X := \times_{i=1}^M H_i \) where each \( H_i \) is the Haar measure on \( \mathcal{U}(d_i) \). Suppose that a real-valued function \( f \) on \( \mathcal{U}_X^M \) is \( L \)-Lipschitz. Then, for arbitrary \( t > 0 \),

\[
\text{Prob}[f(U) \geq f(U') + t] \leq \exp \left( -\frac{(d_{\text{min}} - 2)t^2}{12L^2} \right),
\]

(III.84)

where \( d_{\text{min}} := \min\{d_1, \ldots, d_M\} \).

From Theorem 5, we can easily derive Theorem 4.

**Proof of Theorem 4** Since the support of \( \rho \otimes \rho_B \) is included in the subspace \( \mathcal{H}_X^{M_{\text{ref}}} := \otimes_{m \in M} \mathcal{H}^{(m)} \), the following relation holds for arbitrary \( U \in \mathcal{U}_X^{M_{\text{ref}}} \):

\[
x_A'/(\rho,\rho_B, U) = x_A'/(\rho,\rho_B, U).
\]

(III.85)
where \( \tilde{U} \) defined from \( U \) by \([III.69]\). Therefore, we only have to show

\[
\text{Prob}_{\hat{U}_{\sim H_{\mathcal{M}^*}}}[|x_{A'}(\rho, \rho_B, \tilde{U}) - x_{A'}(\rho, \rho_B, U)|_{H_{\mathcal{M}^*}} > t] \leq 2 \exp \left( -\frac{(N + k C_s - 2)t^2}{48t^2} \right). \tag{III.86}
\]

Note that \( \min_{m \in \mathcal{M}^*} \dim \mathcal{H}^{(m)} = N + k C_s \). Therefore, due to Theorem 5 to show \([III.86]\), it is sufficient to show that \( x_{A'}(\rho, \rho_B, \tilde{U}) \) is 2\(t\)-Lipschitz.

To show that \( x_{A'}(\rho, \rho_B, \tilde{U}) \) is 2\(t\)-Lipschitz, let us take two unitary operations \( \hat{U} \in \mathcal{U}_{\mathcal{M}^*}^\dagger \) and \( \hat{V} \in \mathcal{U}_{\mathcal{M}^*}^\dagger \). We evaluate \( |x_{A'}(\rho, \rho_B, \tilde{U}) - x_{A'}(\rho, \rho_B, \hat{V})| \) as follows:

\[
|x_{A'}(\rho, \rho_B, \tilde{U}) - x_{A'}(\rho, \rho_B, \hat{V})| = |\text{Tr}[X_{A'}(\hat{U}(\rho \otimes \rho_B)\hat{U}^\dagger - \hat{V}(\rho \otimes \rho_B)\hat{V}^\dagger)]|
\leq D_{X_{A'}}||\hat{U}(\rho \otimes \rho_B)\hat{U}^\dagger - \hat{V}(\rho \otimes \rho_B)\hat{V}^\dagger||_1
\leq t||\hat{U}(\rho \otimes \rho_B)\hat{U}^\dagger - \hat{V}(\rho \otimes \rho_B)\hat{V}^\dagger||_1. \tag{III.87}
\]

Therefore, in order to show that \( x_{A'}(\rho, \rho_B, \tilde{U}) \) is 2\(t\)-Lipschitz, we only have to show

\[
||\hat{U}(\rho \otimes \rho_B)\hat{U}^\dagger - \hat{V}(\rho \otimes \rho_B)\hat{V}^\dagger||_1 \leq 2D(\hat{U}, \hat{V}). \tag{III.88}
\]

To show \([III.88]\), we take a purification of \( \rho \otimes \rho_B \), and refer to it as \( |\psi_{ABQ}\rangle \). Due to the monotonicity of the 1 norm and \( ||\phi - \psi||_1 = 2D_F(\phi, \psi) \) for any pure \( \phi \) and \( \psi \),

\[
||\hat{U}(\rho \otimes \rho_B)\hat{U}^\dagger - \hat{V}(\rho \otimes \rho_B)\hat{V}^\dagger||_1 \leq ||\hat{U}\psi_{ABQ}\hat{U}^\dagger - \hat{V}\psi_{ABQ}\hat{V}^\dagger||_1
\]

\[
= 2\sqrt{1 - F^2((\hat{U} \otimes 1_Q)\psi_{ABQ}(\hat{U} \otimes 1_Q)\dagger, (\hat{V} \otimes 1_Q)\psi_{ABQ}(\hat{V} \otimes 1_Q)\dagger)}
\leq 2\sqrt{2(1 - F((\hat{U} \otimes 1_Q)\psi_{ABQ}(\hat{U} \otimes 1_Q)\dagger, (\hat{V} \otimes 1_Q)\psi_{ABQ}(\hat{V} \otimes 1_Q)\dagger))}
\leq 2\sqrt{2 - (\psi_{ABQ}\dagger|\hat{U} \otimes 1_Q\rangle (\hat{V} \otimes 1_Q)\psi_{ABQ})}
\leq 2\sqrt{2 - (\psi_{ABQ}\dagger|\hat{U} \otimes 1_Q\rangle (\hat{V} \otimes 1_Q)\psi_{ABQ})}
\leq 2\sqrt{||(\hat{U} \otimes 1_Q) - (\hat{V} \otimes 1_Q)||\psi_{ABQ})||_2
\leq 2||\hat{U} - \hat{V}||_2||\rho \otimes \rho_B||_\infty^{1/2}. \tag{III.89}
\]

In the final line, we use

\[
||(\hat{U} \otimes 1_Q) - (\hat{V} \otimes 1_Q)||\psi_{ABQ})||_2^2 = \text{Tr}[(\hat{U} \otimes 1_Q) - (\hat{V} \otimes 1_Q)||\psi_{ABQ})\rangle (\psi_{ABQ})\langle(\hat{U} \otimes 1_Q) - (\hat{V} \otimes 1_Q)\rangle]
= \text{Tr}[(\hat{U} - \hat{V})(\rho \otimes \rho_B)(\hat{U} - \hat{V})\dagger]
\leq ||\rho \otimes \rho_B||_\infty||\hat{U} - \hat{V}||_1
\leq ||\rho \otimes \rho_B||_\infty||\hat{U} - \hat{V}||^2_2. \tag{III.90}
\]

where we use the Hölder inequality in the final line. Due to \( ||M_1 \oplus M_2 - M'_1 \oplus M'_2||_2^2 = ||M_1 - M'_1||_2^2 + ||M_2 - M'_2||_2^2 \) and the definition of \( L_2 \)-sum, we can show \( ||\hat{U} - \hat{V}||_2 = D(\hat{U}, \hat{V}) \) as follows:

\[
||\hat{U} - \hat{V}||_2^2 = \sum_{m \in \mathcal{M}^*} ||\hat{U}^{(m)} - \hat{V}^{(m)}||^2_2
= D(\hat{U}, \hat{V}^\dagger), \tag{III.91}
\]

where \( \hat{U}^{(m)} \) and \( \hat{V}^{(m)} \) are defined as \( \hat{U} = (\oplus_{m \in \mathcal{M}^*} \hat{U}^{(m)}) \oplus (\oplus_{m \notin \mathcal{M}^*} I^{(m)}) \) and \( \hat{V} = (\oplus_{m \in \mathcal{M}^*} \hat{V}^{(m)}) \oplus (\oplus_{m \notin \mathcal{M}^*} I^{(m)}) \).

Combining \([III.89], [III.91]\) and \( ||\rho \otimes \rho_B||_\infty \leq 1 \), we obtain \([III.88]\).

Next, we prove that the assumption \([III.20]\) (= [17]) in the main text) is valid for a typical Haar random unitary with the energy conservation. To do so, we use the following theorems, which correspond to the second-order variation of Theorem 3 and Theorem 4.
Theorem 6 For arbitrary eigenstates $|i,a \rangle$ and $|j,b \rangle$ of $X_A$ and $X_B$ whose eigenvalues are $i$ and $j$, the following inequality holds:

$$V_{\rho^\alpha_{i,a,j,b,U}}(X_{\alpha'}) \leq \min\{(1-\gamma)(N+k), \gamma(N+k)\} \cdot 4^{-1} \quad (\alpha' = A', \ B'),$$

(III.92)

where $f(U)$ is the average of the function $f$ with the product Haar measure $H^\infty_M$. Additionally, when $|i,a \rangle \otimes |j,b \rangle$ is included in the subspace $H^M := \otimes_{m \in M} H^{(m)}$, the following equality holds:

$$V_{\rho^\alpha_{i,a,j,b,U}}(X_{\alpha'}) \leq \min\{(1-\gamma)(N+k), \gamma(N+k)\} \cdot 4^{-1} \quad (\alpha' = A', \ B'),$$

(III.93)

where $U$ is a unitary which is described as $U = (\oplus_{m \in M} U^{(m)}) \oplus (\oplus_{n \in N} I^{(m)})$ where $U^{(m)} \in U^{(m)}$, and $f(U)$ is the average of the function $f$ with the product Haar measure $H^\infty_M$.

Theorem 7 For an arbitrary positive number $t$, and arbitrary states $\rho$ and $\rho_B$ on $A$ and $B$ which satisfy that the support of $\rho \otimes \rho_B$ is included in the subspace $H^M := \otimes_{m \in M} H^{(m)}$ for $M := \{s, s+1, \ldots, N+k-s\}$, the following relation holds:

$$\text{Prob}_{U \sim H^\infty_M} \left[ |x_A^{(2)}(\rho, \rho_B, U) - x_B^{(2)}(\rho, \rho_B, U)| > t \right] \leq 2 \exp(-\frac{(N+k-s^2)}{48t^2}),$$

(III.94)

$$\text{Prob}_{U \sim H^\infty_M} \left[ |x_A^{(2)}(\rho, \rho_B, U) - x_B^{(2)}(\rho, \rho_B, U)| > t \right] \leq 2 \exp(-\frac{(N+k-s^2)}{48t^2}),$$

(III.95)

where $x_A^{(2)}(\rho_B, U) := \text{Tr}_{X_B} \rho_B U^\dagger U$, and $\text{Prob}_{U \sim H^\infty_M}[\cdots]$ is the probability that the event $\cdots$ happens when $U$ is drawn from $U^M$ with the measure $H^\infty_M$.

Theorems 3, 6, and 7 guarantee that when $U$ is a typical Haar random unitary with the conservation law of $X$, the assumption (III.20) actually holds with very high probability, in the same way as Theorems 4, 5, and 7 guarantee that when $U$ is a typical Haar random unitary with the conservation law of $X$, the assumption (III.5) holds with very high probability. To be concrete, we can derive the following corollary:

Corollary 2 Let $|i,a \rangle$ and $|j,b \rangle$ be eigenstates of $X_A$ and $X_B$ whose eigenvalues are $i$ and $j$. When $N \geq 10^3$ and $18 \leq i + j \leq N + k$ hold, the following relation holds for an arbitrary real number $\epsilon$ satisfying $1/(N+k)^2 \leq \epsilon \leq 1/(N+k)^2$:

$$\text{Prob}_{U \sim H^\infty_M} \left[ V_{\rho^\alpha_{i,a,j,b,U}}(X_{\alpha'}) \leq (1 + \epsilon) \min\{(1-\gamma)(N+k), \gamma(N+k)\} \cdot 4^{-1} \right] \geq 1 - 4 \exp(-10(N+k)) \quad (\alpha' = A', \ B').$$

(III.96)

Derivation of Proposition 3: Now, let us derive Proposition 3 from Theorem 4 and Corollary 2. First, Corollary 2 directly implies (III.53). Let us derive (III.54) from Theorem 4. We take the parameters $N$, $k$ and $\epsilon$ satisfying $N \geq 10^3$ and $1/(N+k)^2 \leq \epsilon \leq 1/(N+k)^2$. We also set a number $s$ as $s = 18$. To derive (III.54), we only have to set $s$ and $t$ in (III.70) as $s = 18$ and $t = \epsilon$ and show

$$(N+k)^2 \leq \frac{(N+k-s^2)}{48t^2} - \log 2.$$  

(III.97)

Due to $\epsilon \geq 1/(N+k)^2$ and $l \leq N+k$, to show (III.97), we only have to show

$$1 \leq \frac{(N+k-s^2)}{48(N+k)^2} - \frac{\log 2}{(N+k)^2}.$$  

(III.98)

For $N+k \geq 1001$, the right-hand side of (III.98) is monotonically increasing with $N+k$. Due to $\frac{(N+k-s^2)}{48(N+k)^2} \approx 2.8 \times 10^6$, (III.98) holds, and thus (III.54) also holds.
Proof of Corollary 2: Due to $18 \leq i + j \leq N + k - 18$, the state $|i, a\rangle \otimes |j, b\rangle$ is in $H^{M_2=18}$. Therefore, from Theorem 4 and Theorem 7, we obtain the following relations for arbitrary $t_1$ and $t_2$:

\[
\begin{align*}
\text{Prob}_{U \sim H^{M_2=18}} \left[ x_{\alpha'}^2(|i, a\rangle, |j, b\rangle, U) - x_{\alpha'}^2(i, a, |j, b\rangle, U) \right] &> t_2 \right] \leq 2 \exp \left( -\frac{(N+k)C_{18} - 2)\ell_2^2}{48(N+k)^4} \right), \\
\text{Prob}_{\hat{U} \sim H^{M_2=18}} \left[ x_{\alpha'}^2(|i, a\rangle, |j, b\rangle, U) - x_{\alpha'}^2(i, a, |j, b\rangle, U) \right] &> t_1 \right] \leq 2 \exp \left( -\frac{(N+k)C_{18} - 2)\ell_2^2}{48(N+k)^2} \right). 
\end{align*}
\]

Therefore,

\[
\begin{align*}
\text{Prob}_{\hat{U} \sim H^{M_2=18}} \left[ x_{\alpha'}^2(|i, a\rangle, |j, b\rangle, U) \approx t_2 \ x_{\alpha'}^2(i, a, |j, b\rangle, U) \right] \wedge \left( x_{\alpha'}^2(|i, a\rangle, |j, b\rangle, U) \approx t_1 \ x_{\alpha'}(|i, a\rangle, |j, b\rangle, U) \right)
\end{align*}
\]

Due to $f^{\alpha'}_{\lambda}(i, a, j, b, U) = x_{\alpha'}^2(|i, a\rangle, |j, b\rangle, U) - x_{\alpha'}(|i, a\rangle, |j, b\rangle, U)^2$ and $x_{\alpha'}(|i, a\rangle, |j, b\rangle, U) \leq N + k$, we obtain

\[
\begin{align*}
\left( x_{\alpha'}^2(|i, a\rangle, |j, b\rangle, U) \approx t_2 \ x_{\alpha'}^2(i, a, |j, b\rangle, U) \right) \wedge \left( x_{\alpha'}(|i, a\rangle, |j, b\rangle, U) \approx t_1 \ x_{\alpha'}(|i, a\rangle, |j, b\rangle, U) \right)
\end{align*}
\]

\[
\begin{align*}
\Rightarrow \left( V^{\alpha'}_{\lambda}(i, a, j, b, U) \approx t_2 \ t_1 \ N + k \right) \left( V^{\alpha'}_{\lambda}(i, a, j, b, U) \right)
\end{align*}
\]

Therefore, we obtain

\[
\begin{align*}
\text{Prob}_{\hat{U} \sim H^{M_2=18}} \left[ V^{\alpha'}_{\lambda}(i, a, j, b, U) \approx t_2 \ t_1 \ N + k \right) \left( V^{\alpha'}_{\lambda}(i, a, j, b, U) \right)
\end{align*}
\]

Here, let us define $t_1 := \frac{\epsilon}{16(N+k)}$ and $t_2 := \frac{\epsilon}{10}$. Then, with using $1/(N+k)^2 \geq \epsilon \geq 1/(N+k)^3$, we obtain

\[
\begin{align*}
\text{Prob}_{\hat{U} \sim H^{M_2=18}} \left[ V^{\alpha'}_{\lambda}(i, a, j, b, U) \approx \epsilon/4 \right. \left. \left( V^{\alpha'}_{\lambda}(i, a, j, b, U) \right) \right] \geq 1 - 4 \exp \left( -\frac{N+kC_{18} - 2}{48 \times 16^2(N+k)^10} \right).
\end{align*}
\]

Due to $N \geq 10^3$, we obtain the following:

\[
\begin{align*}
\frac{N+kC_{18} - 2}{48 \times 16^2(N+k)^10} \geq \frac{10^3C_{18} - 2}{748 \times 16^2(10)^{33}} \geq 10.
\end{align*}
\]

Therefore, we obtain

\[
\begin{align*}
\frac{N+kC_{18} - 2}{48 \times 16^2(N+k)^10} \geq \frac{10^3C_{18} - 2}{48 \times 16^2(10)^{33}} \geq 10(N + k).
\end{align*}
\]

Therefore, we obtain

\[
\begin{align*}
\text{Prob}_{\hat{U} \sim H^{M_2=18}} \left[ V^{\alpha'}_{\lambda}(i, a, j, b, U) \approx \epsilon/4 \right. \left. \left( V^{\alpha'}_{\lambda}(i, a, j, b, U) \right) \right] \geq 1 - 4 \exp \left( -10(N + k) \right).
\end{align*}
\]

Combining the above inequality, (III.92) and $\min\{(1-\gamma)(N+k), \gamma(N+k)\} \geq 1/4$, we obtain (III.96).

Proof of Theorem 6: We firstly remark that

\[
\hat{U}|i, a\rangle \otimes |j, b\rangle \langle j, b| \hat{U}^\dagger = \hat{U}_{i+j}^{\text{max}}
\]

(III.108)
where $\rho^{\max}_{i+j}$ is the maximally mixed state in eigenspace of $X_{A'} + X_{B'}$ whose eigenvalue is $i + j$. Therefore,

$$p'_{h,h'|i,a,j,b,U} := \text{Tr}_{-\langle h,h'\rangle} \langle U | i, a \rangle \langle j, b | U \rangle$$

$$= \frac{(i + j)(i + j - 1)}{(N + k)(N + k - 1)} |1\rangle \langle 1|_{h} \otimes |1\rangle \langle 1|_{h'} + \frac{(i + j)(N + k - (i + j))}{(N + k)(N + k - 1)} (|1\rangle \langle 0|_{h} \otimes |0\rangle_{h'} + |0\rangle \langle 0|_{h} \otimes |1\rangle_{h'}) + \frac{(N + k - (i + j))(N + k - (i + j - 1))}{(N + k)(N + k - 1)} |0\rangle \langle 0|_{h} \otimes |0\rangle_{h'}$$

(III.109)

where $\text{Tr}_{-\langle h,h'\rangle}$ is the partial trace of the qubits other than the $h$-th and $h'$-th qubits, and $|1\rangle_{h}$ and $|0\rangle_{h}$ are the excited state and the ground state of $X_{h}$.

Here we define $x_{A'}^{(2)}(\rho, \rho B, U) := \text{Tr}_{-\langle A'\rangle} \langle X_{A'}^{2} U \rho \otimes \rho B U \rangle ((\alpha', -\alpha') = (A', B') or (B', A'))$, and obtain

$$\frac{x_{A'}^{(2)}(|i, a\rangle \langle j, b|, |j, b\rangle \langle j, b|, U)}{\sum_{h,h' \in A', h' \neq h'} \langle X_{h} \otimes X_{h'} \rangle p'_{h,h'|h,a,j,b,U} + \sum_{h \in A'} \langle X_{h} \rangle p'_{h,h'|h,a,j,b,U}}$$

$$= \frac{(i + j)(i + j - 1)}{(N + k)(N + k - 1)} l(l - 1) + \frac{i + j}{N + k} l$$

$$= p \left( p + \frac{p - 1}{N + k - 1} \right) l(l - 1) + pl$$

(III.110)

Here we use $p := (i + j)/(N + k)$. Due to Theorem 3, we obtain $x_{A'}(|i, a\rangle \langle j, b|, |j, b\rangle \langle j, b|, U) = (1 - \gamma)(i + j) = lp$. Therefore, we obtain

$$V_{p_{A'}^{\rho}}(X_{A'}) = p \left( p + \frac{p - 1}{N + k - 1} \right) l(l - 1) + pl - p^{2} l^{2}$$

$$= p(1 - p) l - \frac{p(1 - p)}{N + k - 1} l(l - 1)$$

$$= p(1 - p) l \left( 1 - \frac{l - 1}{N + k - 1} \right).$$

(III.111)

Since $0 \leq p \leq 1$ holds, the inequality $p(1 - p) \leq 1/4$ holds. Therefore, to obtain (III.92), we only have to show

$$l \left( 1 - \frac{l - 1}{N + k - 1} \right) \leq \min\{l, N + k - l\}.$$

(III.112)

Due to $\left( 1 - \frac{l - 1}{N + k - 1} \right) \leq 1$, we only have to show that the following inequality holds for $l \in \{1, 2, \ldots, N + k\}$:

$$l \left( 1 - \frac{l - 1}{N + k - 1} \right) \leq N + k - l.$$

(III.113)

To show (III.113), we show that the following inequality holds for $l \in \{1, 2, \ldots, N + k\}$:

$$l \left( 2 - \frac{l - 1}{N + k - 1} \right) \leq N + k.$$

(III.114)

To show this, let us define $\gamma(l) := l \left( 2 - \frac{l - 1}{N + k - 1} \right)$. Then,

$$\frac{d\gamma(l)}{dl} = \left( 2 - \frac{l - 1}{N + k - 1} \right) - \frac{l}{N + k - 1}$$

$$= 2 - \frac{2l - 1}{N + k - 1}.$$

(III.115)

Therefore, for $0 \leq l \leq N + k - 1$, the function $\gamma(l)$ is monotonically increasing with $l$. Furthermore, $\gamma(N + k - 1) = h(N + k) = N + k$ holds. Therefore, (III.114) holds for $l \in \{1, 2, \ldots, N + k\}$. Hence, we obtain (III.92).

We can also derive (III.93) in the same way. We only have to substitute $\hat{U}$ and $\hat{U}_{s}^{M_{l}}$ for $U$ and $\hat{V} \in \hat{U}_{s}^{M_{l}}$. ■
The proof of Theorem 7 is almost the same as that of Theorem 4.

**Proof of Theorem 7.** Since the support of $\rho \otimes \rho_B$ is included in the subspace $H^M := \otimes_{m \in M_A} H^{(m)}$, the following relation holds for arbitrary $U \in U^M_x$:

$$x_{A'}(\rho, \rho_B, U) = x_{A'}(\rho, \rho_B, \tilde{U}),$$

where $\tilde{U}$ defined from $U$ by $\tilde{V} = (\oplus_{m \in M_B} V(m)) \oplus (\oplus_{n \in M_A} I^{(m)})$. Therefore, we only have to show

$$\text{Prob}_{\tilde{U} \sim H^M_x} \left[ |x_{A'}^{(2)}(\rho, \rho_B, \tilde{U}) - x_{A'}^{(2)}(\rho, \rho_B, \tilde{U})|_{H^M_x} > t \right] \leq 2 \exp \left( -\frac{(N+k)C_s - 2t^2}{48t^4} \right).$$

(III.117)

Note that $\min_{m \in M_A}, \dim H^{(m)} = \frac{N+k}{2}$. Therefore, due to Theorem 5, to show (III.117), it is sufficient to show that $x_{A'}^{(2)}(\rho, \rho_B, \tilde{U})$ is 2$^{l_2}$-Lipschitz.

To show that $x_{A'}^{(2)}(\rho, \rho_B, \tilde{U})$ is 2$^{l_2}$-Lipschitz, let us take two unitary operations $\tilde{U} \in U^M_x$ and $\tilde{V} \in U^M_x$. We evaluate $|x_{A'}(\rho, \rho_B, \tilde{U}) - x_{A'}(\rho, \rho_B, \tilde{V})|$ as follows:

$$|x_{A'}^{(2)}(\rho, \rho_B, \tilde{U}) - x_{A'}^{(2)}(\rho, \rho_B, \tilde{V})| = |\text{Tr}[X_{A'}^{(2)}(\tilde{U} \otimes \rho_B) \tilde{U}^\dagger - \tilde{V}(\rho \otimes \rho_B) \tilde{V}^\dagger]|$$

$$\leq D_{X_{A'}}(\rho \otimes \rho_B) \tilde{U}^\dagger - \tilde{V}(\rho \otimes \rho_B) \tilde{V}^\dagger\|_1$$

$$\leq \tilde{l} \|\tilde{U}(\rho \otimes \rho_B)\|_1 - \tilde{V}(\rho \otimes \rho_B)\|_1.$$ (III.118)

Therefore, in order to show that $x_{A'}^{(2)}(\rho, \rho_B, \tilde{U})$ is 2$^{l_2}$-Lipschitz, we only have to show

$$\|\tilde{U}(\rho \otimes \rho_B)\|_1 - \tilde{V}(\rho \otimes \rho_B)\|_1 \leq 2D(\tilde{U}, \tilde{V}).$$

(III.119)

And we have already obtained (III.119) as (III.88). Therefore, $x_{A'}^{(2)}(\rho, \rho_B, \tilde{U})$ is 2$^{l_2}$-Lipschitz, and we have obtained (III.95) by substituting $N + k - l$ for $l$ in the above argument. \[\blacksquare\]

5. Other applications to Hayden-Preskill model with symmetry

Other than [21], there are several applications to Hayden-Preskill model. For example, we can use [5] for non-maximally entangled states for the initial states $AR_A$ and $BR_B$. Noting $\Delta_+ \leq (k + l)/2$, we obtain the following bound

$$\frac{1 - \epsilon}{1 + \epsilon} \times \frac{M(1 - l/(N + k))}{2(\sqrt{\mathcal{F}} + 2(k + l))} \leq \delta.$$ (III.120)

To illustrate the meaning of this inequality, we consider the case of $M \propto k$. Then, we obtain the lower bound (III.120):

$$\text{const.} \times \frac{1 - l/(k + N)}{1 + (2l + \sqrt{\mathcal{F}})/(2k)} \leq \delta.$$ (III.121)

Note that $\mathcal{F} = 4V_{RB}(X_B)$ where $\rho_B := \text{Tr}_{RB}[\rho_{BRB}]$. This inequality shows that when the fluctuation of the conserved quantity of the initial state of the black hole $\tilde{B}$ is not so large, in order to make $\delta$ small, we have to collect information from the Hawking radiation so that $l \gg k$ or $l \approx k + N$. In other words, whenever the fluctuation of the conserved quantity of the black hole is small, then in order to recover the quantum data thrown into the black hole with good accuracy, we have to wait until the black hole is evaporated enough. Note also that if $\sqrt{\mathcal{F}}$ is small, the bound in (III.121) does not become trivial even if $N$ is much larger than $k$.

Appendix Supp.IV: Lower bound of recovery error in the information recovery without using $R_B$

The relations [5] and [6] in the main text describe the limitation of information recovery when one uses the quantum information of $R_B$. We can also discuss the case without using the information of $R_B$. The recovery operation $\mathcal{R}$ in
FIG. S.3. Schematic diagram of the information recovery without using \( R_B \). This case maps the state on the system \( A' \) to \( A \), as seen in the schematic in Fig. S.3. We then define the recovery error as
\[
\tilde{\delta} := \min_{\mathcal{E}} D_F(\rho_{AR_A}, \text{id}_{R_A} \otimes \mathcal{R} \circ \mathcal{E}(\rho_{AR_A})) .
\] (IV.1)

Since \( \tilde{\delta} \geq \delta \), we can substitute \( \tilde{\delta} \) for \( \delta \) in (5) and (6) to get a limitation of recovery in the present setup. Moreover, we can derive a tighter relation than this simple substitution as
\[
\frac{A}{2(\sqrt{F_B} + 4 \Delta_+)} \leq \tilde{\delta} ,
\] (IV.2)

where \( F_B := F_{\rho_B}(X_B) \). Note that \( F_B \leq F \) holds in general. The inequality (IV.2) is the third main relation on the information recovery.

**Proof of (IV.2):** We firstly take a quantum system \( \tilde{B} \) whose dimension is the same as \( B \), and a purification \( |\phi_{BB} \rangle \) of \( \rho_B := \text{Tr}_{R_B}[|\phi_{BB} \rangle \rangle] \). From \( |\phi_{BB} \rangle \) and \( U \), we define a set \( \tilde{\mathcal{I}} := (|\phi_{BB} \rangle \otimes |\eta_{R_B} \rangle, U \otimes 1_B) \). We take the Schmidt decomposition of \( |\phi_{BB} \rangle \) as
\[
|\phi_{BB} \rangle = \sum_i \sqrt{r_i} |l_B \rangle |l'_{\tilde{B}} \rangle ,
\] (IV.3)

and define \( X_{\tilde{B}} \) on \( \tilde{B} \) as
\[
X_{\tilde{B}} := \sum_{i'} 2\sqrt{r_{ii'}} (|l_B \rangle \langle l'_{B'}| + |l'_{B'} \rangle \langle l_B |) .
\] (IV.4)

Then, due to (A4) and (A5),
\[
F_{\rho_B}(X_B) = 4V_{|\phi_{BB} \rangle \otimes |\eta_{R_B} \rangle} (X_B + X_{\tilde{B}}) .
\] (IV.5)

Note that \( \tilde{\mathcal{I}} \) is a Steinspring representation of \( \mathcal{E} \) and that \( U \otimes 1_B(X_A + X_B + X_{\tilde{B}})(U \otimes 1_B)^\dagger = X_{A'} + X_{B'} + X_{\tilde{B}} \). Therefore, we obtain the following inequality from (5):
\[
\frac{A}{2(\sqrt{F_{|\phi_{BB} \rangle \otimes |\eta_{R_B} \rangle}} (X_B + X_{\tilde{B}}) \otimes 1_{R_B}) + 4 \Delta_+} \leq \delta(|\psi_{AR_A} \rangle , \tilde{\mathcal{I}}) \] (IV.6)

Since \( |\phi_{BB} \rangle \otimes |\eta_{R_B} \rangle \) is a tensor product between \( B\tilde{B} \) and \( R_B \), the state of \( B\tilde{B}R_B \) after \( U \) is also another tensor product state between \( B\tilde{B} \) and \( R_B \). Therefore, we obtain
\[
\delta(|\psi_{AR_A} \rangle , \tilde{\mathcal{I}}) = \tilde{\delta} \] (IV.7)

Finally, from (IV.5), we obtain
\[
F_{\rho_B}(X_B) = 4V_{|\phi_{BB} \rangle \otimes |\eta_{R_B} \rangle} (X_B + X_{\tilde{B}}) = 4V_{|\phi_{BB} \rangle \otimes |\eta_{R_B} \rangle} ((X_B + X_{\tilde{B}}) \otimes 1_{R_B}) = F_{|\phi_{BB} \rangle \otimes |\eta_{R_B} \rangle} ((X_B + X_{\tilde{B}}) \otimes 1_{R_B}) .
\] (IV.8)

Therefore, we obtain (IV.2).
Appendix Supp.V: Rederivation of approximated Eastin-Knill theorem as a corollary of \[\text{(G6)}\]

In this subsection, we rederive the approximate Eastin-Knill theorem from our trade-off relation \[\text{(6)}\] and/or \[\text{(G6)}\]. Following the setup for Theorem 1 in Ref. [14], we assume the following three:

- We assume that the code $C$ is covariant with respect to $\{U_{\theta}^L\}_{\theta \in \mathbb{R}}$ and $\{U_{\theta}^P\}_{\theta \in \mathbb{R}}$, where $U_{\theta}^L := e^{i\theta X_L}$ and $U_{\theta}^P := e^{i\theta X_P}$. We also assume that the code $C$ is an isometry.

- We assume that the physical system $P$ is a composite system of subsystems $\{P_i\}_{i=1}^N$, and that $X_P = \sum_i X_{P_i}$. We also assume that the lowest eigenvalue of each $X_{P_i}$ is 0. (We can omit the latter assumption. See the section Supp.VII.)

- We assume that the noise $N$ is the erasure noise in which the location of the noise is known. To be concrete, the noise $N$ is a CPTP-map from $P$ to $P' := PC$ written as follows:

$$
N(...) := \sum_i \frac{1}{N} |i_C\rangle\langle i_C| \otimes |\tau_i\rangle\langle \tau_i| P_i \otimes \text{Tr}_P[...],
$$

where the subsystem $C$ is the register remembering the location of error, and $\{|i_C\rangle\}$ is an orthonormal basis of $C$. The state $|\tau_i\rangle_{P_i}$ is a fixed state in $P_i$.

In general, $N$ is not a covariant operation. However, we can substitute the following covariant operation $\tilde{N}$ for $N$ without changing $\delta_C$:

$$
\tilde{N}(...) := \sum_i \frac{1}{N} |i_C\rangle\langle i_C| \otimes |0_i\rangle\langle 0_i| P_i \otimes \text{Tr}_P[...]
$$

where $|0_i\rangle$ is the eigenvector of $X_{P_i}$ whose eigenvalue is 0. We can easily see that $\tilde{N} \circ C$ and $N \circ C$ are the same in the sense of $\delta_C$ by noting that we can convert the final state of $\tilde{N} \circ C$ to the final state of $N \circ C$ by the following unitary operation:

$$
W := \sum_i |i_C\rangle\langle i_C| \otimes U_{P_i} \otimes j, j \neq i I_{P_i},
$$

where $U_{P_i}$ is a unitary on $P_i$ satisfying $|\tau_i\rangle = U_{P_i}|0_i\rangle$.

Under the above setup, $\tilde{N} \circ C$ is covariant with respect to $\{U_{\theta}^L\}$ and $\{I_C \otimes U_{\theta}^P\}$. Therefore, we can apply \[\text{(5)}\], \[\text{(6)}\], \[\text{G5}\] and \[\text{G6}\] to this situation. Below, we derive the following approximated Eastin-Knill theorem from \[\text{G6}\].

$$
\frac{D_{XL}}{2\delta_C D_{\max}} \leq N + \frac{D_{XL}}{2D_{\max}}.
$$

Here $D_{\max} := \max_i D_{P_i}$. This inequality is the same as the inequality in Theorem 1 of [14], apart from the irrelevant additional term $D_{XL}/2D_{\max}$. (In Theorem 1 of [14], $\frac{D_{XL}}{2\delta_C D_{\max}} \leq N$ is given.) We can also derive a very similar inequality from \[\text{G6}\]. When we use \[\text{(6)}\] instead of \[\text{G6}\], the coefficient 1/2 in the lefthand side of \[\text{(V.4)}\] becomes 1/4.

We remark that although the bound \[\text{(V.4)}\] is a little bit weaker than the bound in Theorem 1 of Ref. [14], it is still remarkable, because \[\text{(V.4)}\] is given as a corollary of more general inequality \[\text{G6}\].

**Proof of \[\text{(V.4)}\]:** We construct an implementation of $\tilde{N} \circ C$ by combining the following implementations of $C$ and $\tilde{N}$. As the implementation of $C$, we take a system $B$ satisfying $LB = P$, a Hermitian operator $X_B$, a symmetric state $\rho_B$ on $B$ with respect to $X_B$, and a unitary $U$ on $LB$ satisfying

$$
U(X_L + X_B)U^\dagger = X_P,
$$

$$
|\rho_B, X_B\rangle = 0.
$$

$$
C(...) = U(... \otimes \rho_B)U^\dagger
$$

The existence of such $B$, $X_B$, $U$, and $\rho_B$ is guaranteed since $C$ is an isometry and any covariant operation is realized by an invariant unitary and a symmetric state (see Appendix A in the main text).
As an implementation of \( \tilde{N} \), we take a composite system \( B_1 := C \tilde{P}_1 \ldots \tilde{P}_N \) where each \( \tilde{P}_i \) is a copy system of \( P_i \) which has \( X_P \) that is equal to \( X_{P_i} \). We also define a state \( \rho_{B_1} \) on \( B_1 \) and a unitary \( V \) on \( P B_1 \) as follows

\[
\rho_{B_1} := \frac{1}{N} \sum_{j=1}^{N} |j\rangle \langle j| \otimes (\otimes_{i=1}^{N} |0_i\rangle \langle 0_i|)_{\tilde{P}_i} \tag{V.8}
\]

\[
V := \sum_{k} |k\rangle \langle k| \otimes S_{\tilde{P}_kp_k} (\otimes_{j,j \neq k} I_{\tilde{P}_j}), \tag{V.9}
\]

where \( S_{\tilde{P}_kp_k} \) is the swap unitary between \( \tilde{P}_k \) and \( P_k \) and \( I_{\tilde{P}_j} \) is the identity operator on \( \tilde{P}_j \). Then, \( \rho_{B_1} \) and \( V \) satisfies

\[
V(X_P \otimes I_P \otimes I_C + I_P \otimes X_P \otimes I_C) V^\dagger = X_P \otimes I_P \otimes I_C + I_P \otimes X_P \otimes I_C, \tag{V.10}
\]

\[
[\rho_{B_1}, X_P \otimes I_C] = 0, \tag{V.11}
\]

\[
\tilde{N}(\ldots) = \text{Tr}_{\tilde{P}}[V(\ldots \otimes \rho_{B_1}) V^\dagger] \tag{V.12}
\]

where \( \tilde{P} = \tilde{P}_1 \ldots \tilde{P}_N \) and \( X_P = \sum_{j=1}^{N} X_{\tilde{P}_j} \).

For the above implementation, from (G6) and \( \delta_C \geq \max_{|\psi_{LR_L}|} \delta \), we obtain the following relation for an arbitrary \( |\psi_{LR_L}\rangle \):

\[
\frac{\mathcal{A}_2}{\delta_C} \leq 2 \sqrt{V_{\rho_P}^{(X_P^n)}} + \Delta_{\text{max}}, \tag{V.13}
\]

where \( \rho_P^{(X)} \) is the final state of \( \tilde{P} \).

To derive \( V.4 \) from \( \text{(G6)} \), let us evaluate \( \mathcal{A}_2 \), \( \Delta_{\text{max}} \) and \( V_{\rho_P^{(X_P^n)}} \) for the following \( |\psi_{LR_L}\rangle \):

\[
|\psi_{LR_L}\rangle := \frac{|0_L\rangle|0_{RL_L}\rangle + |1_L\rangle|1_{RL_L}\rangle}{\sqrt{2}}, \tag{V.14}
\]

where \( |0_L\rangle \) and \( |1_L\rangle \) are the maximum and minimum eigenvectors of \( X_L \). Due to the definition of \( \mathcal{A}_2 \), we obtain

\[
\mathcal{A}_2 \geq \frac{1}{2} \sum_{i=0}^{1} \left| (X_l \langle j_L|l_L\rangle - (X_P \otimes I_C) \mathcal{E}_{(j_L|j_L\rangle}) - (X_L \langle 0_L|0_L\rangle + |1_L\rangle|1_L\rangle)/2 - (X_P \otimes I_C) \mathcal{E}_{(0_L|0_L\rangle + |1_L\rangle|1_L\rangle)/2} \right| \tag{V.15}
\]

Due to \( \text{(V.1)} \) and \( \text{(V.5)} \), for any \( \rho_L \) on \( L \),

\[
(X_P \otimes I_C) \mathcal{E}_{(\rho_L)} = \left( 1 - \frac{1}{N} \right) (X_L \otimes I_L) \rho_L + (X_B \otimes I_L) \rho_B + \frac{1}{N} \sum_{i=1}^{N} (X_{\tilde{P}_i} \otimes I_L) |0_i\rangle \langle 0_i| \\
= \left( 1 - \frac{1}{N} \right) (X_L \otimes I_L) \rho_L + (X_B \otimes I_L) \rho_B. \tag{V.16}
\]

Therefore, we obtain

\[
\mathcal{A}_2 \geq \frac{1}{2N} \sum_{j=0}^{1} \left| (X_L \langle j_L|j_L\rangle - (X_L \langle 0_L|0_L\rangle + |1_L\rangle|1_L\rangle)/2 \right| \\
\geq \frac{D_{X_L}}{2N}. \tag{V.17}
\]

By definition of \( \Delta_{\text{max}} \), we obtain

\[
\Delta_{\text{max}} = \max_{\rho \text{ on the support of } (|0_L\rangle\langle 0_L| + |1_L\rangle\langle 1_L|)/2 \text{ of } \rho} \frac{1}{N} \left| (X_L \rho) - (X_L \langle 0_L|0_L\rangle + |1_L\rangle|1_L\rangle)/2 \right| \leq \frac{D_{X_L}}{2N}. \tag{V.18}
\]
To evaluate $V_{\rho^f_p}(X_{\tilde{P}})$, we note that

$$\rho^f_p = \frac{1}{N} \sum_{h=1}^{N} \rho^f_h \otimes (\otimes_{i:i \neq h} |0_i\rangle \langle 0_i|) \quad (V.19)$$

where $\rho^f_h := \text{Tr}_{-P_h} [C(|0_L\rangle \langle 0_L| + |1_L\rangle \langle 1_L|)/2]$. Therefore,

$$\langle X^2_{\tilde{P}} \rangle_{\rho^f_p} = \frac{\sum_h \langle X^2_{\tilde{P}_h} \rangle_{\rho^f_h}}{N} \quad (V.20)$$

$$\langle X_{\tilde{P}} \rangle_{\rho^f_p} = \frac{\sum_h \langle X_{P_h} \rangle_{\rho^f_h}}{N} \quad (V.21)$$

With using the above, we evaluate $V_{\rho^f_p}(X_{\tilde{P}})$ as follows:

$$V_{\rho^f_p}(X_{\tilde{P}}) = \langle X^2_{\tilde{P}} \rangle_{\rho^f_p} - \langle X_{\tilde{P}} \rangle_{\rho^f_p}^2 = \frac{\sum_h \langle X^2_{\tilde{P}_h} \rangle_{\rho^f_h}}{N} - \left( \frac{\sum_h \langle X_{P_h} \rangle_{\rho^f_h}}{N} \right)^2 = V^c_Q(x) \leq \frac{D^2_{\text{max}}}{4} \quad (V.22)$$

where $V^c_Q(x)$ is the variance of a classical distribution of $Q$ on a set of real numbers $X$ defined as follows:

$$Q(x) := \sum_{h=1}^{N} P_h(x) \quad (V.23)$$

$$P_h(x) := \begin{cases} \langle x_k | \rho^f_h | x_h \rangle & (x \in X_h) \\ 0 & (\text{otherwise}) \end{cases} \quad (V.24)$$

$$X_h := \{ x | \text{eigenvalues of } X_{P_h} \} \quad (V.25)$$

$$X := \bigcup_{h=1}^{N} X_h \quad (V.26)$$

where $|x_h\rangle$ is an eigenvector of $X_{P_h}$ whose eigenvalue is $x$.

Combining the above, we obtain (V.4). □

**Appendix Supp.VI: Generalization of main results to the case of general Lie group symmetry**

In this section, we generalize the results in the main text to the case of general Lie group symmetries. In the first subsection, we derive a variation of the main results ((5) and (6) in the main text) for the case of the conservation law of $X$, as preliminary. In the variation, we use $A_V$ which represents the variance of the change of local conserved quantity $X$ instead of $A$. In the second subsection, we extend the variation to the case of general symmetries.

1. **Variance-type lower bound of recovery error for the cases of $U(1)$ and $\mathbb{R}$**

In this subsection, we derive a variation of the main results for the case of the conservation law of $X$. We consider Setup 1 with the conservation law of $X$: $X_A + X_B = U^\dagger (X_A + X_B U)$. For an arbitrary decomposition of $\rho_A := \sum_j p_j \rho_{j,A}$, we define the following quantity:

$$A_V(\{p_j, \rho_{j,A}\}, \mathcal{E}) := \sum_j p_j \Delta^2_j \quad (VI.1)$$
Hereafter, we abbreviate $A_V(\{p_j, \rho_{j,A}\}, \mathcal{E})$ as $A_V$. We remark that the quantity $A_V$ depends on the decomposition of $\rho_A$, unlike $A$.

For $A_V$, the following trade-off relation holds:

$$\frac{A_V}{8\delta^2} \leq F + B,$$  \hspace{1cm} (VI.2)

$$\frac{A_V}{8\delta^2} \leq F_f + B,$$  \hspace{1cm} (VI.3)

where $\delta$, $F$ and $F_f$ are the same as in (5) and (6), and $B$ is defined as follows:

$$B := \sum_j \frac{\Delta_j^2}{2} + 8(V_{\rho_A}(X_A) + V_{E(\rho_A)}(X_{A'})).$$  \hspace{1cm} (VI.4)

**Proof of (VI.2) and (VI.3):** To derive (VI.2) and (VI.3), we use the following mean-variance-distance trade-off relation which holds for arbitrary states $\rho$ and $\sigma$ and an arbitrary Hermitian operator $X$ [2]:

$$\text{Tr}[(\rho - \sigma)X]^2 \leq D_F(\rho, \sigma)^2 ((\sqrt{V_\rho(X)} + \sqrt{V_\sigma(X)})^2 + \text{Tr}((\rho - \sigma)X|^2).$$  \hspace{1cm} (VI.5)

With using (VI.5), Lemma 1 and (F1)–(F3), we derive (VI.2) and (VI.3), in the very similar way to (5) and (6).

Let us take an arbitrary decomposition of $\rho_A$ as $\rho_A = \sum_j p_j \rho_{j,A}$. Then, the following equation follows from (F1):

$$|\Delta_j| = |(X_{B'})_{\rho_{j,B'}} - (X_{B'})_{\rho_{B'}}|.$$  \hspace{1cm} (VI.6)

We firstly evaluate $A_V$ as follows:

$$A_V = \sum_j p_j (\langle X_{B'} \rangle_{\rho_{j,B'}} - \langle X_{B'} \rangle_{\rho_{B'}})^2$$

$$\leq \sum_j q_j D_F(\rho_{j,B'}, \rho_{B'})^2 \left( (\sqrt{V_{\rho_{j,B'}}(X_{B'})} + \sqrt{V_{\rho_{B'}}(X_{B'})})^2 + \Delta_j^2 \right).$$  \hspace{1cm} (VI.7)

Here we use (VI.6) in (a), (F1) in (b).

Second, we evaluate $(\sqrt{V_{\rho_{j,B'}}(X_{B'})} + \sqrt{V_{\rho_{B'}}(X_{B'})})^2$ in (VI.7) as follows:

$$\left( \sqrt{V_{\rho_{j,B'}}(X_{B'})} + \sqrt{V_{\rho_{B'}}(X_{B'})} \right)^2 \leq 4 \left( \sqrt{V_{\rho_{B}}(X_{B})} + \sqrt{V_{\rho_{A}}(X_{A})} + \sqrt{V_{\rho_{A'}}(X_{A'})} \right)^2$$

$$\leq 4 \left( 2V_{\rho_{B}}(X_{B}) + 4V_{\rho_{A}}(X_{A}) + V_{\rho_{A'}}(X_{A'}) \right)$$

$$= 2(F + 8(V_{\rho_{A}}(X_{A}) + V_{\rho_{A'}}(X_{A'}))).$$  \hspace{1cm} (VI.8)

Here we use (F2) and $(x + y)^2 \leq 2(x^2 + y^2)$. By combining (VI.7), (VI.8), Lemma 1 and $\Delta_j^2 \leq \sum_j \Delta_j^2$, we obtain (VI.2):

$$A_V \leq 8\delta^2 \left( F + 8(V_{\rho_{A}}(X_{A}) + V_{\rho_{A'}}(X_{A'})) + \sum_j \frac{\Delta_j^2}{2} \right).$$  \hspace{1cm} (VI.9)

To derive (VI.3), we evaluate $(\sqrt{V_{\rho_{j,B'}}(X_{B'})} + \sqrt{V_{\rho_{B'}}(X_{B'})})^2$ in (VI.7) in a different way:

$$\left( \sqrt{V_{\rho_{j,B'}}(X_{B'})} + \sqrt{V_{\rho_{B'}}(X_{B'})} \right)^2 \leq \left( \sqrt{V_{\rho_{B}}(X_{B})} + \sqrt{V_{\rho_{A}}(X_{A})} + \sqrt{V_{\rho_{A'}}(X_{A'})} \right)^2$$

$$\leq 4 \left( \sqrt{V_{\rho_{B}}(X_{B})} + \sqrt{V_{\rho_{A}}(X_{A})} + \sqrt{V_{\rho_{A'}}(X_{A'})} \right)^2$$

$$\leq 4 \left( 2V_{\rho_{B}}(X_{B}) + 4V_{\rho_{A}}(X_{A}) + V_{\rho_{A'}}(X_{A'}) \right)$$

$$= 2(F + 8(V_{\rho_{A}}(X_{A}) + V_{\rho_{A'}}(X_{A'}))).$$  \hspace{1cm} (VI.10)
Here we use \( \{F_2, F_3\} \) and \( (x + y)^2 \leq 2(x^2 + y^2) \).

By combining (VI.7), (VI.10), Lemma 1 and \( \Delta_j^2 \leq \sum_j \Delta_j^2 \), we obtain (VI.3):

\[
\mathcal{A}_V \leq 8\delta^2 \left( F_f + 8(V_{\rho_A}(X_A) + V_{\rho_A}(X_{A'}) + \frac{\sum_j \Delta_j^2}{2}) \right). \tag{VI.11}
\]

2. Main results for general symmetry: Limitations of recovery error for general Lie groups

Now, we introduce the generalized version of the main results. We consider Setup 1, and assume that \( U \) is restricted by some Lie group symmetry. To be more concrete, we take an arbitrary Lie group \( G \) and its unitary representations \( \{V_{g,\alpha}\}_{g \in G} (\alpha = A, B, A', B') \). We assume that \( U \) satisfies the following relation:

\[
U(V_{g,A} \otimes V_{g,B}) = (V_{g,A'} \otimes V_{g,B'})U, \quad g \in G. \tag{VI.12}
\]

Let \( \{X^{(a)}\} (\alpha = A, B, A', B') \) be an arbitrary basis of Lie algebra corresponding to \( \{V_{g,\alpha}\}_{g \in G} \). Then, for an arbitrary decomposition \( \rho_A = \sum_j p_j \rho_{j,A} \), the following matrix inequalities hold:

\[
\frac{\mathcal{A}_V}{8\delta^2} \leq \tilde{F} + \tilde{B}, \tag{VI.13}
\]

\[
\frac{\mathcal{A}_V}{8\delta^2} \leq \tilde{F}_f + \tilde{B}, \tag{VI.14}
\]

where \( \preceq \) is the inequality for matrices, and \( \mathcal{A}_V \) and \( \tilde{B} \) are matrices whose components are defined as follows:

\[
\mathcal{A}_{V,ab} := \sum_j p_j \Delta_j^{(a)} \Delta_j^{(b)} \tag{VI.15}
\]

\[
\Delta_j^{(a)} := \left( (X^{(a)})_{\rho_j} - (X^{(a)})_{\rho_A} \right) - \left( (X^{(a)})_{\rho_A} - (X^{(a)})_{\rho_A} \right) \tag{VI.16}
\]

\[
\tilde{B}_{ab} := 8(Cov_{\rho_A}(X^{(a)}_{\rho_A} : X^{(b)}_{\rho_A}) + Cov_{\rho_A}(X^{(a)}_{\rho_A} : X^{(b)}_{\rho_A})) + \frac{\sum_j \Delta_j^{(a)} \Delta_j^{(b)}}{2}. \tag{VI.17}
\]

and \( \tilde{F} \) and \( \tilde{F}_f \) are the Fisher information matrices

\[
\tilde{F} := \tilde{F}_{\phi_{R_B}}(\{X^{(a)}_{\rho_A} \otimes 1_{R_B}\}), \tag{VI.18}
\]

\[
\tilde{F}_f := \tilde{F}_{\phi_{R_B}}(\{X^{(a)}_{\rho_A} \otimes 1_{R_B}\}), \tag{VI.19}
\]

where the Fisher information matrix \( \tilde{F}_{\xi}(\{X^{(a)}\}) \) is defined as follows for a given state \( \xi \) and given Hermite operators \( \{X^{(a)}\} \):

\[
\tilde{F}_{\xi}(\{X^{(a)}\})_{ab} = 2 \sum_{i,i'} \frac{(r_i - r_{i'})^2}{r_i + r_{i'}} X^{(a)}_{ii'} X^{(b)}_{ii'} \tag{VI.20}
\]

Here, \( r_i \) is the \( i \)-th eigenvalue of the density matrix \( \xi \) with the eigenvector \( \psi_i \), and \( X^{(a)}_{ii'} := \langle \psi_i | X^{(a)} | \psi_{i'} \rangle \).

**Proof of (VI.13) and (VI.14):** We first show (VI.13). Since \( \mathcal{A}_V, \tilde{F} \) and \( \tilde{B} \) are real symmetric matrices, we only have to show the following relation holds for arbitrary real vector \( \lambda \):

\[
\lambda^T \frac{\mathcal{A}_V}{8\delta^2} \lambda \leq \lambda^T (\tilde{F} + \tilde{B}) \lambda. \tag{VI.21}
\]

By definition of \( \mathcal{A}_V, \tilde{F} \) and \( \tilde{B} \), the inequality (VI.21) is equivalent to (VI.2) whose \( X_A, X_{A'} \) and \( X_B \) are substituted by \( X^{(a)}_{\lambda} = \sum \lambda_a X^{(a)}_{\lambda_a} (\alpha = A, A', B) \) and \( \{\lambda_a\} \) are the components of \( \lambda \). Therefore, we only have to show that the following equality holds for arbitrary \( \lambda \):

\[
U(X_{A,\lambda} + X_{B,\lambda})U^\dagger = X_{A',\lambda} + X_{B',\lambda}. \tag{VI.22}
\]
Due to (VI.12), for any $a$, the following relation holds:

$$U(X_A^{(a)} + X_B^{(a)}) = (X_A^{(a)} + X_B^{(a)})U. \tag{VI.23}$$

Therefore, (VI.22) holds, and thus we obtain (VI.13). We can obtain (VI.14) in the same way.

\[\Box\]

3. Limitations of recovery error for general symmetry in information recovery without using $R_B$

In this subsection, we extend (VI.13) and (VI.14) to the case of information recoveries without using $R_B$. Let us consider the almost same setup as in the subsection Supp.VI.2. The difference between the present setup and the setup in the subsection Supp.VI.2 is that the recovery operation $\mathcal{R}$ is a CPTP-map $\mathcal{A}'$ to $\mathcal{A}$. Then, the recovery error is $\delta$ which is defined in (IV.1).

As is explained in the section Supp.IV, since the inequality $\delta \geq \delta$ holds in general, we can substitute $\delta$ for $\delta$ in (VI.13) and (VI.14). Moreover, we can derive the following more strong inequality from (VI.2):

$$\frac{\hat{A}_V}{8\delta^2} \leq \hat{F}_B + \hat{B}, \tag{VI.24}$$

where $\hat{F}_B := \tilde{F}_B((X_B^{(a)}))$.

The proof of (VI.24) is very similar to the proof of (IV.2):

**Proof of (VI.24):** As in the proof of (VI.13), since $\hat{A}_V$, $\hat{F}_B$ and $\hat{B}$ are real symmetric matrices, we only have to show the following inequality for an arbitrary real vector $\lambda$:

$$X^T \frac{\hat{A}_V}{8\delta^2} \lambda \leq X^T (\hat{F}_B + \hat{B}) \lambda. \tag{VI.25}$$

We take a quantum system $\tilde{B}$ whose dimension is the same as $B$, and a purification $|\phi_{B\tilde{B}}\rangle$ of $\rho_{B} := \text{Tr}_{RB}[\phi_{B\tilde{B}}]$. From $|\phi_{B\tilde{B}}\rangle$ and $U$, we define a set $\tilde{I} := (|\phi_{B\tilde{B}}\rangle \otimes |\eta_{\tilde{B}}\rangle, U \otimes 1_{\tilde{B}})$. We take the Schmidt decomposition of $|\phi_{B\tilde{B}}\rangle$ as

$$|\phi_{B\tilde{B}}\rangle = \sum_i \sqrt{r_i} |l_B\rangle |l_{\tilde{B}}\rangle, \tag{VI.26}$$

and define $\{X_B^{(a)}\}$ on $\tilde{B}$ corresponding to $\{X_B^{(a)}\}$ as

$$X_B^{(a)} := \sum_{a'} \frac{2\sqrt{r_{a'a}}}{r_{a} + r_{a'}} \langle l_B | X_B^{(a)} | l_{\tilde{B}} \rangle | l_{a'} \rangle | l_{\tilde{B}} \rangle. \tag{VI.27}$$

Note that $\tilde{I}$ is a Steinspring representation of $\mathcal{E}$ and that $U \otimes 1_{\tilde{B}}(X_A^{(a)} + X_B^{(a)} + X_B^{(a)})(U \otimes 1_{\tilde{B}})^\dagger = X_A^{(a')} + X_B^{(a')} + X_B^{(a)}$ for any $a$. Therefore, we obtain the following inequality from (VI.2) by substituting $X_A^{(a)} := \sum a' \lambda_a X_A^{(a')}$ for $X_A$, $X_B^{(a)} := \sum a' \lambda_a X_B^{(a')}$ for $X_B$, $X_A^{(a')} := \sum a \lambda_a X_A^{(a')}$ for $X_A'$, and $X_B^{(a')} := \sum a \lambda_a (X_B^{(a')} + X_B^{(a)})$ for $X_B'$:

$$\frac{\hat{A}_V^{(\lambda)}(\psi_{ARA}, \tilde{I})}{8\delta(\psi_{ARA}, \tilde{I})} \leq F^{(\lambda)}|\phi_{B\tilde{B}}\rangle \otimes |\eta_{\tilde{B}}\rangle + B^{(\lambda)}. \tag{VI.28}$$

Here $A_V^{(\lambda)}(\psi_{ARA}, \tilde{I})$, $F^{(\lambda)}|\phi_{B\tilde{B}}\rangle \otimes |\eta_{\tilde{B}}\rangle$ and $B^{(\lambda)}$ are $A_V$, $F$ and $B$ for $(|\phi_{B\tilde{B}}\rangle \otimes |\eta_{\tilde{B}}\rangle, U \otimes 1_{\tilde{B}})$ and $X_\alpha^{(\lambda)} (\alpha = A, B, \tilde{A}, A', B')$.

Since both of $I$ and $\tilde{I}$ gives the same CPTP-map $\mathcal{E}$, and due to the definitions of $A_V^{(\lambda)}(\psi_{ARA}, \tilde{I})$ and $B^{(\lambda)}$,

$$A_V^{(\lambda)}(\psi_{ARA}, \tilde{I}) = \lambda^T A_V \lambda,$$

$$B^{(\lambda)} = \lambda^T B \lambda.$$
Similarly due to (A4),

\[ X^T \tilde{F}_{\rho_B}(\{X_B^{(a)}\})A = F_{\rho_B}(\sum_a \lambda_a X_B^{(a)}) = 4V_{\phi_B} \left( \sum_a \lambda_a (X_B^{(a)} + X_{B'}^{(a)}) \right) = F^{(A)}_{\rho_B}. \]  

(VI.31)

Moreover, since |\phi_{B\tilde{B}}\rangle \otimes |\eta_{RB}\rangle is a tensor product between B\tilde{B} and RB, the state of B\tilde{B}RB after U is also another tensor product state between B\tilde{B} and RB. Therefore, we obtain

\[ \delta(\psi_{AR}, \tilde{I}) = \bar{\delta} \]  

(VI.32)

Combining the above, we obtain (VI.24).

4. Applications of the limitations of recovery error for general symmetries

As the cases of U(1) and \( \Re \), we can use the inequalities (VI.13), (VI.14) and (VI.24) (and (VI.14) whose \( \delta \) is substituted by \( \bar{\delta} \)) to various phenomena.

- As (5) and (6), we can apply (VI.13) and (VI.14) to information recovery from scrambling with general symmetry.
- As (IV.2), we can apply (VI.13) to implementation of general unitary dynamics and covariant error correcting codes with covariant errors. With using \( \delta_U \) and \( \delta_C \), we obtain

\[ \frac{\tilde{A}_V}{8\tilde{\delta}_C^2} \leq \tilde{B} \]  

(VII.34)

Appendix Supp.VII: limitations of recovery error for the case where the conservation law is weakly violated

In this section, we consider the case where the conservation law of \( X \) is violated. We show that our results hold even in such cases. We consider Setup 1 with the following violated global conservation law:

\[ X_A + X_B = U^\dagger (X_A' + X_B') U + Z. \]  

(VII.1)

Here Z is some perturbation term which describes the strength of the violation of global conservation law. Then, the following two relations hold:

\[ \frac{\mathcal{A} - \mathcal{A}_Z}{2(\sqrt{\mathcal{F}} + 2(\sqrt{V_{\rho_A}}(X_A) + \sqrt{V_{\rho_B}}(X_A'))) + \mathcal{A}^{(2)} + \mathcal{A}_Z^{(2)} + 2\sqrt{V_Z}} \leq \delta, \]  

(VII.2)

\[ \frac{\mathcal{A} - \mathcal{A}_Z}{2(\sqrt{\mathcal{F}} + \mathcal{A}^{(2)} + \mathcal{A}_Z^{(2)})} \leq \delta. \]  

(VII.3)

Here \( V_Z := V_{\rho_A \otimes \rho_B}(Z) \) and

\[ \mathcal{A}_Z := \max_{\{p_j, \rho_j, A\}} \sum_j p_j |(Z)_{\rho_j, A} \otimes \rho_B - (Z)_{\rho_A \otimes \rho_B}|, \]  

(VII.4)

\[ \mathcal{A}_Z^{(2)} := \max_{\{p_j, \rho_j, A\}} \sqrt{\sum_j p_j |(Z)_{\rho_j, A} \otimes \rho_B - (Z)_{\rho_A \otimes \rho_B}|^2}, \]  

(VII.5)

\[ \mathcal{A}^{(2)} := \max_{\{p_j, \rho_j, A\}} \sqrt{\sum_j p_j |\Delta_j|^2}, \]  

(VII.6)

where \( \{p_j, \rho_j, A\} \) runs \( \rho_A = \sum_j p_j \rho_j, A \).
To simplify (VII.2) and (VII.3), we can use the following relations (we prove them in the end of this section):

\[
A_Z \leq A_Z^{(2)} \leq \sqrt{V_Z},
\]

\[
A^{(2)} \leq \Delta_{\text{max}} \leq 2\Delta_+,
\]

\[
\sqrt{V_{\rho_A}(X_A)} + \sqrt{V_{\rho_A}(X_A')} \leq \Delta_+, \tag{VII.9}
\]

\[
A_Z \leq M_{\rho_A}(Z_A),
\]

\[
A_Z^{(2)} \leq \sqrt{V_{\rho_A}(Z_A)}. \tag{VII.11}
\]

where \( Z_A := \text{Tr}_B[(1_A \otimes \rho_B)Z] \) and \( M_{\rho_A}(Z_A) := \langle [Z_A - (Z_A)_A]_A \rangle_F \). For example, by using (VII.7), (VII.8) and (VII.11), we obtain the following inequalities from (VII.2) and (VII.3):

\[
\frac{A - \sqrt{V_Z}}{2(\sqrt{\mathcal{F}} + 4\Delta_+ + 3\sqrt{V_Z})} \leq \delta, \tag{VII.12}
\]

\[
\frac{A - \sqrt{V_Z}}{2(\sqrt{\mathcal{F}} + \Delta_{\text{max}} + \sqrt{V_Z})} \leq \delta. \tag{VII.13}
\]

We remark that we have introduced (VII.12) in the section III A of the main text. Similarly, the following relations also hold:

\[
\frac{A_2 - A_Z}{\sqrt{\mathcal{F}} + 2(\sqrt{V_{\rho_A}(X_A)} + \sqrt{V_{\rho_A}(X_A')})} \leq \delta, \tag{VII.14}
\]

\[
\frac{A_2 - A_Z}{\sqrt{\mathcal{F}} + A^{(2)} + A_Z^{(2)}} \leq \delta. \tag{VII.15}
\]

These inequalities have two important messages. First, when \( Z = \mu I \) where \( \mu \) is an arbitrary real number, the inequalities (VII.12) and (VII.13) are valid, since in that case \( A_Z = V_Z = V_{\rho_A}(Z_A) = 0 \) holds. Therefore, we can omit the assumption that the lowest eigenvalue of \( X_A \) is 0, which is used in the re-derivation of the approximate Eastin-Knill theorem in the section Supp. V. Second, our trade-off relations become trivial only when \( A \leq A_Z \). As we show in the section 3 in the main text, the inequality \( A \geq M_{\gamma}(1 - \epsilon) \) holds in the Hayden-Preskill black hole model. Therefore, when \( M_{\gamma} \) is not so large, our message on black holes does not radically change. Even when the global conservation law is weakly violated, black holes are foggy mirrors.

**Proof of (VII.2), (VII.3), (VII.14) and (VII.15):** Hereafter we use the abbreviation \( X_{AB} = X_A + X_B \) and \( X_{A'B'} = X_A' + X_B' \). Then, for an arbitrary state \( \xi \) on \( AB \), we can transform \( U_{\xi U^\dagger}(X_{A'B'}) \) as follows

\[
V_{U_{\xi U^\dagger}(X_{A'B'})} = (X_{A'B'})^2_{U_{\xi U^\dagger}} - (X_{A'B'})^2_{U_{\xi U^\dagger}},
\]

\[
= ((U^\dagger X_{A'B'})U)^2_{\xi} - (U^\dagger X_{A'B'})U)^2_{U_{\xi U^\dagger}} \tag{VII.16}
\]

\[
= ((X_{AB} - Z)^2_{\xi} - (X_{AB} - Z)_{\xi})^2 \tag{VII.17}
\]

\[
= V_{\xi}(X_{AB} - Z) - V_{\xi}(X_{AB}) - 2\text{Cov}_{\xi}(X_{AB} : Z) + V_{\xi}(Z). \tag{VII.18}
\]

Due to \(|\text{Cov}_{\xi}(X_{AB} : Z)| \leq \sqrt{V_{\xi}(X_{AB})}\sqrt{V_{\xi}(Z)}\), we obtain

\[
\left(\sqrt{V_{\xi}(X_{AB})} - \sqrt{V_{\xi}(Z)}\right)^2 \leq V_{U_{\xi U^\dagger}(X_{A'B'})} \leq \left(\sqrt{V_{\xi}(X_{AB})} + \sqrt{V_{\xi}(Z)}\right)^2 \tag{VII.19}
\]

Now, let us set \( \xi = \xi_A \otimes \xi_B, \xi_{A'}^U := \text{Tr}_B[U(\xi_A \otimes \xi_B)U^\dagger] \) and \( \xi_{B'}^U := \text{Tr}_A[U(\xi_A \otimes \xi_B)U^\dagger] \). Then,

\[
V_{U_{\xi U^\dagger}(X_{A'B'})} = V_{\xi_{A'}^U}(X_{A'}) + 2\text{Cov}_{U_{\xi U^\dagger}(X_{A'} : X_{B'})} + V_{\xi_{B'}^U}(X_{B'}). \tag{VII.20}
\]

Due to \(|\text{Cov}_{U_{\xi U^\dagger}(X_{A'} : X_{B'})}| \leq \sqrt{V_{\xi_{A'}^U}(X_{A'})}\sqrt{V_{\xi_{B'}^U}(X_{B'})}\), we obtain

\[
\left(\sqrt{V_{\xi_{A'}^U}(X_{A'})} - \sqrt{V_{\xi_{B'}^U}(X_{B'})}\right)^2 \leq V_{U_{\xi U^\dagger}(X_{A'B'})} \leq \left(\sqrt{V_{\xi_{A'}^U}(X_{A'})} + \sqrt{V_{\xi_{B'}^U}(X_{B'})}\right)^2 \tag{VII.21}
\]
Substituting \( \xi = \xi_A \otimes \xi_B \) into (VII.17) and combining it with (VII.19), we obtain
\[
\sqrt{V_{\xi_B'}(X_B')} \leq \sqrt{V_{\xi_B}(X_B)} + \sqrt{V_{\xi_A}(X_A)} + \sqrt{V_{\xi_A'}(X_A')} + \sqrt{V_{\xi_B \otimes \xi_B}(Z)}.
\]
(VII.20)

Due to (VII.1), we obtain
\[
(X_A)_{\xi_A} - \langle X_A' \rangle_{\xi_A'} = -\langle X_B \rangle_{\xi_B} + \langle X_B' \rangle_{\xi_B'} + (Z)_{\xi_A \otimes \xi_B}.
\]
(VII.21)

Therefore, for the decomposition \( \rho_A = \sum_j p_j \rho_j \) such that \( A = \sum_j p_j |\Delta_j| \), we obtain
\[
|\langle X_B' \rangle_{\rho_{j,B'}} - \langle X_B \rangle_{\rho_{j,B}}| + |\langle Z \rangle_{(\rho_j,A-\rho_A)\otimes \rho_B}| \leq |\Delta_j| \leq |\langle X_B' \rangle_{\rho_{j,B'}} - \langle X_B' \rangle_{\rho_{j,B'}}| + |\langle Z \rangle_{(\rho_j,A-\rho_A)\otimes \rho_B}|
\]
(VII.22)

By using (VII.20) and (VII.22) instead of (F2) and (F1), we obtain (VII.2) by the same way as (5). We choose an ensemble \( \{p_j, \rho_j, A\} \) satisfying \( A = \sum_j p_j |\Delta_j| \). Then, we obtain
\[
A = \sum_j p_j |\Delta_j|
\]
\[
\leq \sum_j p_j (|\langle X_B' \rangle_{\rho_{j,B'}} - \langle X_B \rangle_{\rho_{j,B}}| + |\langle Z \rangle_{(\rho_j,A-\rho_A)\otimes \rho_B}|)
\]
\[
\leq \sum_j p_j |\langle Z \rangle_{(\rho_j,A-\rho_A)\otimes \rho_B}| + \sum_j p_j D_F(\rho_{j,B},\rho_{j,B'}) \left( \sqrt{V_{\rho_{j,B'}}(X_B')} + \sqrt{V_{\rho_{j,B}}(X_B')} + |\langle X_B' \rangle_{\rho_{j,B'}} - \langle X_B' \rangle_{\rho_{j,B'}}| \right)
\]
\[
\leq A_Z + \sum_j p_j D_F(\rho_{j,B},\rho_{j,B'}) \left( 2\sqrt{V_{\rho_B}(X_B)} + \sqrt{V_{\rho_{j,A}}(X_A)} + \sqrt{V_{\rho_{j,A'}}(X_A')} + \sqrt{V_{\rho_A}(X_A)} + \sqrt{V_{\rho_{A'}}(X_A')}
\right.
\]
\[
+ 2\sqrt{V_Z} + |\Delta_j| + |\langle Z \rangle_{(\rho_j,A-\rho_A)\otimes \rho_B}|
\]
\[
\leq A_Z + 2\delta \left( \sqrt{\mathcal{F}} + \sqrt{V_{\rho_B}(X_A)} + \sqrt{V_{\rho_{A'}}(X_A')} + 2\sqrt{V_Z} \right)
\]
\[
+ \sum_j p_j D_F(\rho_{j,B}',\rho_{j,B'})^2 \left( \sum_j p_j V_{\rho_{j,B}}(X_A) + \sum_j p_j V_{\rho_{j,A'}}(X_A) + \sum_j p_j |\Delta_j|^2 + \sum_j p_j |\langle Z \rangle_{(\rho_j,A-\rho_A)\otimes \rho_B}|^2 \right)
\]
\[
\leq A_Z + 2\delta \left( \sqrt{\mathcal{F}} + 2\left( \sqrt{V_{\rho_B}(X_A)} + \sqrt{V_{\rho_{A'}}(X_A')} \right) + 2\sqrt{V_Z} + A^{(2)} + A_Z^{(2)} \right).
\]
(VII.23)

Similarly, we derive (VII.3) as follows:
\[
A = \sum_j p_j |\Delta_j|
\]
\[
\leq \sum_j p_j (|\langle X_B' \rangle_{\rho_{j,B'}} - \langle X_B \rangle_{\rho_{j,B}}| + |\langle Z \rangle_{(\rho_j,A-\rho_A)\otimes \rho_B}|)
\]
\[
\leq \sum_j p_j |\langle Z \rangle_{(\rho_j,A-\rho_A)\otimes \rho_B}| + \sum_j p_j D_F(\rho_{j,B},\rho_{j,B'}) \left( \sqrt{V_{\rho_{j,B'}}(X_B')} + \sqrt{V_{\rho_{j,B}}(X_B')} + |\langle X_B' \rangle_{\rho_{j,B'}} - \langle X_B' \rangle_{\rho_{j,B'}}| \right)
\]
\[
\leq A_Z + 2\delta \sqrt{V_{\rho_{j,B'}}(X_B')}
\]
\[
+ \sum_j p_j D_F(\rho_{j,B}',\rho_{j,B'})^2 \left( \sum_j p_j V_{\rho_{j,B}}(X_B') + \sum_j p_j |\Delta_j|^2 + \sum_j p_j |\langle Z \rangle_{(\rho_j,A-\rho_A)\otimes \rho_B}|^2 \right)
\]
\[
\leq A_Z + 2\delta \left( \sqrt{\mathcal{F}} + A^{(2)} + A_Z^{(2)} \right).
\]
(VII.24)

We can show (VII.14) and (VII.15) in the same way.
Finally, we prove \((\text{VII.7})\)–\((\text{VII.11})\).

**Proof of \((\text{VII.7})\)–\((\text{VII.11})\):** The inequalities \((\text{VII.8})\) and \((\text{VII.9})\) are easily obtained from the definition. So, we prove \((\text{VII.7})\), \((\text{VII.10})\) and \((\text{VII.11})\). We firstly show \((\text{VII.7})\) and \((\text{VII.11})\). Since the square of the average is smaller than the average of square, the inequality \(A_Z \leq A_Z^{(2)}\) in \((\text{VII.7})\) clearly holds. We can easily derive the remaining parts of \((\text{VII.7})\) and \((\text{VII.11})\) from the following inequality holds for arbitrary Hermitian \(Y\) and state \(\xi\) and its decomposition \(\xi = \sum_l q_l \xi_l\):

\[
\sum_l q_l (\langle Y \rangle_{\xi_l} - \langle Y \rangle_{\xi})^2 \leq V_{\xi}(Y) \tag{VII.25}
\]

We obtain \((\text{VII.25})\) as follows:

\[
V_{\xi}(Y) = \langle Y^2 \rangle_{\xi} - \langle Y \rangle_{\xi}^2 \\
= \sum_l q_l \langle Y^2 \rangle_{\xi_l} - \left( \sum_l q_l \langle Y \rangle_{\xi_l} \right)^2 \\
\geq \sum_l q_l \langle Y \rangle_{\xi_l}^2 - \left( \sum_l q_l \langle Y \rangle_{\xi_l} \right)^2 \\
= \sum_l q_l \left( \langle Y \rangle_{\xi_l} - \sum_l q_l \langle Y \rangle_{\xi_l} \right)^2 \\
= \sum_l q_l (\langle Y \rangle_{\xi_l} - \langle Y \rangle_{\xi})^2. \tag{VII.26}
\]

Similarly, we can easily derive \((\text{VII.10})\) from the following inequality holds for arbitrary Hermitian \(Y\) and state \(\xi\) and its decomposition \(\xi = \sum_l q_l \xi_l\):

\[
\sum_l q_l |\langle Y \rangle_{\xi_l} - \langle Y \rangle_{\xi}| \leq M_{\xi}(Y) \tag{VII.27}
\]

We obtain \((\text{VII.27})\) as follows:

\[
M_{\xi}(Y) = \langle |Y - \langle Y \rangle_{\xi}| \rangle_{\xi} \\
= \sum_l q_l \langle |Y - \langle Y \rangle_{\xi}| \rangle_{\xi_l} \\
\geq \sum_l q_l |\langle Y - \langle Y \rangle_{\xi} \rangle_{\xi_l}| \\
= \sum_l q_l (\langle Y \rangle_{\xi_l} - \langle Y \rangle_{\xi}). \tag{VII.28}
\]

where we use the inequality \(|\langle H \rangle_{\zeta}| \leq \langle |H| \rangle_{\zeta}\) which holds for arbitrary Hermitian \(H\) and state \(\zeta\) in (a).

\[\blacksquare\]

**Appendix Supp.VIII: Derivations of \((35)\)–\((37)\) and \((38)\) in the main text**

In this section, we derive \((35)\)–\((37)\) and \((38)\) in the main text.

1. **Model**

For the reader's convenience, we write down our model again. Let us consider four qubits, \(A\), \(R_A\), \(B_1\) and \(R_B_1\). We also take a natural number \(b\) and a \(b + 2\)-dimensional system \(B_2\). For \(A\), \(B_1\), \(B_2\), we define Hermitian operators \(X_A\), \(X_B\),
We calculate $A$ and numerical ways. Here, to check our bound, we calculate each term in the RHS of (VIII.9), and give upper bounds of $\delta$ above model:

\[ X_{B_1} := |1\rangle \langle 1|_{B_1}, \]
\[ X_{B_2} := \sum_{x=1}^{b+2} 2x|\langle x|_{B_2}, \]

where $\{ |x\rangle_A \}_{x=0,1}, \{ |x\rangle_{B_1} \}_{x=0,1}, \{ |x\rangle_{B_2} \}_{x=0,...,d+2}$ are orthogonal basis on $A$, $B_1$, and $B_2$, respectively.

On the above system, we prepare the following initial states:

\[ \|\psi_{AR_A}\rangle := \frac{|00\rangle_{AR_A} + |11\rangle_{AR_A}}{\sqrt{2}}, \]
\[ \|\phi_{B_1,R_{B_1}}\rangle := \frac{|00\rangle_{B_1} + |11\rangle_{B_1}}{\sqrt{2}}, \]
\[ \|\phi_{B_2}\rangle := \frac{1}{\sqrt{b}} \sum_{x=1}^{b} |x\rangle_{B_2}. \]

And we prepare the following unitary $U_{AB_1B_2}$ on $AB_1B_2$:

\[ U_{AB_1B_2} := \sum_{1 \leq k \leq b+1} (|1\rangle \langle 1|_{AB_1} \otimes |k-1\rangle \langle k|) + |01\rangle \langle 01|_{AB_1} \otimes |k\rangle \langle k| + |00\rangle \langle 00|_{AB_1} \otimes |k\rangle \langle k| - |01\rangle \langle 01|_{AB_1} \otimes |k-1\rangle \langle k|) + |00\rangle \langle 00|_{AB_1} \otimes |0\rangle \langle 0| + |01\rangle \langle 01|_{AB_1} \otimes |0\rangle \langle 0| + |10\rangle \langle 10|_{AB_1} \otimes |0\rangle \langle 0| + |11\rangle \langle 11|_{AB_1} \otimes |b+1\rangle \langle b+1|. \]

After the unitary $U$, we perform a CPTP map on $AR_B$ to recover the initial state $\|\psi_{AR_A}\rangle$ on $AR_A$. Then, following our framework, we define the minimum recovery error as follows:

\[ \delta = \min_{\mathcal{R}} D_F(\|\psi_{AR_A}\rangle, id_{R_A} \otimes \mathcal{R} \circ \mathcal{E}_{A \rightarrow AR_B}(\|\psi_{AR_A}\rangle)). \]

Here $\mathcal{E}_{A \rightarrow AR_B}(\cdots) := \text{Tr}_{B_1B_2}[U_{AB_1B_2}(\cdots \otimes \phi_{B_2} \otimes \phi_{R_{B_1}})U_{AB_1B_2}^\dagger].$

Since $U_{AB_1B_2}$ conserves $X_A + X_{B_1} + X_{B_2}$ and $\rho_A := \text{Tr}_{R_B}[\|\psi_{AR_A}\rangle] = \frac{|0\rangle \langle 0| + |1\rangle \langle 1|}{2}$, our bound is applicable to the above model:

\[ \delta \geq \frac{A_2}{\sqrt{F_f} + \Delta_{\max}}. \]

Here, to check our bound, we calculate each term in the RHS of (VIII.9), and give upper bounds of $\delta$ with analytical and numerical ways.

\[ a. \textit{analytical lower bound: Derivations of (35) - (37)} \]

Let us calculate $A_2$, $F_f$ and $\Delta_{\max}$ in the RHS of (VIII.9). The calculations correspond to the derivations of (35) - (37). We firstly calculate $A_2$. The definition of $A_2$ is as follows:

\[ A_2 := \sum_{i=0,1} \frac{|\langle X_A |i\rangle \langle i| - \langle X_A |\mathcal{E}_{A \rightarrow A}(i)\rangle \langle i|)|^2}{2}, \]

FIG. S.4. Schematic diagram of the concrete model for numerical check.

$X_{B_1}$ and $X_{B_2}$ as follows:

\[ X_A := |1\rangle \langle 1|_A, \]
\[ X_{B_1} := \sum_{x=1}^{b+2} 2x|\langle x|_{B_1}, \]

\[ X_{B_2} := |1\rangle \langle 1|_{B_2}, \]

\[ \|\psi_{AR_A}\rangle := \frac{|00\rangle_{AR_A} + |11\rangle_{AR_A}}{\sqrt{2}}, \]
\[ \|\phi_{B_1,R_{B_1}}\rangle := \frac{|00\rangle_{B_1} + |11\rangle_{B_1}}{\sqrt{2}}, \]
\[ \|\phi_{B_2}\rangle := \frac{1}{\sqrt{b}} \sum_{x=1}^{b} |x\rangle_{B_2}. \]
where $\mathcal{E}_{A \rightarrow A}(\ldots) := \text{Tr}_{B_1B_2R_1} [U_{AB_1B_2}(\ldots \otimes \phi_{B_2} \otimes \phi_{BR_1}) U^\dagger_{AB_1B_2}]$. We remark that

\begin{align}
\mathcal{E}_{A \rightarrow A}(|0\rangle|0\rangle) &= |1\rangle|1\rangle, \\
\mathcal{E}_{A \rightarrow A}(|1\rangle|1\rangle) &= |0\rangle|0\rangle.
\end{align}

Therefore, we obtain

\[ A_2 = 1. \]  

Next, we evaluate $\Delta_{\text{max}}$. Due to $\Delta_{\text{max}} := \max_{\rho} \sup_{|X_A\rangle = (X_A)e^{-\mathcal{E}_{A \rightarrow A}(\rho)}$ and $\Delta_X = 1$, we obtain

\[ \Delta_{\text{max}} \leq 1. \]  

Next, we evaluate $\mathcal{F}_f$. Note that

\[ \mathcal{F}_f = 4\rho'_{B_1B_2}(X_{B_1} + X_{B_2}), \]

where $\rho'_{B_1B_2} := \text{Tr}_{AR_1R_2} [1_{RA} \otimes U_{AB_1B_2}(\psi_{RA} \otimes \phi_{B_2} \otimes \phi_{BR_1}) 1_{RA} \otimes U^\dagger_{AB_1B_2}]$. To evaluate $\rho'_{B_1B_2}(X_{B_1} + X_{B_2})$, we firstly note that

\begin{align}
U_{AB_1B_2}|\psi_{RA}\rangle &\otimes |\phi_{B_2}\rangle \otimes |\phi_{BR_1}\rangle \\
&= \frac{1}{2}(|0101\rangle_{RA} |\phi_{B_2}^{(-1)}\rangle_{R_1} + |0110\rangle_{RA} |\phi_{B_2}\rangle_{R_1} |\phi_{B_2}\rangle_{R_1} + |1001\rangle_{RA} |\phi_{B_2}\rangle_{R_1} |\phi_{B_2}\rangle_{R_1} + |1010\rangle_{RA} |\phi_{B_2}^{(+1)}\rangle_{R_1}),
\end{align}

where $|\phi_{B_2}^{(-1)}\rangle := \frac{1}{\sqrt{b}} \sum_{x=0}^{b-1} |x\rangle_{B_2}$ and $|\phi_{B_2}^{(+1)}\rangle := \frac{1}{\sqrt{b}} \sum_{x=2}^{b+1} |x\rangle_{B_2}$. Therefore, we obtain

\[ \rho'_{B_1B_2} = \frac{1}{4}(|0\rangle|0\rangle \otimes \phi_{B_2}^{(-1)} + |0\rangle|0\rangle \otimes \phi_{B_2} + |1\rangle|1\rangle \otimes \phi_{B_2} + |0\rangle|0\rangle \otimes \phi_{B_2}^{(+1)}). \]

Therefore, with using $V_{\sum_{x} q_{\rho_{x}}(W) = V_{\langle q_{\rho_{x}} \rangle}(|\langle W |\rangle_{\rho_{x}}\rangle) + \sum_{x} q_{\rho_{x}}(X)\rangle, we obtain

\[ V_{\rho'_{B_1B_2}}(X_{B_1} + X_{B_2}) = \frac{5}{4} + V_{\rho_{B_2}}(X_{B_2}) \]

\begin{align}
&= \frac{5}{4} + \frac{1}{b} \sum_{k=1}^{b} k^2 - \frac{(\sum_{k=1}^{b} k)^2}{b^2} \\
&= \frac{5}{4} + \frac{(b+1)(2b+1)}{6} - \frac{(b+1)^2}{4} \\
&= \frac{b^2 + 14}{12}.
\end{align}

Therefore, we obtain

\[ \mathcal{F}_f = \frac{b^2 + 14}{3}. \]

Hence, $\delta$ is bounded as follows:

\[ \delta \geq \frac{1}{\sqrt{\mathcal{F}_f + 1}} \]

\[ = \frac{1}{\sqrt{\frac{b^2 + 14}{3} + 1}}. \]

b. Analytical upper bound:

Next, we give (38) in the main text, an analytical upper bound for $\delta$. To derive the upper bound, we define the following $V_{ARB_1}$ on $ARB_1$.

\[ V_{ARB_1} := |11\rangle\langle 00|_{ARB_1} + |10\rangle\langle 10|_{ARB_1} + |01\rangle\langle 10|_{ARB_1} + |00\rangle\langle 11|_{ARB_1}. \]
Since the unitary operation $\mathcal{R}_{AB_1}^{V_{AB_1} \to A}(\psi) := \text{Tr}_{B_1} [V_{AB_1} (\psi) V_{AB_1}^\dagger]$ is a CPTP map from $AB_1$ to $A$, we obtain
\begin{equation}
\delta \leq D_F(\psi_{AR_A}, \text{id}_{RA} \otimes \mathcal{R}_{AB_1}^{V_{AB_1} \to A} \circ \mathcal{E}_{A \to AB_1}(\psi_{AR_A})).
\end{equation}

To evaluate the RHS of the above inequality, note the following relations:
\begin{align}
\text{id}_{RA} \otimes \mathcal{R}_{AB_1}^{V_{AB_1} \to A} \circ \mathcal{E}_{A \to AB_1}(\psi_{AR_A}) &= \text{Tr}_{B_1 B_2} [V_{AB_1} U_{AB_2 B_2} \psi_{AR_A} \otimes \phi_{B_2} \otimes \phi_{BR_1} U_{AB_1 B_2}^\dagger V_{AB_1}^\dagger], \quad \text{(VIII.23)} \\
V_{AB_1} U_{AB_2} |\psi_{RA} \rangle \otimes |\phi_{BR_1} \rangle &= \frac{1}{\sqrt{b}} \left( |0111_1 R_A B_{B_1} B_1 \rangle - |1000_1 R_A B_{B_1} B_1 \rangle \right) + |1111_1 R_A B_{B_1} B_1 \rangle |\phi_{B_2} \rangle + |1100_1 R_A B_{B_1} B_1 \rangle |\phi_{B_2}^{(+)} \rangle \\
&= \sqrt{\frac{b-2}{b}} |\psi_{AR_A} \rangle |\phi_{BR_1} \rangle |\phi_{B_2}^{'} \rangle \\
&+ \frac{1}{2} \sqrt{\frac{b-2}{b}} |0011_1 R_A B_{B_1} B_1 \rangle (|0 \rangle + |1 \rangle) + |0000_1 R_A B_{B_1} B_1 \rangle (|1 \rangle + |b \rangle) + |1111_1 R_A B_{B_1} B_1 \rangle (|1 \rangle + |b \rangle) + |1100_1 R_A B_{B_1} B_1 \rangle (|b \rangle + |b + 1 \rangle) \\
\text{(VIII.24)}
\end{align}

where $|\phi_{B_2}^{'} \rangle := \sqrt{\frac{1}{b-2}} \sum_{x=2}^{b-1} |x \rangle$ and we use (VIII.16) in (a).

Using the above, we evaluate $F(\psi_{AR_A}, \text{id}_{RA} \otimes \mathcal{R}_{AB_1}^{V_{AB_1} \to A} \circ \mathcal{E}_{A \to AB_1}(\psi_{AR_A})$) as follows:
\begin{align}
F^2(\psi_{AR_A}, \text{id}_{RA} \otimes \mathcal{R}_{AB_1}^{V_{AB_1} \to A} \circ \mathcal{E}_{A \to AB_1}(\psi_{AR_A})) = &\langle \psi_{AR_A} | \text{id}_{RA} \otimes \mathcal{R}_{AB_1}^{V_{AB_1} \to A} \circ \mathcal{E}_{A \to AB_1}(\psi_{AR_A}) | \psi_{AR_A} \rangle \\
= &\langle \psi_{AR_A} | \text{id}_{RA} \otimes \mathcal{R}_{AB_1}^{V_{AB_1} \to A} \circ \mathcal{E}_{A \to AB_1}(\psi_{AR_A}) | \psi_{AR_A} \rangle \\
= &\langle \psi_{AR_A} | \text{id}_{RA} \otimes \mathcal{R}_{AB_1}^{V_{AB_1} \to A} \circ \mathcal{E}_{A \to AB_1}(\psi_{AR_A}) | \psi_{AR_A} \rangle \\
= &\langle \psi_{AR_A} | \text{id}_{RA} \otimes \mathcal{R}_{AB_1}^{V_{AB_1} \to A} \circ \mathcal{E}_{A \to AB_1}(\psi_{AR_A}) | \psi_{AR_A} \rangle \\
\geq &\langle \psi_{AR_A} | \phi_0 \rangle \langle \phi_0 | V_{AB_1} U_{AB_2 B_2} \psi_{AR_A} \otimes \phi_{B_2} \otimes \phi_{BR_1} U_{AB_1 B_2}^\dagger V_{AB_1}^\dagger | \phi_0 \rangle |\psi_{AR_A} \rangle \\
= &\frac{b-2}{b}.
\end{align}

where $\{\phi_j\}$ is a orthonormal basis of $R_{B_1} B_1 B_2$ satisfying $\phi_0 = \phi_{BR_1} \otimes \phi_{B_2}$. Due to $DF = \sqrt{1 - F^2}$, we obtain
\begin{equation}
D_F(\psi_{AR_A}, \text{id}_{RA} \otimes \mathcal{R}_{AB_1}^{V_{AB_1} \to A} \circ \mathcal{E}_{A \to AB_1}(\psi_{AR_A})) \leq \sqrt{\frac{2}{b}}.
\end{equation}

Therefore, we obtain
\begin{equation}
\delta \leq \sqrt{\frac{2}{b}}.
\end{equation}

[1] E. Meckes. Concentration of measure and the compact classical matrix groups, https://www.math.ias.edu/files/wam/Haar_notes-revised.pdf, (2014).
[2] T. Nishiyama, A Tight Lower Bound for the Hellinger Distance with Given Means and Variances, arxiv:2010.13548 (2020).
[3] M. Tomamichel, A Framework for Non-Asymptotic Quantum Information Theory, PhD. Thesis, arXiv:1203.2142 (2012).
[4] G. Gour, R. W. Spekkens, The resource theory of quantum reference frames: manipulations and monotones, New J. Phys. 10, 033023 (2008).
[5] I. Marvian, R. W. Spekkens, The theory of manipulations of pure state asymmetry: basic tools and equivalence classes of states under symmetric operations, New J. Phys. 15, 033001 (2013).
[6] I. Marvian Symmetry, Asymmetry and Quantum Information, PhD thesis, (2012).
[7] M. Keyl and R.F. Werner, Optimal cloning of pure states, testing single clones, J.Math.Phys. 40 3283, (1999).
[8] M. Ozawa, Conservative quantum computing, Phys. Rev. Lett. 89, 057902 (2002).
[9] M. Ozawa, Uncertainty principle for quantum instruments and computing, Int. J. Quant. Inf. 1, 569 (2003).
[10] T. Karasawa and M. Ozawa, Conservation-law-induced quantum limits for physical realizations of the quantum not gate, Phys. Rev. A 75, 032324 (2007).

[11] T. Karasawa, J. Gea-Banacloche, M. Ozawa, Gate fidelity of arbitrary single-qubit gates constrained by conservation laws, J. Phys. A: math. Theor. 42, 225303 (2009).

[12] H. Tajima, N. Shiraishi and K. Saito, Uncertainty Relations in Implementation of Unitary Operations, Phys. Rev. Lett. 121, 110403 (2018).

[13] H. Tajima, N. Shiraishi and K. Saito, Coherence cost for violating conservation laws, Phys. Rev. Research. 2, 043374 (2020).

[14] P. Faist, S. Nezami, V. V. Albert, G. Salton, F. Pastawski, P. Hayden, and J. Preskill, Continuous Symmetries and Approximate Quantum Error Correction, Phys. Rev. X 10, 041018 (2020).