Supplementary Material: Simulating non-physical actions via exponentiation of Hermitian-preserving maps

CONTENTS

Appendix A: Preliminaries

1. Notations and definitions

In this material, we use letters $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{H}$, $\mathcal{K}$, and $\mathcal{R}$ to represent Hilbert spaces. The symbol $\mathcal{N}$ is used to denote the Hermitian-preserving map that we aim to exponentiate. The symbols $\mathcal{U}$, $\mathcal{Q}$, and $\mathcal{I}$ are used to represent the CPTP maps that are directly implementable on physical systems. Scratch letters like $\mathscr{H}$ and $\mathscr{F}$ are used to represent sets. $\mathbb{I}$ denotes the identity matrix. We denote $L(\mathcal{H})$ as the set of all linear operators on the Hilbert space $\mathcal{H}$, and $D(\mathcal{H})$ as the set of all density operators on $\mathcal{H}$. Furthermore, $T(\mathcal{X}, \mathcal{Y})$ represents the set of all linear maps from $L(\mathcal{X})$ to $L(\mathcal{Y})$. For a pure state $|\psi\rangle$, we sometimes omit the Dirac notation and denote the density matrix $|\psi\rangle\langle\psi|$ by ψ for brevity. For a unitary operator, $U \in L(\mathcal{H})$, we denote the corresponding unitary channel as $[U]$, which satisfies $[U](\sigma) = U\sigma U^\dagger$ for $\sigma \in D(\mathcal{H})$. The notation of $\log(\cdot)$ in this work represents the natural logarithm. For a Hermitian matrix A with spectral decomposition $A = \sum_i \lambda_i |\alpha_i\rangle\langle\alpha_i|$, we denote $A^+ := \sum_{i:\lambda_i>0} \lambda_i |\alpha_i\rangle\langle\alpha_i|$ as the positive part of A and $A^- := -\sum_{i:\lambda_i<0} \lambda_i |\alpha_i\rangle\langle\alpha_i|$ as the negative part of A.

When drawing tensor network diagrams, we use rounded rectangles to represent tensors and use lines of different colors to represent indices belonging to different systems. While drawing quantum circuits, we use regular rectangles to represent quantum operations and black lines to represent quantum systems.

We denote the operator norm as $\|\cdot\|_\infty$, which represents the maximum singular value of a matrix, and the trace norm as $\|\cdot\|_1$, which is the sum of singular values of a matrix. For two quantum states ρ and σ, the trace distance and fidelity between them are $\|\rho - \sigma\|_1$ and $F(\rho, \sigma) := \left(\operatorname{tr}\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right)^2$, respectively. For a Hermitian-preserving map $\mathcal{Q} \in T(\mathcal{X}, \mathcal{Y})$, its diamond norm is defined as $\|\mathcal{Q}\|_\diamond := \sup_{\sigma_{\mathcal{XR}} \in D(\mathcal{XR}), \mathcal{R}} \|(\mathcal{Q} \otimes \mathcal{I}_\mathcal{R})(\sigma_{\mathcal{XR}})\|_1$, where the symbol $\mathcal{XR}$ represents the Hilbert space $\mathcal{X} \otimes \mathcal{R}$, $\mathcal{I}_\mathcal{R}$ is the identity channel on $\mathcal{R}$, and the supremum is taken over the reference system $\mathcal{R}$ of an arbitrary dimension and all the states $\sigma_{\mathcal{XR}} \in D(\mathcal{XR})$. The diamond distance of two Hermitian-preserving maps $\mathcal{Q}_1$ and $\mathcal{Q}_2$ is defined as $\|\mathcal{Q}_1 - \mathcal{Q}_2\|_\diamond$.

2. Auxiliary lemmas

The essential property of the diamond distance is its subadditivity:

Lemma 1 (Proposition 3.48 in [1]). *Suppose $\mathcal{Q}_0, \mathcal{Q}_0' \in T(\mathcal{X}, \mathcal{Y})$ and $\mathcal{Q}_1, \mathcal{Q}_1' \in T(\mathcal{Y}, \mathcal{R})$ are CPTP maps, then the following inequality holds:*

$$\|\mathcal{Q}_1 \circ \mathcal{Q}_0 - \mathcal{Q}_1' \circ \mathcal{Q}_0'\|_\diamond \leq \|\mathcal{Q}_1 - \mathcal{Q}_1'\|_\diamond + \|\mathcal{Q}_0 - \mathcal{Q}_0'\|_\diamond. \tag{A1}$$

The following lemma provides an upper bound on the diamond distance between two channels corresponding to two Hamiltonian evolutions:

Lemma 2. *Let H_1 and H_2 be two Hermitian operators, and $t > 0$. Then, it holds that:*

$$\left\|\left[e^{-iH_1t}\right] - \left[e^{-iH_2t}\right]\right\|_\diamond \leq 2t\|H_1 - H_2\|_\infty \exp\left(t \max\left\{\|H_1\|_\infty, \|H_2\|_\infty\right\}\right). \tag{A2}$$

Proof. By Lemma 5 in the appendix of [2], we have

$$\left\|\left[e^{-iH_1t}\right] - \left[e^{-iH_2t}\right]\right\|_\diamond \leq 2\left\|e^{-iH_1t} - e^{-iH_2t}\right\|_\infty. \tag{A3}$$

Moreover,

$$
\begin{aligned}
&\left\|e^{-iH_1t}-e^{-iH_2t}\right\|_\infty\\
=&\left\|\sum_{m=0}^{\infty}\frac{1}{m!}(-iH_1t)^m-\sum_{m=0}^{\infty}\frac{1}{m!}(-iH_2t)^m\right\|_\infty\\
\leq&\sum_{m=1}^{\infty}\frac{1}{m!}t^m\|H_1^m-H_2^m\|_\infty\\
=&\sum_{m=1}^{\infty}\frac{1}{m!}t^m\left\|\sum_{l=0}^{m-1}H_1^{m-1-l}(H_1-H_2)H_2^l\right\|_\infty\\
\leq&\sum_{m=1}^{\infty}\frac{1}{m!}t^m\sum_{l=0}^{m-1}\|H_1\|_\infty^{m-1-l}\|H_1-H_2\|_\infty\|H_2\|_\infty^l\\
\leq& t\|H_1-H_2\|_\infty\sum_{m=1}^{\infty}\frac{t^{m-1}}{(m-1)!}\max\left\{\|H_1\|_\infty,\|H_2\|_\infty\right\}^{m-1}\\
=& t\|H_1-H_2\|_\infty\exp\Big(t\max\left\{\|H_1\|_\infty,\|H_2\|_\infty\right\}\Big).
\end{aligned}
\tag{A4}
$$

□

When measuring an observable over two quantum states, the difference in the expectation values can be upper bounded by applying the *Hölder inequality* [1] to $\left|\mathrm{Tr}\left(O(\sigma_1-\sigma_2)\right)\right|$:

Lemma 3. *Let $\sigma_1,\sigma_2\in D(\mathcal{H})$ be two quantum states, then the difference in the expectation values when measuring O satisfies*

$$
|\mathrm{Tr}(O\sigma_1)-\mathrm{Tr}(O\sigma_2)|\leq\|O\|_\infty\|\sigma_1-\sigma_2\|_1. \tag{A5}
$$

Appendix B: Hamiltonian in HME

1. Why partial transpose?

We note that the Hamiltonian used in HME is not the Choi matrix of $\mathcal{N}$, but the partial transposition of the Choi matrix. An intuitive way to grasp the role of the partial transposition operation in the Hamiltonian is by considering it as a result of concatenating the identity map. By rewriting $\mathcal{N}(\rho)=\mathcal{N}\circ\mathcal{I}(\rho)$, we can interpret HME as a process of first encoding ρ into the evolution using the identity map $\mathcal{I}$, and then applying the map $\mathcal{N}$ to it. Since the state-encoding procedure is equivalent to the density matrix exponentiation algorithm, the Hamiltonian for exponentiating the identity map is the SWAP operator, which can also be derived using Theorem 1 in the main context. Then, to perform the map $\mathcal{N}$ on the encoded state, we concatenate the Choi matrix of $\mathcal{N}$ to the SWAP operator, resulting in the final Hamiltonian of HME, $\Lambda_{\mathcal{N}}^{T_1}$. The concatenation with the SWAP operator results in the partial transposition of $\Lambda_{\mathcal{N}}$, which can be visualized using the red dashed box in Fig. 2(b) in the main context, where the cross of the black lines above $\Lambda_{\mathcal{N}}$ represents the SWAP operator.

2. Freedom of Hamiltonian in HME

As shown in Theorem 1 in the main context, $\Lambda_{\mathcal{N}}^{T_1}$ is a Hamiltonian that can be used to exponentiate the Hermitian-preserving map $\mathcal{N}\in T(\mathcal{H},\mathcal{K})$. In this section, we point out that $\Lambda_{\mathcal{N}}^{T_1}$ is not the unique choice. Specifically, we prove that all the Hamiltonians that can be used to exponentiate $\mathcal{N}$ form the set

$$
\mathscr{H}_{\mathcal{N}}=\left\{\Lambda_{\mathcal{N}}^{T_1}+M\otimes\mathbb{I}_{\mathcal{K}}:M\in\mathcal{L}(\mathcal{H}),M^\dagger=M\right\}. \tag{B1}
$$

Note that the equation $\mathrm{Tr}_1\left(e^{-iH\Delta t}(\rho\otimes\sigma)e^{iH\Delta t}\right)=e^{-i\mathcal{N}(\rho)\Delta t}\sigma e^{i\mathcal{N}(\rho)\Delta t}+\mathcal{O}(\Delta t^2)$ is equivalent to

$$
\mathrm{Tr}_1\left([H,\rho\otimes\sigma]\right)=[\mathcal{N}(\rho),\sigma]. \tag{B2}
$$

Thus, any H satisfying Eq. (B2) for all $\rho \in D(\mathcal{H})$ and $\sigma \in D(\mathcal{K})$ is a valid choice for exponentiating $\mathcal{N}$. It can be easily verified that any Hamiltonian in $\mathscr{H}_{\mathcal{N}}$ satisfies Eq. (B2) for all ρ and σ, as $\mathrm{Tr}_1\left([M \otimes \mathbb{I}_{\mathcal{K}}, \rho \otimes \sigma]\right) = 0$ always holds.

On the other hand, if H satisfies Eq. (B2) for all ρ and σ, we have

$$\mathrm{Tr}_1\left(\left[H - \Lambda_{\mathcal{N}}^{T_1}, \rho \otimes \sigma\right]\right) = 0,\ \forall \rho \in D(\mathcal{H}),\ \sigma \in D(\mathcal{K}). \tag{B3}$$

Since any matrix $A \in L(\mathcal{H} \otimes \mathcal{K})$ can be expressed as a complex linear combination of matrices of the form $\rho \otimes \sigma$, Eq. (B3) implies that

$$\mathrm{Tr}_1\left(\left[H - \Lambda_{\mathcal{N}}^{T_1}, A\right]\right) = 0,\ \forall A \in L(\mathcal{H} \otimes \mathcal{K}). \tag{B4}$$

Based on Lemma 4 presented below, we conclude that $H - \Lambda_{\mathcal{N}}^{T_1} = M \otimes \mathbb{I}_{\mathcal{K}}$ for some matrix $M \in L(\mathcal{H})$. As H should be a valid Hamiltonian, M must be a Hermitian matrix, which implies $H \in \mathscr{H}_{\mathcal{N}}$.

Lemma 4. *Suppose a matrix $B \in L(\mathcal{H} \otimes \mathcal{K})$ satisfies*

$$\mathrm{Tr}_1\left([B, A]\right) = 0,\ \forall A \in L(\mathcal{H} \otimes \mathcal{K}), \tag{B5}$$

where Tr_1 denotes the partial trace over the first system, $\mathcal{H}$. Then $B = M \otimes \mathbb{I}_{\mathcal{K}}$ for some matrix $M \in L(\mathcal{H})$.

Proof. For all $|i\rangle\langle j| \otimes |k\rangle\langle l| \in L(\mathcal{H} \otimes \mathcal{K})$, we have

$$\mathrm{Tr}_1\left(\left[B, |i\rangle\langle j| \otimes |k\rangle\langle l|\,\right]\right) = 0. \tag{B6}$$

Thus, for all i and j, we have

$$\left[\,\left(\langle j| \otimes \mathbb{I}_{\mathcal{K}}\right) B \left(|i\rangle \otimes \mathbb{I}_{\mathcal{K}}\right), |k\rangle\langle l|\,\right] = 0,\ \forall k, l, \tag{B7}$$

which implies that $\left(\langle j| \otimes \mathbb{I}_{\mathcal{K}}\right) B \left(|i\rangle \otimes \mathbb{I}_{\mathcal{K}}\right) = M_{ji}\mathbb{I}_{\mathcal{K}}$ for some constant $M_{ji} \in \mathbb{C}$. By defining $M \in L(\mathcal{H})$ as a matrix with elements in the j-th row and i-th column given by M_{ji}, we have $B = M \otimes \mathbb{I}_{\mathcal{K}}$. □

An open problem is whether the freedom in choosing the Hamiltonian can help to reduce the sample complexity of HME. Although for all Hermitian-preserving maps discussed in this paper, $\Lambda_{\mathcal{N}}^{T_1}$ is a suitable choice, it does not exclude the possibility that for certain Hermitian-preserving maps $\mathcal{N}$, there exist better choices within the set $\mathscr{H}_{\mathcal{N}}$ that could potentially reduce the sample complexity of realizing $e^{-i\mathcal{N}(\cdot)t}$.

Appendix C: Sample Complexity Upper Bound

1. Sample complexity upper bound of HME

Theorem 1 (Upper bound of sample complexity)**.** *Let $\mathcal{N}$ be an arbitrary Hermitian-preserving map, the HME algorithm shown in Fig. 1(c) in the main context requires at most $\mathcal{O}\left(\epsilon^{-1}\|H\|_\infty^2 t^2\right)$ copies of sequentially inputting state ρ to ensure that $\|\mathcal{Q}_t - \mathcal{U}_t\|_\diamond \le \epsilon$ holds for arbitrary ρ. Here, $H = \Lambda_{\mathcal{N}}^{T_1}$, $\mathcal{Q}_t = \mathcal{Q}_{\Delta t}^{\circ K}$ represents the real channel with $\mathcal{Q}_{\Delta t}(\sigma) := \mathrm{Tr}_1\left[e^{-iH\Delta t}(\rho \otimes \sigma)e^{iH\Delta t}\right]$ and $K = t/\Delta t$, $\mathcal{U}_t$ is the ideal evolution channel corresponding to $e^{-i\mathcal{N}(\rho)t}$, and $\|\cdot\|_\diamond$ denotes the diamond norm.*

Proof. As $\mathcal{Q}_t = \mathcal{Q}_{\Delta t}^{\circ K}$ and $\mathcal{U}_t = \mathcal{U}_{\Delta t}^{\circ K}$, we can use Lemma 1 iteratively to obtain

$$\|\mathcal{Q}_t - \mathcal{U}_t\|_\diamond = \left\|\mathcal{Q}_{\Delta t}^{\circ K} - \mathcal{U}_{\Delta t}^{\circ K}\right\|_\diamond \le K\|\mathcal{Q}_{\Delta t} - \mathcal{U}_{\Delta t}\|_\diamond. \tag{C1}$$

Therefore, the error only accumulates additively, and it is sufficient to analyze the error of a single step: $\|\mathcal{Q}_{\Delta t} - \mathcal{U}_{\Delta t}\|_\diamond$. Let $\mathcal{N} \in T(\mathcal{H}, \mathcal{K})$ be a Hermitian-preserving map from $L(\mathcal{H})$ to $L(\mathcal{K})$. Take arbitrary reference system $\mathcal{R}$ and arbitary

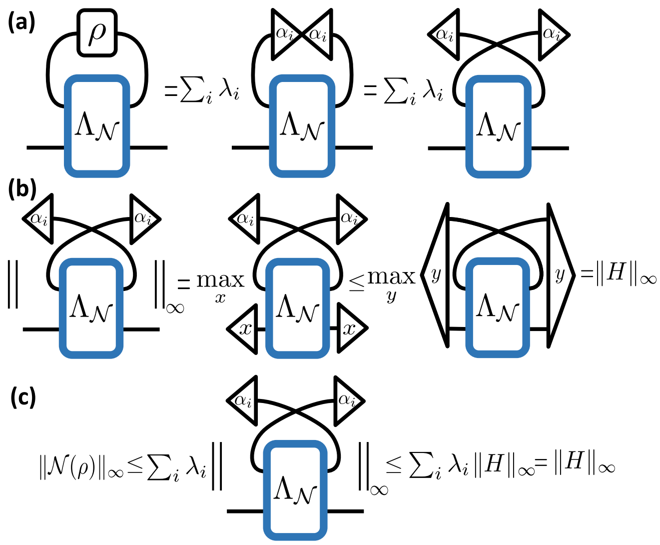

FIG. 1. Proof of $\|\mathcal{N}(\rho)\|_\infty \leq \|H\|_\infty$ using tensor diagrams. The triangles stand for vectors and rounded rectangles stand for matrices. (a) We decompose $\rho = \sum_i \lambda_i |\alpha_i\rangle\langle\alpha_i|$ and represent $\mathcal{N}(\rho)$ with $\Lambda_{\mathcal{N}}^{T_1}$, λ_i, and $|\alpha_i\rangle$ by dragging two triangles across each other. (b) Using the fact that $\|M\|_\infty = \max\{\langle s|\, M\, |s\rangle : s \in \mathcal{X}, \|s\| = 1\}$ for Hermitian matrix $M \in L(\mathcal{X})$, we derive $\|\mathcal{N}(|\alpha_i\rangle\langle\alpha_i|)\|_\infty \leq \|H\|_\infty$. (c) By applying the convexity of the operator norm, we prove that $\|\mathcal{N}(\rho)\|_\infty \leq \|H\|_\infty$.

$\sigma_{\mathcal{K}\mathcal{R}} \in D(\mathcal{K}\mathcal{R})$, we have

$$
\begin{aligned}
&(\mathcal{Q}_{\Delta t} \otimes \mathcal{I}_{\mathcal{R}})(\sigma_{\mathcal{K}\mathcal{R}}) \\
=&\operatorname{Tr}_1\left[(e^{-iH\Delta t} \otimes \mathbb{I}_{\mathcal{R}})(\rho \otimes \sigma_{\mathcal{K}\mathcal{R}})(e^{iH\Delta t} \otimes \mathbb{I}_{\mathcal{R}})\right] \\
=&\operatorname{Tr}_1\left[e^{-i(H\otimes\mathbb{I}_{\mathcal{R}})\Delta t}(\rho \otimes \sigma_{\mathcal{K}\mathcal{R}})e^{i(H\otimes\mathbb{I}_{\mathcal{R}})\Delta t}\right] \\
=&\operatorname{Tr}_1\left[\sum_{m=0}^{\infty} \frac{1}{m!}\operatorname{ad}^m_{(-i(H\otimes\mathbb{I}_{\mathcal{R}})\Delta t)}(\rho \otimes \sigma_{\mathcal{K}\mathcal{R}})\right] \\
=&\sigma_{\mathcal{K}\mathcal{R}} - i\Delta t\left[\mathcal{N}(\rho) \otimes \mathbb{I}_{\mathcal{R}}, \sigma_{\mathcal{K}\mathcal{R}}\right] \\
&+ \operatorname{Tr}_1\left[\sum_{m=2}^{\infty} \frac{1}{m!}\operatorname{ad}^m_{(-i(H\otimes\mathbb{I}_{\mathcal{R}})\Delta t)}(\rho \otimes \sigma_{\mathcal{K}\mathcal{R}})\right].
\end{aligned}
\tag{C2}
$$

Here the notation ad is defined as

$$
\begin{aligned}
\operatorname{ad}_A(B) &= [A, B], \\
\operatorname{ad}^m_A(B) &= [A, \operatorname{ad}^{m-1}_A(B)],
\end{aligned}
\tag{C3}
$$

for square matrices A and B of the same size. The trace norm of the third term in the last equation of Eq. (C2) satisfies

$$\begin{aligned}
&\left\|\mathrm{Tr}_1\left[\sum_{m=2}^{\infty}\frac{1}{m!}\,\mathrm{ad}^m_{(-i(H\otimes\mathbb{I}_{\mathcal{R}})\Delta t)}(\rho\otimes\sigma_{\mathcal{KR}})\right]\right\|_1\\
\leq&\left\|\sum_{m=2}^{\infty}\frac{1}{m!}\,\mathrm{ad}^m_{(-i(H\otimes\mathbb{I}_{\mathcal{R}})\Delta t)}(\rho\otimes\sigma_{\mathcal{KR}})\right\|_1\\
\leq&\sum_{m=2}^{\infty}\frac{1}{m!}\left\|\mathrm{ad}^m_{(-i(H\otimes\mathbb{I}_{\mathcal{R}})\Delta t)}(\rho\otimes\sigma_{\mathcal{KR}})\right\|_1\\
=&\sum_{m=2}^{\infty}\frac{1}{m!}(\Delta t)^m\left\|\mathrm{ad}^m_{(H\otimes\mathbb{I}_{\mathcal{R}})}(\rho\otimes\sigma_{\mathcal{KR}})\right\|_1\\
\leq&\sum_{m=2}^{\infty}\frac{1}{m!}(\Delta t)^m 2^m\|H\otimes\mathbb{I}_{\mathcal{R}}\|_\infty^m\|\rho\otimes\sigma_{\mathcal{KR}}\|_1\\
=&\sum_{m=2}^{\infty}\frac{1}{m!}(2\|H\|_\infty\Delta t)^m\\
=&e^{2\|H\|_\infty\Delta t}-1-2\|H\|_\infty\Delta t\\
\leq&4\|H\|_\infty^2\Delta t^2,\ \text{ when } \|H\|_\infty\Delta t\in(0,0.8].
\end{aligned}\tag{C4}$$

The first inequality in Eq. (C4) is due to the non-increasing property of trace norm under partial trace. The third inequality in Eq. (C4) is obtained by iteratively applying the inequalities $\|AB\|_1\leq\|A\|_\infty\|B\|_1$ and $\|AB\|_1\leq\|A\|_1\|B\|_\infty$, which hold for all square matrices A and B of the same size. The second equality in Eq. (C4) follows from $\|\rho\otimes\sigma_{\mathcal{KR}}\|_1=1$ and $\|H\otimes\mathbb{I}_{\mathcal{R}}\|_\infty=\|H\|_\infty$.

On the other hand, we have

$$\begin{aligned}
&(\mathcal{U}_{\Delta t}\otimes\mathcal{I}_{\mathcal{R}})(\sigma_{\mathcal{KR}})\\
=&(e^{-i\mathcal{N}(\rho)\Delta t}\otimes\mathbb{I}_{\mathcal{R}})(\sigma_{\mathcal{KR}})(e^{i\mathcal{N}(\rho)\Delta t}\otimes\mathbb{I}_{\mathcal{R}})\\
=&e^{-i(\mathcal{N}(\rho)\otimes\mathbb{I}_{\mathcal{R}})\Delta t}(\sigma_{\mathcal{KR}})e^{i(\mathcal{N}(\rho)\otimes\mathbb{I}_{\mathcal{R}})\Delta t}\\
=&\sum_{m=0}^{\infty}\frac{1}{m!}\,\mathrm{ad}^m_{(-i(\mathcal{N}(\rho)\otimes\mathbb{I}_{\mathcal{R}})\Delta t)}(\sigma_{\mathcal{KR}})\\
=&\sigma_{\mathcal{KR}}-i\Delta t\,[\mathcal{N}(\rho)\otimes\mathbb{I}_{\mathcal{R}},\sigma_{\mathcal{KR}}]\\
&\quad+\sum_{m=2}^{\infty}\frac{1}{m!}\,\mathrm{ad}^m_{(-i(\mathcal{N}(\rho)\otimes\mathbb{I}_{\mathcal{R}})\Delta t)}(\sigma_{\mathcal{KR}}).
\end{aligned}\tag{C5}$$

The trace norm of the third term in the last equation of Eq. (C5) satisfies

$$\begin{aligned}
&\left\|\sum_{m=2}^{\infty}\frac{1}{m!}\,\mathrm{ad}^m_{(-i(\mathcal{N}(\rho)\otimes\mathbb{I}_{\mathcal{R}})\Delta t)}(\sigma_{\mathcal{KR}})\right\|_1\\
\leq&\sum_{m=2}^{\infty}\frac{1}{m!}(\Delta t)^m\left\|\mathrm{ad}^m_{(\mathcal{N}(\rho)\otimes\mathbb{I}_{\mathcal{R}})}(\sigma_{\mathcal{KR}})\right\|_1\\
\leq&\sum_{m=2}^{\infty}\frac{1}{m!}(\Delta t)^m 2^m\|\mathcal{N}(\rho)\otimes\mathbb{I}_{\mathcal{R}}\|_\infty^m\|\sigma_{\mathcal{KR}}\|_1\\
=&\sum_{m=2}^{\infty}\frac{1}{m!}(2\|\mathcal{N}(\rho)\|_\infty\Delta t)^m\\
\leq&\sum_{m=2}^{\infty}\frac{1}{m!}(2\|H\|_\infty\Delta t)^m\\
\leq&4\|H\|_\infty^2\Delta t^2,\ \text{ when } \|H\|_\infty\Delta t\in(0,0.8],
\end{aligned}\tag{C6}$$

where the third inequality is given by $\|\mathcal{N}(\rho)\|_\infty \leq \|H\|_\infty$. To gain a better understanding of this inequality, we provide a graphical proof in Fig. 1. Combining Eq. (C2)-(C6), we can conclude that when $\|H\|_\infty \Delta t \in (0, 0.8]$, it always holds true that

$$\begin{aligned}
&\|(\mathcal{Q}_{\Delta t} \otimes \mathcal{I}_{\mathcal{R}})(\sigma_{\mathcal{KR}}) - (\mathcal{U}_{\Delta t} \otimes \mathcal{I}_{\mathcal{R}})(\sigma_{\mathcal{KR}})\|_1 \\
\leq& \left\| \mathrm{Tr}_1 \left[\sum_{m=2}^{\infty} \frac{1}{m!} \mathrm{ad}^m_{(-i(H\otimes \mathbb{I}_{\mathcal{R}})\Delta t)} (\rho \otimes \sigma_{\mathcal{KR}}) \right] \right\|_1 \\
&+ \left\| \sum_{m=2}^{\infty} \frac{1}{m!} \mathrm{ad}^m_{(-i(\mathcal{N}(\rho)\otimes \mathbb{I}_{\mathcal{R}})\Delta t)} (\sigma_{\mathcal{KR}}) \right\|_1 \\
\leq& 8\|H\|_\infty^2 \Delta t^2,
\end{aligned} \tag{C7}$$

regardless of the dimension of the reference system $\mathcal{R}$. By taking the supremum over $\mathcal{R}$ and $\sigma_{\mathcal{KR}}$, we have $\|\mathcal{Q}_{\Delta t} - \mathcal{U}_{\Delta t}\|_\diamond \leq 8\|H\|_\infty^2 \Delta t^2$. Substituting this result into Eq. (C1), when $K \in [1.25\|H\|_\infty t, \infty)$, we have

$$\|\mathcal{Q}_t - \mathcal{U}_t\|_\diamond \leq 8K\|H\|_\infty^2 \Delta t^2 = 8K^{-1}\|H\|_\infty^2 t^2. \tag{C8}$$

So, in order to guarantee $\|\mathcal{Q}_t - \mathcal{U}_t\|_\diamond \leq \epsilon$, it suffices to set

$$\begin{aligned}
K =& \max\left\{ \lceil 8\epsilon^{-1}\|H\|_\infty^2 t^2 \rceil, \lceil 1.25\|H\|_\infty t \rceil \right\} \\
\leq& \mathcal{O}\left(\epsilon^{-1}\|H\|_\infty^2 t^2 \right).
\end{aligned} \tag{C9}$$

□

By specializing $\mathcal{N} = \mathcal{I}$ and $H = S$, the sample complexity upper bound provided by Theorem 1 recovers the previously obtained upper bound of $\mathcal{O}\left(\epsilon^{-1}t^2\right)$ for density matrix exponentiation [3, 4], as $\|S\|_\infty = 1$. This upper bound has been proven to be tight [4]. While, $\mathcal{O}\left(\epsilon^{-1}\|H\|_\infty^2 t^2\right)$ may not be a tight bound in certain cases. For instance, when considering the partial transposition map, we observe that the sample complexity of exponentiating it with HME can be lower than what is predicted by Theorem 1, based on the specific properties of the partial transposition map. We provide the proof in the following section.

2. Sample complexity upper bound for exponentiating the partial transposition map

Proposition 1. *Setting the Hamiltonian to be $H_P = \Phi_A^+ \otimes S_B$, the HME algorithm requires at most $\mathcal{O}(\epsilon^{-1} d_A t^2)$ copies of sequentially inputting state ρ to ensure that*

$$\left\| \mathcal{Q}_t - [e^{-i\rho^{T_A} t}] \right\|_\diamond \leq \epsilon \tag{C10}$$

holds for arbitrary ρ.

To prove Proposition 1, we need to make some preliminary preparations. When $k \geq 1$, the expression for H_P^k can be written as:

$$H_P^k = \begin{cases} d_A^{k-1} \Phi_A^+ \otimes \mathbb{I}_B & \text{if } k \text{ even}, \\ d_A^{k-1} \Phi_A^+ \otimes S_B & \text{if } k \text{ odd}, \end{cases} \tag{C11}$$

where $\mathbb{I}_B$ represents the identity operator acting on two copies of the system B.

For two operators $P \in L(\mathcal{H}_1 \otimes \mathcal{H}_2)$ and $Q \in L(\mathcal{H}_2 \otimes \mathcal{H}_3)$, we define the *convolution* of P and Q as

$$P * Q := (\mathbb{I}_{\mathcal{H}_1} \otimes \langle \Phi^+ | \otimes \mathbb{I}_{\mathcal{H}_3})(P \otimes Q)(\mathbb{I}_{\mathcal{H}_1} \otimes |\Phi^+\rangle \otimes \mathbb{I}_{\mathcal{H}_3}), \tag{C12}$$

where $|\Phi^+\rangle$ represents the unnormalized maximally entangled state of $\mathcal{H}_2 \otimes \mathcal{H}_2$ in the computational basis. The following lemma provides an inequality concerning the norm of the convolution of two positive operators.

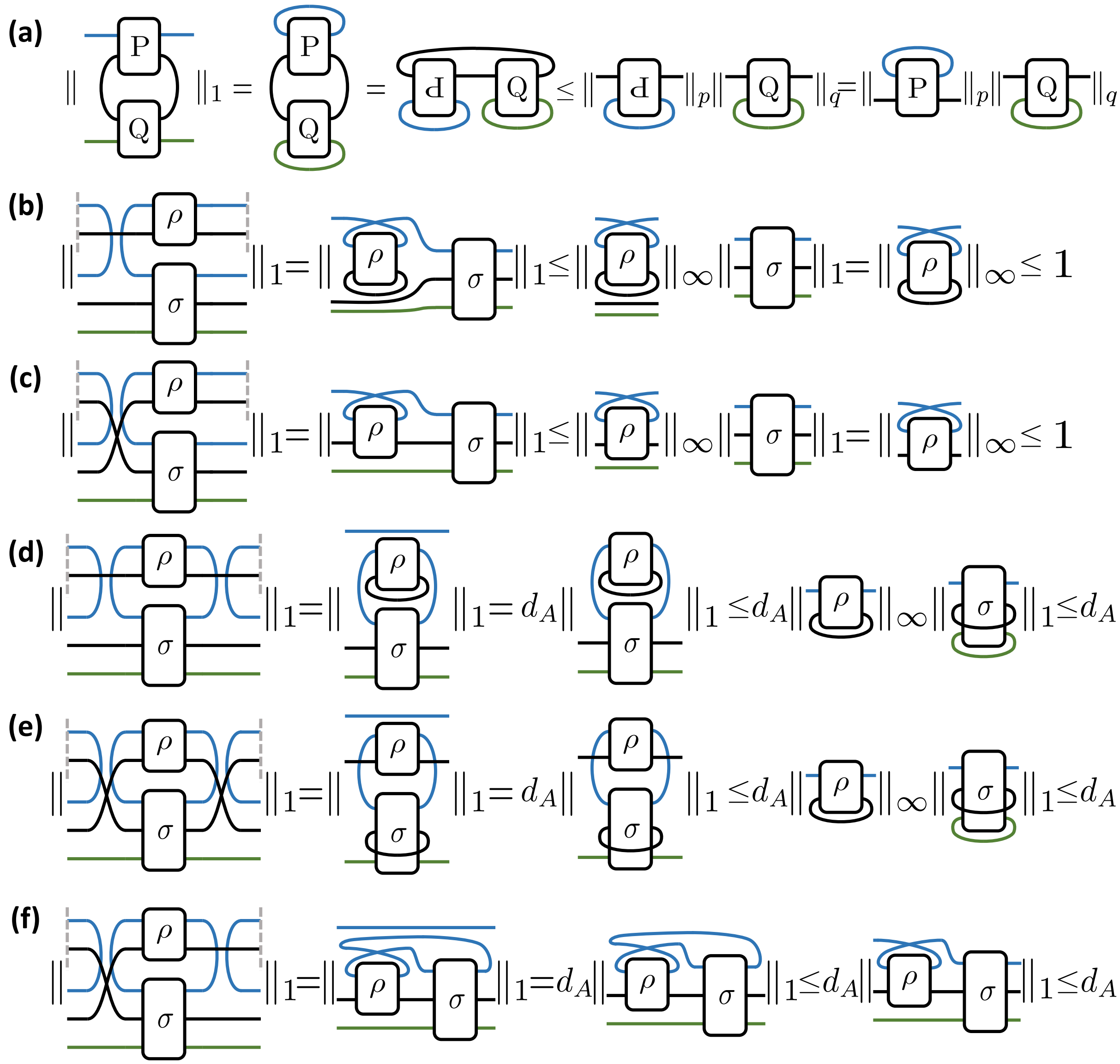

FIG. 2. (a) A tensor network version of Eq. (C14), where the inverted P box denotes the original P box after transposition and exchanging the upper and lower legs. The gray dashed lines denote the trace function, which is represented by connecting legs in tensor network diagrams. (b)-(f) are diagrams for the proof of Lemma 6, divided into five cases according to the values of k_1 and k_2. (b) k_1 even, $k_2 = 0$. The last inequality holds because $\mathrm{Tr}_2(\rho)^T$ is still a quantum state. (c) k_1 odd, $k_2 = 0$. The last inequality holds because $\|\rho^{T_A}\|_\infty \leq \sum_i \lambda_i \left\| |\alpha_i\rangle\langle\alpha_i|^{T_A} \right\|_\infty \leq \sum_i \lambda_i = 1$, where $\rho = \sum_i \lambda_i |\alpha_i\rangle\langle\alpha_i|$ is the spectral decomposition, and the eigenvalues of $|\alpha_i\rangle\langle\alpha_i|^{T_A}$ are in the range $[-1/2, 1]$ [5]. (d) k_1 even, k_2 even. The first inequality is given by Lemma 5. (e) k_1 odd, k_2 odd. The first inequality is given by Lemma 5. (f) k_1 odd, k_2 even. The first inequality is given by the non-increasing property of trace norm under partial trace. The last equality follows the derivation in (c). Other cases, such as '$k_1 = 0$, k_2 odd', can be derived by taking the Hermitian conjugate in the above derivations.

Lemma 5. *For positive operators $P \in L(\mathcal{H}_1 \otimes \mathcal{H}_2)$ and $Q \in L(\mathcal{H}_2 \otimes \mathcal{H}_3)$, the trace norm of $P * Q$ satisfies*

$$\|P * Q\|_1 \leq \|\mathrm{Tr}_{\mathcal{H}_1}(P)\|_p \|\mathrm{Tr}_{\mathcal{H}_3}(Q)\|_q, \tag{C13}$$

where $p, q \in [1, \infty]$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. $P, Q \geq 0$ implies $P * Q \geq 0$. We have

$$
\begin{aligned}
&\|P * Q\|_1 = \mathrm{Tr}_{\mathcal{H}_1,\mathcal{H}_3}(P * Q) = \mathrm{Tr}_{\mathcal{H}_1}(P) * \mathrm{Tr}_{\mathcal{H}_3}(Q) \\
=& \mathrm{Tr}\left(\mathrm{Tr}_{\mathcal{H}_1}(P)^T \, \mathrm{Tr}_{\mathcal{H}_3}(Q)\right) = \mathrm{Tr}\left(\overline{\mathrm{Tr}_{\mathcal{H}_1}(P)}^{\dagger} \, \mathrm{Tr}_{\mathcal{H}_3}(Q)\right) \\
\leq& \left\|\overline{\mathrm{Tr}_{\mathcal{H}_1}(P)}\right\|_p \|\mathrm{Tr}_{\mathcal{H}_3}(Q)\|_q = \|\mathrm{Tr}_{\mathcal{H}_1}(P)\|_p \|\mathrm{Tr}_{\mathcal{H}_3}(Q)\|_q,
\end{aligned}
\tag{C14}
$$

where $\overline{M}$ denotes the complex conjugate of an operator M. The first equality holds due to the positivity of $P * Q$. The inequality is given by the *Hölder inequality* [1]. A tensor network representation of Eq. (C14) is shown in Fig. 2(a). □

Taking advantage of the structure of H_P, we can provide a more refined estimation for the third term in the last equation of Eq. (C2).

Lemma 6. *For non-negative integers k_1 and k_2 such that $k_1 + k_2 \geq 1$, we have*

$$
\left\|\mathrm{Tr}_1\left[(H_P \otimes \mathbb{I}_{\mathcal{R}})^{k_1}(\rho \otimes \sigma_{\mathcal{KR}})(H_P \otimes \mathbb{I}_{\mathcal{R}})^{k_2}\right]\right\|_1 \leq d_A^{k_1+k_2-1}
\tag{C15}
$$

for any quantum states ρ and $\sigma_{\mathcal{KR}}$.

Proof. When k_1 and k_2 are odd, we have

$$
\begin{aligned}
&(H_P \otimes \mathbb{I}_{\mathcal{R}})^{k_1}(\rho \otimes \sigma_{\mathcal{KR}})(H_P \otimes \mathbb{I}_{\mathcal{R}})^{k_2} \\
=& d_A^{k_1+k_2-2}(\Phi_A^+ \otimes S_B \otimes \mathbb{I}_{\mathcal{R}})(\rho \otimes \sigma_{\mathcal{KR}})(\Phi_A^+ \otimes S_B \otimes \mathbb{I}_{\mathcal{R}}).
\end{aligned}
\tag{C16}
$$

Taking partial trace over the first system and estimating the trace norm, we have

$$
\begin{aligned}
&\left\|\mathrm{Tr}_1\left[(\Phi_A^+ \otimes S_B \otimes \mathbb{I}_{\mathcal{R}})(\rho \otimes \sigma_{\mathcal{KR}})(\Phi_A^+ \otimes S_B \otimes \mathbb{I}_{\mathcal{R}})\right]\right\|_1 \\
&= \left\|\mathbb{I}_{d_A} \otimes \left(S\rho S * \mathrm{Tr}_B(\sigma_{\mathcal{KR}})\right)\right\|_1 \\
&= d_A \|S\rho S * \mathrm{Tr}_B(\sigma_{\mathcal{KR}})\|_1 \\
&\leq d_A \|\mathrm{Tr}_B(\rho)\|_\infty \|\mathrm{Tr}_{B,\mathcal{R}}(\sigma)\|_1 \leq d_A.
\end{aligned}
\tag{C17}
$$

In the second and third lines of Eq. (C17), the symbol S denotes the SWAP operator between the two systems on which the density matrix ρ acts. The tensor network derivation of this expression can be found in Fig. 2(e). The proofs for other cases follow a similar approach and are provided in Fig. 2. □

Proof of Proposition 1. The trace norm of the third term in the last equation of Eq. (C2) can be estimated as

$$
\begin{aligned}
&\left\|\mathrm{Tr}_1\left[\sum_{m=2}^{\infty}\frac{1}{m!}\mathrm{ad}^m_{(-i(H_P\otimes\mathbb{I}_{\mathcal{R}})\Delta t)}(\rho \otimes \sigma_{\mathcal{KR}})\right]\right\|_1 \\
\leq& \sum_{m=2}^{\infty}\frac{1}{m!}(\Delta t)^m\left\|\mathrm{Tr}_1\left[\mathrm{ad}^m_{(H_P\otimes\mathbb{I}_{\mathcal{R}})}(\rho \otimes \sigma_{\mathcal{KR}})\right]\right\|_1 \\
\leq& \sum_{m=2}^{\infty}\frac{1}{m!}(2\Delta t)^m d_A^{m-1} \\
=& d_A^{-1}\sum_{m=2}^{\infty}\frac{1}{m!}(2d_A\Delta t)^m \\
=& \mathcal{O}(d_A \Delta t^2),
\end{aligned}
\tag{C18}
$$

where the second inequality is given by

$$
\left\|\mathrm{Tr}_1\left[(H_P \otimes \mathbb{I}_{\mathcal{R}})^{k}(\rho \otimes \sigma_{\mathcal{KR}})(H_P \otimes \mathbb{I}_{\mathcal{R}})^{m-k}\right]\right\|_1 \leq d_A^{m-1},
\tag{C19}
$$

which can be derived from Lemma 6 and holds for $0 \leq k \leq m$.

On the other hand, we have

$$\begin{aligned}&\left\|\sum_{m=2}^{\infty}\frac{1}{m!}\operatorname{ad}^m_{\left(-i\left(\rho^{T_A}\otimes\mathbb{I}_{\mathcal{R}}\right)\Delta t\right)}(\sigma_{\mathcal{KR}})\right\|_1\\ \leq&\mathcal{O}(\left\|\rho^{T_A}\right\|_\infty^2\Delta t^2)=\mathcal{O}(\Delta t^2).\end{aligned}\tag{C20}$$

Combining Eq. (C2) (C5) (C18) (C20), we can conclude that in order to guarantee $\left\|\mathcal{Q}_t-\left[e^{-i\rho^{T_A}t}\right]\right\|_\diamond\leq\epsilon$, the sample complexity K we need is at most $\mathcal{O}(\epsilon^{-1}d_At^2)$. □

Note that the proof of Proposition 1 can be easily generalized to the controlled version. Therefore, we need at most $\mathcal{O}(\epsilon^{-1}d_At^2)$ copies of ρ to implement controlled-$e^{-i\rho^{T_A}t}$ up to error ϵ in diamond distance.

Appendix D: Sample Complexity Lower Bound

1. Sample complexity lower bound of HME

Theorem 2 (Lower bound of sample complexity)**.** *Let $\mathcal{N}\in T(\mathcal{H},\mathcal{K})$ be a Hermitian-preserving map and set $0<\epsilon\leq 1/6$, and $t\geq\frac{15\pi\epsilon}{4R_*}$. The minimum number of ρ needed to realize the evolution of $e^{-i\mathcal{N}(\rho)t}$ using protocols shown in Fig. 3(a) in the main context with ϵ accuracy in diamond distance satisfies*

$$f_{\mathcal{N}}(\epsilon,t)\geq\Omega\left(\epsilon^{-1}R_*^2t^2\right),\tag{D1}$$

where $R_*:=\max_{A\in\mathscr{F}}R\left[\mathcal{N}(A)\right]$. $R[\cdot]=\lambda_{max}(\cdot)-\lambda_{min}(\cdot)$ denotes the spectral gap defined as the difference between the largest and the smallest eigenvalues of the processed matrix. The feasible region $\mathscr{F}$ is defined as

$$\mathscr{F}=\Big\{A\in L(\mathcal{H}):A^\dagger=A,\operatorname{Tr}(A)=0,\|A\|_1=1,\left[\mathcal{N}(A^+),\mathcal{N}(A^-)\right]=0\Big\},\tag{D2}$$

where A^+ and A^- are the positive and negative parts of A.

As an illustrative example, we consider the identity map $\mathcal{N}=\mathcal{I}$. According to the definition of the feasible region in Eq. (D2), it is straightforward to verify that $\frac{1}{2}(|\alpha\rangle\langle\alpha|-|\beta\rangle\langle\beta|)$, where $|\alpha\rangle$ and $|\beta\rangle$ are arbitrary orthonormal pure states, is a feasible point that maximizes $R[\mathcal{N}(A)]=R[A]$ with $R_*=1$. Thus, based on Theorem 2, the lower bound for realizing the evolution of $e^{-i\rho t}$ is $\Omega(\epsilon^{-1}t^2)$, which meets the upper bound of the complexity of HME in Theorem 1. This fact demonstrates that Theorem 2 covers the sample complexity lower bound given in Ref. [4].

We will utilize the following two lemmas in our proof.

Lemma 7 (Holevo-Helstrom theorem [6, 7])**.** *Given two states, $\rho_0,\rho_1\in D(\mathcal{H})$, for any two-outcomes POVM $\{M_0,M_1\}$, we have*

$$\frac{1}{2}\operatorname{Tr}(M_0\rho_0)+\frac{1}{2}\operatorname{Tr}(M_1\rho_1)\leq\frac{1}{2}+\frac{1}{4}\|\rho_0-\rho_1\|_1.\tag{D3}$$

Furthermore, equality can be achieved by an appropriate choice of M_0,M_1.

Lemma 8 (Channel analogue of Holevo-Helstrom theorem [1])**.** *Let $\mathcal{Q}_0,\mathcal{Q}_1\in T(\mathcal{X},\mathcal{Y})$ be two CPTP maps. For any ancillary system $\mathcal{R}$, any two-outcomes POVM $\{M_0,M_1\}$ on $\mathcal{Y}\otimes\mathcal{R}$, and any input state $\sigma\in D(\mathcal{X}\otimes\mathcal{R})$, we have*

$$\frac{1}{2}\operatorname{Tr}\left(M_0\left(\mathcal{Q}_0\otimes\mathcal{I}_{\mathcal{R}}\right)(\sigma)\right)+\frac{1}{2}\operatorname{Tr}\left(M_1\left(\mathcal{Q}_1\otimes\mathcal{I}_{\mathcal{R}}\right)(\sigma)\right)\leq\frac{1}{2}+\frac{1}{4}\|\mathcal{Q}_0-\mathcal{Q}_1\|_\diamond.\tag{D4}$$

Furthermore, if $\dim(\mathcal{R})\geq\dim(\mathcal{X})$, the equality can be achieved by an appropriate choice of σ and $\{M_0,M_1\}$.

Proof of Theorem 2. First, let us choose two states given by

$$\rho_\pm:=|A|\pm\lambda A=(1\pm\lambda)A^++(1\mp\lambda)A^-,\tag{D5}$$

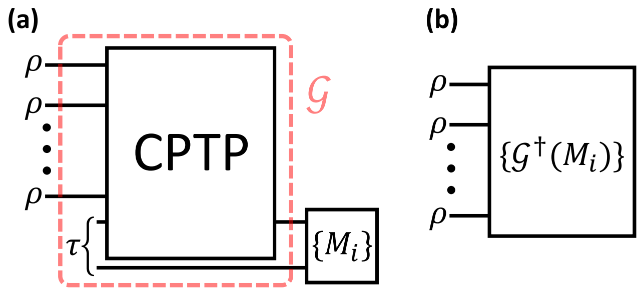

FIG. 3. (a) Multiple copies of ρ together with a CPTP map induce a channel $\mathcal{Q}^{\rho}$ on the last input state. Here we demonstrate the optimal method to distinguish two different induced channels $\mathcal{Q}^{\rho_\pm}$, which is composed of first acting $\mathcal{Q}^{\rho}$ on a subsystem of a larger processed state τ and measuring it using a two-outcome POVM measurement. (b) The channel discrimination task shown in (a) is equivalent to a state discrimination task, as the CPTP map, the processed state τ, and the POVM measurement $\{M_i\}$ can be considered as a large POVM $\{\mathcal{G}^\dagger(M_i)\}$ acting on $\rho_\pm^{\otimes K}$. Note that $\mathcal{G}^\dagger$ is a completely positive unital map, thus $\{\mathcal{G}^\dagger(M_i)\}$ is a valid POVM.

where A is an element of the feasible region $\mathscr{F}$ and $0<\lambda<1$. Note that $\|A\|_1=1$ and $\mathrm{Tr}(A)=0$ imply $\mathrm{Tr}(\rho_\pm)=1$, and $0<\lambda<1$ implies $\rho_\pm\geq 0$, together ensuring that $\rho_\pm$ are valid quantum states. Denoting $A=\sum_x a_x |\alpha_x\rangle\!\langle\alpha_x|$ as the spectral decomposition of A, we have

$$\rho_\pm=\sum_{x:a_x>0}(1\pm\lambda)a_x|\alpha_x\rangle\!\langle\alpha_x|+\sum_{x:a_x<0}(1\mp\lambda)(-a_x)|\alpha_x\rangle\!\langle\alpha_x|. \tag{D6}$$

Thus, the fidelity between ρ_+ and ρ_- is given by

$$\begin{aligned} F(\rho_+,\rho_-)&=\left(\sum_x\sqrt{(1+\lambda)(1-\lambda)a_x^2}\right)^2\\ &=\left(1-\lambda^2\right)\left(\sum_x|a_x|\right)^2=1-\lambda^2. \end{aligned} \tag{D7}$$

Denote $R=R[\mathcal{N}(A)]$ for brevity. Take $t_\lambda=\frac{\pi}{2\lambda R}$. We can compute the diamond norm distance as

$$\begin{aligned} &\left\|\left[e^{-i\mathcal{N}(\rho_+)t_\lambda}\right]-\left[e^{-i\mathcal{N}(\rho_-)t_\lambda}\right]\right\|_\diamond\\ =&\max_{\sigma\in D(\mathcal{KR}),\mathcal{R}}\left\|\left(\left[e^{-i\mathcal{N}(\rho_+)t_\lambda}\right]-\left[e^{-i\mathcal{N}(\rho_-)t_\lambda}\right]\right)\otimes\mathcal{I}_{\mathcal{R}}(\sigma)\right\|_1\\ =&\max_{|u\rangle}\left\|\left[e^{-i\mathcal{N}(\rho_+)t_\lambda}\right]|u\rangle\!\langle u|-\left[e^{-i\mathcal{N}(\rho_-)t_\lambda}\right]|u\rangle\!\langle u|\right\|_1\\ =&\max_{|u\rangle}2\left(1-\left|\langle u|e^{i\mathcal{N}(\rho_+)t_\lambda}e^{-i\mathcal{N}(\rho_-)t_\lambda}|u\rangle\right|^2\right)^{\frac{1}{2}}\\ =&2\left(1-\min_{|u\rangle}\left|\langle u|e^{i\pi\frac{\mathcal{N}(A)}{R}}|u\rangle\right|^2\right)^{\frac{1}{2}}=2. \end{aligned} \tag{D8}$$

The second equality of Eq. (D8) holds because $\left[e^{-i\mathcal{N}(\rho_+)t_\lambda}\right]$ and $\left[e^{-i\mathcal{N}(\rho_-)t_\lambda}\right]$ are unitary channels. The fourth equality of Eq. (D8) holds because

$$[\mathcal{N}(\rho_+),\mathcal{N}(\rho_-)]=4\lambda\left[\mathcal{N}(A^+),\mathcal{N}(A^-)\right]=0, \tag{D9}$$

where the second equality follows from the fact that $A\in\mathscr{F}$. By setting $|u\rangle=\frac{1}{\sqrt{2}}(|L\rangle+|S\rangle)$, where $|L\rangle$ and $|S\rangle$ are the eigenstates corresponding to the largest and smallest eigenvalues of $\mathcal{N}(A)$, we can demonstrate that $\left|\langle u|e^{i\pi\frac{\mathcal{N}(A)}{R}}|u\rangle\right|^2=0$. Hence, the last equality of Eq. (D8) holds.

Now for the non-physical map $\mathcal{N}$, desired accuracy $1/3$, and evolution time t_λ, suppose there exists a CPTP map in the general protocol shown in Fig. 3(a) int the main text with K copies of the inputting state ρ, such that for all

$\rho \in D(\mathcal{H})$, the diamond distance between the induced channel $\mathcal{Q}^{\rho}$ and the target unitary channel $\left[e^{-i\mathcal{N}(\rho)t_\lambda}\right]$ satisfies $\left\|Q^{\rho} - \left[e^{-i\mathcal{N}(\rho)t_\lambda}\right]\right\|_\diamond \leq 1/3$. By selecting $\rho_\pm$ as the inputting states for the general protocol, the diamond distance between the channels induced by them satisfies

$$\begin{aligned}\|\mathcal{Q}^{\rho_+} - \mathcal{Q}^{\rho_-}\|_\diamond \geq & \left\|\left[e^{-i\mathcal{N}(\rho_+)t_\lambda}\right] - \left[e^{-i\mathcal{N}(\rho_-)t_\lambda}\right]\right\|_\diamond \\ & - \left\|Q^{\rho_+} - \left[e^{-i\mathcal{N}(\rho_+)t_\lambda}\right]\right\|_\diamond \\ & - \left\|Q^{\rho_-} - \left[e^{-i\mathcal{N}(\rho_-)t_\lambda}\right]\right\|_\diamond \geq \frac{4}{3}.\end{aligned} \tag{D10}$$

According to Lemma 8, if we take an ancillary system $\mathcal{R}$ such that $\dim(\mathcal{R}) = \dim(\mathcal{K})$, then there exist a state $\tau \in D(\mathcal{K} \otimes \mathcal{R})$ and a POVM $\{M_0, M_1\}$ on $L(\mathcal{K} \otimes \mathcal{R})$ such that

$$\begin{aligned}&\frac{1}{2}\operatorname{Tr}\left(M_0\left(\mathcal{Q}^{\rho_+} \otimes \mathcal{I}_\mathcal{R}\right)(\tau)\right) + \frac{1}{2}\operatorname{Tr}\left(M_1\left(\mathcal{Q}^{\rho_-} \otimes \mathcal{I}_\mathcal{R}\right)(\tau)\right) \\ &\qquad = \frac{1}{2} + \frac{1}{4}\|\mathcal{Q}^{\rho_+} - \mathcal{Q}^{\rho_-}\|_\diamond \geq \frac{5}{6}.\end{aligned} \tag{D11}$$

Therefore, it is possible to distinguish between $\mathcal{Q}^{\rho_+}$ and $\mathcal{Q}^{\rho_-}$, thus also $\rho_\pm$, with a success probability of at least $5/6$ using just a single copy of $\mathcal{Q}^{\rho_+}$ or $\mathcal{Q}^{\rho_-}$.

On the other hand, as shown in Fig. 3(b), the CPTP channel, the processed state τ, and the POVM operators can be unified into a global POVM measurement acting on the multiple input states ρ, labeled as $\{\mathcal{G}^\dagger(M_i)\}_{i=0,1}$, where $\mathcal{G}$ represents a new CPTP channel. Thus

$$\operatorname{Tr}\left(M_i\left(\mathcal{Q}^{\rho} \otimes \mathcal{I}_\mathcal{R}\right)(\tau)\right) = \operatorname{Tr}\left(\mathcal{G}^\dagger\left(M_i\right)\rho^{\otimes K}\right) \tag{D12}$$

holds for $i = 0, 1$ and all $\rho \in D(\mathcal{H})$. Combining Lemma 7, Eq. (D11), and Eq. (D12), we obtain

$$\begin{aligned}\frac{5}{6} &\leq \frac{1}{2}\operatorname{Tr}\left(\mathcal{G}^\dagger\left(M_0\right)\rho_+^{\otimes K}\right) + \frac{1}{2}\operatorname{Tr}\left(\mathcal{G}^\dagger\left(M_1\right)\rho_-^{\otimes K}\right) \\ &\leq \frac{1}{2} + \frac{1}{4}\left\|\rho_+^{\otimes K} - \rho_-^{\otimes K}\right\|_1.\end{aligned} \tag{D13}$$

This implies that

$$\begin{aligned}\frac{4}{3} &\leq \left\|\rho_+^{\otimes K} - \rho_-^{\otimes K}\right\|_1 \leq 2\sqrt{1 - F\left(\rho_+^{\otimes K}, \rho_-^{\otimes K}\right)} \\ &= 2\sqrt{1 - F\left(\rho_+, \rho_-\right)^K} = 2\sqrt{1 - \left(1 - \lambda^2\right)^K},\end{aligned} \tag{D14}$$

where the second inequality follows from the relation between trace distance and fidelity, and the first equality is based on the property of fidelity [8]. We thus have

$$K \geq \log(1.8)\left(\log\frac{1}{1-\lambda^2}\right)^{-1} \geq \frac{\log(1.8)}{2}\frac{1}{\lambda^2} = \frac{2\log(1.8)}{\pi^2}R^2 t_\lambda^2, \tag{D15}$$

where the second inequality holds for all $\lambda \in (0, 0.8]$. Denote the constant as $C = 2\log(1.8)/\pi^2$, by adjusting the value of λ within the range of $(0, 0.8]$, Eq. (D15) implies that for all $t \in [\frac{5\pi}{8R}, \infty)$, we have

$$f_\mathcal{N}(1/3, t) \geq CR^2t^2. \tag{D16}$$

To derive the lower bound for $f_\mathcal{N}(\epsilon, t)$, we need to explore the properties of this function. Suppose a global circuit shown in Fig. 3(a) in the main context can realize $[e^{-i\mathcal{N}(\rho)t}]$ with ϵ accuracy, meaning that $\left\|\mathcal{Q}^{\rho} - [e^{-i\mathcal{N}(\rho)t}]\right\|_\diamond \leq \epsilon$. According to the subadditivity of diamond distance, if we concatenate this global circuit for m times, we can realize the evolution of $[e^{-i\mathcal{N}(\rho)mt}]$ with accuracy $m\epsilon$. This implies an important property of $f_\mathcal{N}(\epsilon, t)$:

$$f_\mathcal{N}(m\epsilon, mt) \leq m f_\mathcal{N}(\epsilon, t). \tag{D17}$$

For all $\epsilon \in (0, \frac{1}{6}]$ and $t \in [\frac{15\pi\epsilon}{4R}, \infty)$, we take $m = \lceil \frac{1}{6\epsilon} \rceil$, ensuring $m\epsilon \leq \frac{1}{3}$ and $mt \geq \frac{5\pi}{8R}$. Thus,

$$\begin{aligned}f_\mathcal{N}(\epsilon, t) \geq & \frac{1}{m}f_\mathcal{N}\left(m\epsilon, mt\right) \geq \frac{1}{m}f_\mathcal{N}\left(1/3, mt\right) \\ \geq & \frac{1}{m}CR^2(mt)^2 = CmR^2t^2 \geq \frac{C}{6}\epsilon^{-1}R^2t^2,\end{aligned} \tag{D18}$$

where the third inequality is based on Eq. (D16). The above equation holds for all $A \in \mathscr{F}$. By taking the maximum value over the feasible region, we define

$$R_* := \max_{A \in \mathscr{F}} R\left[\mathcal{N}(A)\right], \tag{D19}$$

and obtain

$$f_{\mathcal{N}}(\epsilon, t) \geq \frac{C}{6}\epsilon^{-1} R_*^2 t^2 \geq \Omega\left(\epsilon^{-1} R_*^2 t^2\right), \tag{D20}$$

which holds for all $\epsilon \in (0, \frac{1}{6}]$ and all $t \in [\frac{15\pi\epsilon}{4R_*}, \infty)$. $\square$

It can be observed that Theorem 2 holds significance only if the feasible region $\mathscr{F}$ exists. The following proposition presents a sufficient condition for the non-emptiness of $\mathscr{F}$.

Proposition 2. *For a Hermitian-preserving map $\mathcal{N} \in T(\mathcal{H}, \mathcal{K})$, suppose there exists a non-zero Hermitian matrix $P \in L(\mathcal{H})$ such that $\mathcal{N}(P)$ commutes with all matrices in the set $\left\{\mathcal{N}(M) : M \in L(\mathcal{H}), M^\dagger = M\right\}$, then the feasible region $\mathscr{F}$ defined in Theorem 2 is non-empty.*

Proof. If $\operatorname{rank} P = 1$, we can select a positive matrix $Q \in L(\mathcal{H})$ whose non-zero eigenstates are orthogonal to the non-zero eigenstate of P, one can check that

$$A = \frac{|P|}{2\|P\|_1} - \frac{Q}{2\|Q\|_1} \tag{D21}$$

is an element of $\mathscr{F}$.

For the case where $\operatorname{rank} P > 1$ and $P \geq 0$ ($P \leq 0$), we can decompose $P = P_1 + P_2$ using two non-zero matrices P_1 and P_2, where $P_1, P_2 \geq 0$ ($P_1, P_2 \leq 0$) and have orthogonal support. One can check that

$$A = \frac{|P_1|}{2\|P_1\|_1} - \frac{|P_2|}{2\|P_2\|_1} \tag{D22}$$

is an element of $\mathscr{F}$.

If both P^+ and P^-, the positive and negative parts of P, are non-zero, then one can check that

$$A = \frac{P^+}{2\|P^+\|_1} - \frac{P^-}{2\|P^-\|_1} \tag{D23}$$

is an element of $\mathscr{F}$. $\square$

Consider the following two easily satisfied conditions:

1. $\mathcal{N}$ is not injective,
2. $\dim(\mathcal{H}) \geq \dim(\mathcal{K})$.

If one of these conditions is satisfied, then the condition required by Proposition 2 is also satisfied, implying that $\mathscr{F}$ is non-empty. When $\mathcal{N}$ satisfies condition 1, we can choose a non-zero Hermitian matrix $P \in L(\mathcal{H})$ such that $\mathcal{N}(P) = 0$. As a result, $\mathcal{N}(P)$ commutes with all matrices in $L(\mathcal{K})$. On the other hand, when $\mathcal{N}$ satisfies condition 2, we only need to consider the case when $\dim(\mathcal{H}) = \dim(\mathcal{K})$ and $\mathcal{N}$ is injective, because otherwise $\mathcal{N}$ is not injective and satisfies condition 1. In this case $\mathcal{N}$ is also surjective, so we can choose a non-zero Hermitian matrix $P \in L(\mathcal{H})$ such that $\mathcal{N}(P) = \mathbb{I}_{\mathcal{K}}$, and thus $\mathcal{N}(P)$ commutes with all matrices in $L(\mathcal{K})$.

Note that when $\mathscr{F}$ is empty, Theorem 2 will yield the trivial lower bound $f_{\mathcal{N}}(\epsilon, t) \geq 0$. However, for Hermitian-preserving maps discussed in this paper, Theorem 2 can always provide meaningful lower bounds. The most challenging condition to fulfill required by the feasible region $\mathscr{F}$ is $[\mathcal{N}(A^+), \mathcal{N}(A^-)] = 0$. This condition is intended to facilitate the computation of $\langle u| e^{i\mathcal{N}(\rho_+)t_\lambda} e^{-i\mathcal{N}(\rho_-)t_\lambda} |u\rangle$ in Eq. (D8). Alternatively, more powerful mathematical tools, such as the Baker-Campbell-Hausdorff formula, can be employed to compute this quantity. Utilizing such tools has the potential to eliminate the constraint $[\mathcal{N}(A^+), \mathcal{N}(A^-)] = 0$ from the feasible region $\mathscr{F}$ and derive improved lower bounds for $f_{\mathcal{N}}(\epsilon, t)$.

2. Example: inverting local amplitude damping noise channel

We now calculate the lower bound on the sample complexity for exponentiating the inverse channel of local amplitude damping noise. Our analysis demonstrates that even with highly global protocols, mitigating the effects of local amplitude damping noise requires an exponential sample complexity. The lower bound presented in this subsection and the upper bound provided by Theorem 1 are asymptotically consistent for all variables: the desired accuracy ϵ, the evolution time t, the number of qubits n and the noise intensity γ. This result demonstrates that HME is an asymptotically optimal protocol for exponentiating certain maps.

Let $\mathcal{E}_\gamma$ denote the single-qubit amplitude damping noise with a damping rate $0 < \gamma < 1$. It can be expressed as $\mathcal{E}_\gamma(\sigma) = E_0\sigma E_0^\dagger + E_1\sigma E_1^\dagger$, where

$$\begin{aligned} E_0 &= |0\rangle\langle 0| + \sqrt{1-\gamma}\, |1\rangle\langle 1|\,, \\ E_1 &= \sqrt{\gamma}\, |0\rangle\langle 1|\,. \end{aligned} \tag{D24}$$

The inverse of $\mathcal{E}_\gamma$ acts as

$$\mathcal{E}_\gamma^{-1}(\sigma) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{1-\gamma}} \end{pmatrix} \sigma \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{1-\gamma}} \end{pmatrix} + \frac{-\gamma}{1-\gamma} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \sigma \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{D25}$$

Denote $\mathcal{N}_{\gamma,n} := \left(\mathcal{E}_\gamma^{\otimes n}\right)^{-1} = \left(\mathcal{E}_\gamma^{-1}\right)^{\otimes n}$ as the inverse of the n-qubit local amplitude damping channel we aim to exponentiate.

To compute the lower bound for exponentiating $\mathcal{N}_{\gamma,n}$, we need to find an element in its feasible region. We observe that

$$\begin{aligned} \mathcal{N}_{\gamma,n}(|0\rangle\langle 0|^{\otimes n}) &= |0\rangle\langle 0|^{\otimes n}\,, \\ \mathcal{N}_{\gamma,n}(|1\rangle\langle 1|^{\otimes n}) &= \left(\frac{-\gamma}{1-\gamma}\, |0\rangle\langle 0| + \frac{1}{1-\gamma}\, |1\rangle\langle 1|\right)^{\otimes n}, \end{aligned} \tag{D26}$$

which implies that $\left[\mathcal{N}_{\gamma,n}(|0\rangle\langle 0|^{\otimes n}), \mathcal{N}_{\gamma,n}(|1\rangle\langle 1|^{\otimes n})\right] = 0$. One can show that $A_n := \frac{1}{2}\,|1\rangle\langle 1|^{\otimes n} - \frac{1}{2}\,|0\rangle\langle 0|^{\otimes n}$ is an element of the feasible region $\mathscr{F}$. Acting $\mathcal{N}_{\gamma,n}$ on A_n, we obtain

$$\begin{aligned} \mathcal{N}_{\gamma,n}\left(A_n\right) =& \frac{1}{2} \sum_{s\in\{0,1\}^n} \left(\frac{-\gamma}{1-\gamma}\right)^{n-|s|} \left(\frac{1}{1-\gamma}\right)^{|s|} |s\rangle\langle s| - \frac{1}{2}\, |0\rangle\langle 0|^{\otimes n} \\ =& \frac{1}{2}\left(\frac{1}{1-\gamma}\right)^n |1\rangle\langle 1|^{\otimes n} + \left[\frac{1}{2}\left(\frac{-\gamma}{1-\gamma}\right)^n - \frac{1}{2}\right] |0\rangle\langle 0|^{\otimes n} + \cdots\,, \end{aligned} \tag{D27}$$

where $|s|$ denote the number of 1's in the string s. Since $0 < \gamma < 1$, we have

$$\begin{aligned} R_* \geq& R\left[\mathcal{N}_{\gamma,n}\left(A_n\right)\right] \\ \geq& \frac{1}{2}\left(\frac{1}{1-\gamma}\right)^n - \left[\frac{1}{2}\left(\frac{-\gamma}{1-\gamma}\right)^n - \frac{1}{2}\right] \\ \geq& \Omega\left(\left(\frac{1}{1-\gamma}\right)^n\right), \end{aligned} \tag{D28}$$

which implies

$$f_{\mathcal{N}_{\gamma,n}}(\epsilon, t) \geq \Omega\left(\epsilon^{-1}\left(\frac{1}{1-\gamma}\right)^{2n} t^2\right). \tag{D29}$$

Through some straightforward calculations, we find that

$$\left\|\Lambda_{\mathcal{N}_{\gamma,n}}^{T_1}\right\|_\infty = \left(\frac{1}{1-\gamma}\right)^n. \tag{D30}$$

Therefore, according to Theorem 1, HME is a protocol with sample complexity $\mathcal{O}\left(\epsilon^{-1}\left(\frac{1}{1-\gamma}\right)^{2n} t^2\right)$.

Appendix E: Robustness

1. Robustness of HME

In practical situations, the performance of HME can be affected by various factors. For example, quantum devices may exhibit unpredictable and unavoidable noise, and the realization of the Hamiltonian evolution $e^{-iH\Delta t}$ may be inaccurate due to approximation errors in Hamiltonian simulation [9, 10]. Hence, the robustness of HME is another important indicator that needs in-depth analysis. We mainly consider two kinds of errors, the input state error, and the Hamiltonian error. These errors can lead to biases and variations in the input states and the evolution Hamiltonians at each step. The error of HME caused by these factors can be bounded by:

Theorem 3 (Robustness)**.** *Suppose the input states and Hamiltonians of HME vary at each step, denoted as $\{\rho'_k\}_{k=1}^K$ and $\{H'_k\}_{k=1}^K$. If the number of steps satisfies $K \geq 2t\|H\|_\infty$ and $K \geq 2t\|H'_k\|_\infty$ for $k = 1, \cdots, K$, then the diamond distance between the noisy channel $\mathcal{Q}'_t$ constructed using $\{\rho'_k\}_{k=1}^K$, and $\{H'_k\}_{k=1}^K$ and the noiseless channel $\mathcal{Q}_t$ constructed using fixed ρ and H, is bounded by*

$$\|\mathcal{Q}'_t - \mathcal{Q}_t\|_\diamond \leq 4t\left(D_H + \|H\|_\infty D_S\right), \tag{E1}$$

where $D_H := \frac{1}{K}\sum_{k=1}^K \|H'_k - H\|_\infty$ and $D_S := \frac{1}{K}\sum_{k=1}^K \|\rho'_k - \rho\|_1$ are the average divergences between the Hamiltonians and input states, respectively.

In practice, the difference between $\|H\|_\infty$ and $\|H'_k\|_\infty$ is relatively small and the condition $\|H\|_\infty t \geq 1$ is usually satisfied. Thus, if we choose $K = \Theta\left(\epsilon^{-1}\|H\|_\infty^2 t^2\right)$, as requires in Theorem 1, the requirement of $K \geq 2t\|H\|_\infty$ and $K \geq 2t\|H'_k\|_\infty$ for $k = 1, \cdots, K$ can be naturally satisfied. In addition to the diversities of Hamiltonians and input states, the error also depends on the norm of the evolution Hamiltonian. According to Eq. (E1), $\|H\|_\infty$ acts as a multiplier that amplifies the error caused by the diversity of input states. Similar to the sample complexity upper bound, we can derive a better estimation of the robustness of certain maps by leveraging their specific properties. For instance, in the case of the partial transposition map, the robustness is not related to $\|H_P\|_\infty$. We prove this in the next section.

Proof. Likewise the proof of Theorem 1 in Appendix C 1, to derive the distance between the whole channel $\|\mathcal{Q}'_t - \mathcal{Q}_t\|$, we need to first derive the distance of a step $\left\|\mathcal{Q}'_{k,\Delta t} - \mathcal{Q}_{\Delta t}\right\|$, where

$$\mathcal{Q}'_{k,\Delta t}(\cdot) := \mathrm{Tr}_1\left[(e^{-iH'_k\Delta t}\otimes \mathbb{I}_\mathcal{R})(\rho'_k \otimes \cdot)(e^{iH'_k\Delta t}\otimes \mathbb{I}_\mathcal{R})\right] \tag{E2}$$

with H'_k and ρ'_k varying for each step, and

$$\mathcal{Q}'_t = \mathcal{Q}'_{K,\Delta t}\circ \mathcal{Q}'_{K-1,\Delta t}\circ\cdots\circ \mathcal{Q}'_{1,\Delta t}. \tag{E3}$$

Then, for arbitrary reference system $\mathcal{R}$ and any state $\sigma_{\mathcal{KR}} \in D(\mathcal{KR})$, we have

$$\begin{aligned}
&\left\|\mathcal{Q}'_{k,\Delta t}\otimes \mathcal{I}_\mathcal{R}(\sigma_{\mathcal{KR}}) - \mathcal{Q}_{\Delta t}\otimes \mathcal{I}_\mathcal{R}(\sigma_{\mathcal{KR}})\right\|_1\\
=&\left\|\mathrm{Tr}_1\left[(e^{-iH'_k\Delta t}\otimes \mathbb{I}_\mathcal{R})(\rho'_k\otimes\sigma_{\mathcal{KR}})(e^{iH'_k\Delta t}\otimes \mathbb{I}_\mathcal{R})\right] - \mathrm{Tr}_1\left[(e^{-iH\Delta t}\otimes \mathbb{I}_\mathcal{R})(\rho\otimes\sigma_{\mathcal{KR}})(e^{iH\Delta t}\otimes \mathbb{I}_\mathcal{R})\right]\right\|_1\\
\leq&\left\|\mathrm{Tr}_1\left[(e^{-iH'_k\Delta t}\otimes \mathbb{I}_\mathcal{R})(\rho'_k\otimes\sigma_{\mathcal{KR}})(e^{iH'_k\Delta t}\otimes \mathbb{I}_\mathcal{R})\right] - \mathrm{Tr}_1\left[(e^{-iH\Delta t}\otimes \mathbb{I}_\mathcal{R})(\rho'_k\otimes\sigma_{\mathcal{KR}})(e^{iH\Delta t}\otimes \mathbb{I}_\mathcal{R})\right]\right\|_1\\
&+\left\|\mathrm{Tr}_1\left[(e^{-iH\Delta t}\otimes \mathbb{I}_\mathcal{R})(\rho'_k\otimes\sigma_{\mathcal{KR}})(e^{iH\Delta t}\otimes \mathbb{I}_\mathcal{R})\right] - \mathrm{Tr}_1\left[(e^{-iH\Delta t}\otimes \mathbb{I}_\mathcal{R})(\rho\otimes\sigma_{\mathcal{KR}})(e^{iH\Delta t}\otimes \mathbb{I}_\mathcal{R})\right]\right\|_1\\
\leq&\left\|(e^{-iH'_k\Delta t}\otimes \mathbb{I}_\mathcal{R})(\rho'_k\otimes\sigma_{\mathcal{KR}})(e^{iH'_k\Delta t}\otimes \mathbb{I}_\mathcal{R}) - (e^{-iH\Delta t}\otimes \mathbb{I}_\mathcal{R})(\rho'_k\otimes\sigma_{\mathcal{KR}})(e^{iH\Delta t}\otimes \mathbb{I}_\mathcal{R})\right\|_1\\
&+\left\|\mathrm{Tr}_1\left[(e^{-iH\Delta t}\otimes \mathbb{I}_\mathcal{R})\left((\rho'_k-\rho)\otimes\sigma_{\mathcal{KR}}\right)(e^{iH\Delta t}\otimes \mathbb{I}_\mathcal{R})\right]\right\|_1\\
\leq&\left\|[e^{-iH'_k\Delta t}] - [e^{-iH\Delta t}]\right\|_\diamond + \left\|\mathrm{Tr}_1\left[(e^{-iH\Delta t}\otimes \mathbb{I}_\mathcal{R})\left((\rho'_k-\rho)\otimes\sigma_{\mathcal{KR}}\right)(e^{iH\Delta t}\otimes \mathbb{I}_\mathcal{R})\right]\right\|_1,
\end{aligned} \tag{E4}$$

where the last inequality is given by the definition of the diamond norm. Consider the first term of the last line, we have

$$\begin{aligned}
&\left\|[e^{-iH'_k\Delta t}]-[e^{-iH\Delta t}]\right\|_\diamond\\
\leq&2\Delta t\|H'_k-H\|_\infty\exp\left(\Delta t\max\left\{\|H'_k\|_\infty,\|H\|_\infty\right\}\right)\\
\leq&2\sqrt{e}\Delta t\|H'_k-H\|_\infty,
\end{aligned}\tag{E5}$$

where the first inequality is given by Lemma 2, and the second inequality holds when

$$\Delta t\max\left\{\|H'_k\|_\infty,\|H\|_\infty\right\}\in(0,0.5].\tag{E6}$$

The second term of the last equation of Eq. (E4) satisfies

$$\begin{aligned}
&\left\|\operatorname{Tr}_1\left[(e^{-iH\Delta t}\otimes\mathbb{I}_{\mathcal{R}})\left((\rho'_k-\rho)\otimes\sigma_{\mathcal{KR}}\right)(e^{iH\Delta t}\otimes\mathbb{I}_{\mathcal{R}})\right]\right\|_1\\
=&\left\|\sum_{m=0}^{\infty}\frac{1}{m!}\operatorname{Tr}_1\left[\operatorname{ad}^m_{(-i(H\otimes\mathbb{I}_{\mathcal{R}})\Delta t)}((\rho'_k-\rho)\otimes\sigma_{\mathcal{KR}})\right]\right\|_1\\
\leq&\|\operatorname{Tr}_1\left[(\rho'_k-\rho)\otimes\sigma_{\mathcal{KR}})\right]\|_1+\sum_{m=1}^{\infty}\frac{1}{m!}\left\|\operatorname{ad}^m_{(-i(H\otimes\mathbb{I}_{\mathcal{R}})\Delta t)}((\rho'_k-\rho)\otimes\sigma_{\mathcal{KR}})\right\|_1\\
\leq&\sum_{m=1}^{\infty}\frac{1}{m!}(2\Delta t\|H\|_\infty)^m\|(\rho'_k-\rho)\otimes\sigma_{\mathcal{KR}}\|_1\\
=&\|\rho'_k-\rho\|_1\left(e^{2\Delta t\|H\|_\infty}-1\right)\\
\leq&4\Delta t\|H\|_\infty\|\rho'_k-\rho\|_1,
\end{aligned}\tag{E7}$$

where the last inequality holds when $\Delta t\|H\|_\infty\in(0,0.5]$.

Combining Eq. (E4) (E5) (E7), we can conclude that

$$\begin{aligned}
&\left\|\mathcal{Q}'_{k,\Delta t}\otimes\mathcal{I}_{\mathcal{R}}(\sigma_{\mathcal{KR}})-\mathcal{Q}_{\Delta t}\otimes\mathcal{I}_{\mathcal{R}}(\sigma_{\mathcal{KR}})\right\|_1\\
\leq&\Delta t\left(2\sqrt{e}\|H'_k-H\|_\infty+4\|H\|_\infty\|\rho'_k-\rho\|_1\right),
\end{aligned}\tag{E8}$$

which holds for all choices of $\mathcal{R}$ and $\sigma_{\mathcal{KR}}\in D(\mathcal{KR})$. Thus we have

$$\begin{aligned}
&\left\|\mathcal{Q}'_{k,\Delta t}-\mathcal{Q}_{\Delta t}\right\|_\diamond\\
\leq&\Delta t\left(2\sqrt{e}\|H'_k-H\|_\infty+4\|H\|_\infty\|\rho'_k-\rho\|_1\right).
\end{aligned}\tag{E9}$$

By the subadditivity property of diamond distance (Lemma 1), we have

$$\begin{aligned}
\|\mathcal{Q}'_t-\mathcal{Q}_t\|_\diamond\leq&\sum_{k=1}^{K}\left\|\mathcal{Q}'_{k,\Delta t}-\mathcal{Q}_{\Delta t}\right\|_\diamond\\
\leq&t\left(2\sqrt{e}\frac{1}{K}\sum_{k=1}^{K}\|H'_k-H\|_\infty+4\|H\|_\infty\frac{1}{K}\sum_{k=1}^{K}\|\rho'_k-\rho\|_1\right)\\
\leq&4t\left(D_H+\|H\|_\infty D_S\right).
\end{aligned}\tag{E10}$$

According to the derivation, the requirement for the validity of Eq. (E10) is that

$$\frac{t}{K}\max\left\{\|H\|_\infty,\|H'_1\|_\infty,\cdots,\|H'_K\|_\infty\right\}\leq0.5,\tag{E11}$$

that is, $K\geq2t\|H\|_\infty$ and $K\geq2t\|H'_k\|_\infty$ for $k=1,\cdots,K$. $\square$

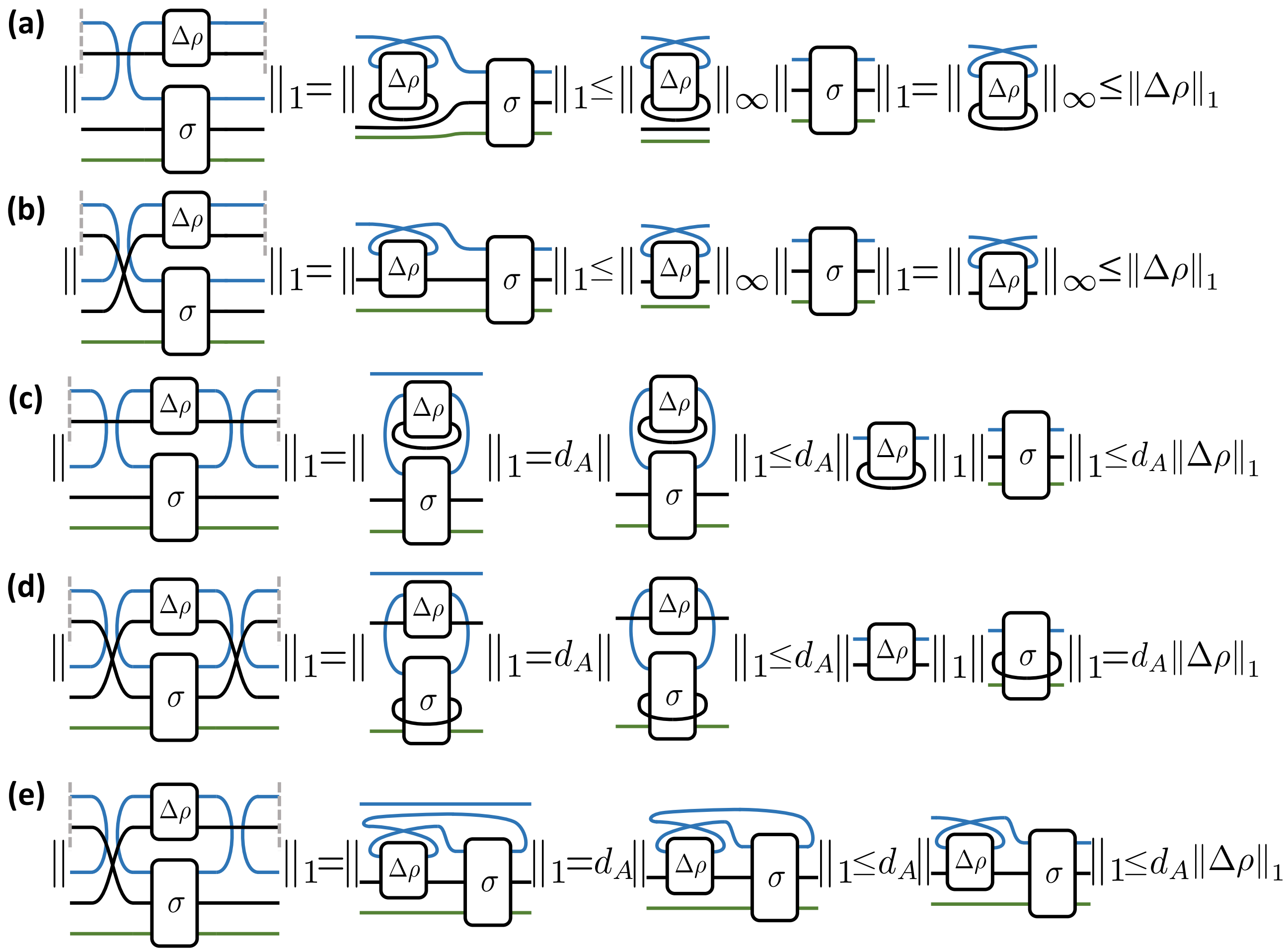

FIG. 4. Diagrams for the proof of Lemma 10, divided into five cases according to the values of k_1 and k_2. We denote $\rho_1 - \rho_2$ by $\Delta\rho$. Other cases, such as '$k_1 = 0$, k_2 odd' can be derived by taking the Hermitian conjugate of the derivations listed in the figure. (a) k_1 even, $k_2 = 0$. The last inequality is given by $\left\|\mathrm{Tr}_2(\Delta\rho)^T\right\|_\infty = \|\mathrm{Tr}_2(\Delta\rho)\|_\infty \le \|\mathrm{Tr}_2(\Delta\rho)\|_1 \le \|\Delta\rho\|_1$. (b) k_1 odd, $k_2 = 0$. The last inequality holds because $\left\|(\Delta\rho)^{T_A}\right\|_\infty \le \sum_i |\lambda_i| \left\| |\alpha_i\rangle\langle\alpha_i|^{T_A} \right\|_\infty \le \sum_i |\lambda_i| = \|\Delta\rho\|_1$, where $\Delta\rho = \sum_i \lambda_i |\alpha_i\rangle\langle\alpha_i|$ is the spectral decomposition of $\Delta\rho$. (c) k_1 even, k_2 even. The first inequality is given by Lemma 9. (d) k_1 odd, k_2 odd. The inequality is given by Lemma 9. (f) k_1 odd, k_2 even. The last inequality follows the derivation in (b).

2. Robustness for exponentiating the partial transposition map

For certain specific Hermitian-preserving maps, we can obtain more accurate estimations of the robustness of HME. Again we use the partial transposition map as an example to illustrate this.

Proposition 3. *Setting the Hamiltonian to be $H_P = \Phi_A^+ \otimes S_B$, following the noise assumptions in Theorem 3, we have*

$$\|\mathcal{Q}_t' - \mathcal{Q}_t\|_\diamond \le 4t\,(D_H + D_S)\,. \tag{E12}$$

To prove this proposition, we first need to present two lemmas, similar to what we have done in Appendix C 2.

Lemma 9. *For $P \in L(\mathcal{H}_1 \otimes \mathcal{H}_2)$ and $Q \in L(\mathcal{H}_2 \otimes \mathcal{H}_3)$, we have*

$$\|P * Q\|_1 \le \|P\|_1 \|Q\|_1. \tag{E13}$$

Proof. Let $P = \sum_i \lambda_i |\alpha_i\rangle\langle\alpha_i|$ and $Q = \sum_i \mu_i |\beta_i\rangle\langle\beta_i|$ be the spectral decomposition of P and Q respectively. We have

$$\begin{aligned}\|P * Q\|_1 &\leq \sum_{i,j} |\lambda_i| \cdot |\mu_j| \cdot \| |\alpha_i\rangle\langle\alpha_i| * |\beta_j\rangle\langle\beta_j| \|_1 \\ &\leq \sum_{i,j} |\lambda_i| \cdot |\mu_j| = \|P\|_1 \|Q\|_1,\end{aligned} \tag{E14}$$

where the second inequality can be derived by Lemma 5. □

Lemma 10. *For non-negative integers k_1 and k_2 such that $k_1 + k_2 \geq 1$, we have*

$$\begin{aligned}&\left\| \mathrm{Tr}_1 \left[(H_P \otimes \mathbb{I}_{\mathcal{R}})^{k_1} \left((\rho_1 - \rho_2) \otimes \sigma_{\mathcal{KR}} \right) (H_P \otimes \mathbb{I}_{\mathcal{R}})^{k_2}) \right] \right\|_1 \\ &\leq d_A^{k_1 + k_2 - 1} \|\rho_1 - \rho_2\|_1\end{aligned} \tag{E15}$$

for any quantum states ρ_1, ρ_2 and $\sigma_{\mathcal{KR}}$.

Proof. Denote $\rho_1 - \rho_2$ by $\Delta\rho$ for brevity. By Eq. (C11), when k_1 and k_2 are odd, we have

$$\begin{aligned}&(H_P \otimes \mathbb{I}_{\mathcal{R}})^{k_1} \left(\Delta\rho \otimes \sigma_{\mathcal{KR}} \right) (H_P \otimes \mathbb{I}_{\mathcal{R}})^{k_2} \\ =& d_A^{k_1 + k_2 - 2} (\Phi_A^+ \otimes S_B \otimes \mathbb{I}_{\mathcal{R}}) (\Delta\rho \otimes \sigma_{\mathcal{KR}}) (\Phi_A^+ \otimes S_B \otimes \mathbb{I}_{\mathcal{R}}).\end{aligned} \tag{E16}$$

Taking partial trace on the first system and estimating the trace norm, we have

$$\begin{aligned}&\left\| \mathrm{Tr}_1 \left[(\Phi_A^+ \otimes S_B \otimes \mathbb{I}_{\mathcal{R}}) (\Delta\rho \otimes \sigma_{\mathcal{KR}}) (\Phi_A^+ \otimes S_B \otimes \mathbb{I}_{\mathcal{R}}) \right] \right\|_1 \\ &= \|\mathbb{I}_{d_A} \otimes [S\Delta\rho S * \mathrm{Tr}_B(\sigma_{\mathcal{KR}})]\|_1 \\ &= d_A \|[S\Delta\rho S * \mathrm{Tr}_B(\sigma_{\mathcal{KR}})]\|_1 \\ &\leq d_A \|S\Delta\rho S\|_1 \|\mathrm{Tr}_B(\sigma)\|_1 \leq d_A \|\Delta\rho\|_1,\end{aligned} \tag{E17}$$

where the symbol S denotes the SWAP operator between the two systems on which the density matrix ρ acts. The tensor network derivation of this expression can be found in Fig. 4(d). The proofs of the other cases follow a similar approach and are presented in Fig. 4. □

Proof of Proposition 3. The second term of the last equation in Eq. (E4) satisfies

$$\begin{aligned}&\left\| \mathrm{Tr}_1 \left[(e^{-iH_P \Delta t} \otimes \mathbb{I}_{\mathcal{R}}) \left((\rho_k' - \rho) \otimes \sigma_{\mathcal{KR}} \right) (e^{iH_P \Delta t} \otimes \mathbb{I}_{\mathcal{R}}) \right] \right\|_1 \\ &\leq \sum_{m=1}^{\infty} \frac{1}{m!} \left\| \mathrm{Tr}_1 \left[\mathrm{ad}^m_{(-i(H_P \otimes \mathbb{I}_{\mathcal{R}})\Delta t)} ((\rho_k' - \rho) \otimes \sigma_{\mathcal{KR}}) \right] \right\|_1 \\ &\leq \sum_{m=1}^{\infty} \frac{1}{m!} (\Delta t)^m \left\| \mathrm{Tr}_1 \left[\mathrm{ad}^m_{(H_P \otimes \mathbb{I}_{\mathcal{R}})} ((\rho_k' - \rho) \otimes \sigma_{\mathcal{KR}}) \right] \right\|_1 \\ &\leq \sum_{m=1}^{\infty} \frac{1}{m!} (2\Delta t)^m d_A^{m-1} \|\rho_k' - \rho\|_1 \\ &= \frac{1}{d_A} \|\rho_k' - \rho\|_1 \left(e^{2 d_A \Delta t} - 1 \right) \\ &\leq 4\Delta t \|\rho_k' - \rho\|_1,\end{aligned} \tag{E18}$$

where the third inequality is given by Lemma 10, and the last inequality holds when $d_A \Delta t \in (0, 0.5]$.

Combining Eq. (E4), (E5) and (E18), we obtain

$$\begin{aligned}&\left\| \mathcal{Q}'_{k,\Delta t} \otimes \mathcal{I}_{\mathcal{R}}(\sigma_{\mathcal{KR}}) - \mathcal{Q}_{\Delta t} \otimes \mathcal{I}_{\mathcal{R}}(\sigma_{\mathcal{KR}}) \right\|_1 \\ \leq& \Delta t \left(2\sqrt{e} \|H_k' - H\|_\infty + 4\|\rho_k' - \rho\|_1 \right),\end{aligned} \tag{E19}$$

which holds for arbitrary $\mathcal{R}$ and $\sigma_{\mathcal{KR}} \in D(\mathcal{KR})$. Thus by the definition and the subadditivity property of diamond distance, we have

$$\begin{aligned}\left\|\mathcal{Q}_t' - \mathcal{Q}_t\right\|_\diamond \leq& \sum_{k=1}^{K} \left\|\mathcal{Q}_{k,\Delta t}' - \mathcal{Q}_{\Delta t}\right\|_\diamond \\ \leq& t\left(2\sqrt{e}\frac{1}{K}\sum_{k=1}^{K}\left\|H_k' - H\right\|_\infty + 4\frac{1}{K}\sum_{k=1}^{K}\left\|\rho_k' - \rho\right\|_1\right) \\ \leq& 4t\left(D_H + D_S\right).\end{aligned} \tag{E20}$$

□

Appendix F: Entanglement Detection

1. A brief introduction to quantum phase estimation

To facilitate our subsequent discussions, we now give a brief introduction of the functions and properties of the quantum phase estimation algorithm, which is employed to analyze the eigenvalue, $e^{i\varphi}$, and eigenstate of a given unitary U [8, 11]. As depicted in Fig. 4(a) in the main context, the number of ancilla qubits is determined by the desired estimation accuracy of target eigenvalues. While in some special cases when the binary representation of $\varphi/2\pi$ has a finite length, a finite number of ancilla qubits can estimate φ accurately. If a sufficient number of ancilla qubits is employed and the input state σ is one of the eigenstates of U, measuring the ancilla qubits will yield the corresponding eigenvalue of σ. While, if σ is not an eigenstate of U, the measurement of the ancilla qubits will randomly output one of the eigenvalues of U and σ will collapse into the corresponding eigenstate, assuming no degeneracy. The probability of obtaining some eigenvalue is proportional to the fidelity between σ and the corresponding eigenstate.

2. Constant upper bound for the HME-based protocol

Theorem 4. *Let $H_R = \mathbb{I}_A \otimes S_B - S_{AB}$, where $\mathbb{I}_A$, S_B, and S_{AB} stand for the identity operator acting on two copies of system A, SWAP operator acting on two copies of system B, and SWAP operator acting on two copies of the joint system AB, respectively. Choosing $t = \pi$, the circuit depicted in Fig. 4(b) in the main context requires at most $\mathcal{O}(1)$ copies of ρ to accomplish the task described in Fact 1 in the main context with a success probability of at least $2/3$.*

Proof. Let $\mathscr{P}_0 = \{\psi_A \otimes \psi_B\}$ and $\mathscr{P}_1 = \{\psi_{AB}\}$ be the sets of product pure states and global pure states, respectively. An unknown state ρ is selected through the following procedure: first, one randomly chooses $\mathscr{P}_0$ or $\mathscr{P}_1$ with equal probabilities, and then randomly selects a state from the chosen set according to the Haar measure. Our task is to determine which set the unknown state belongs to, using multiple copies of this state.

To accomplish this state discrimination task, we utilize $K+1$ copies of ρ to exponentiate the circuit depicted in Fig. 4(b) in the main context for K sequential steps. Let $\mathcal{Q}^\rho$ be the channel constructed by HME to approximate $\left[\text{C-}e^{-i\rho^{R_A}\pi}\right]$. By using $K = \mathcal{O}(1)$ copies of ρ, we ensure that $\left\|\mathcal{Q}^\rho - \left[\text{C-}e^{-i\rho^{R_A}\pi}\right]\right\|_\diamond \leq 1/12$. Upon measuring the ancilla qubit, if the outcome is $|i\rangle$ $(i = 0, 1)$, we infer that ρ belongs to the set $\mathscr{P}_i$.

Let us first consider the case when $\rho = \psi_{AB} \in \mathscr{P}_1$. Define

$$\begin{aligned}\Sigma :=& \left[\text{C-}e^{-i(-\psi_{AB})\pi}\right]\left(|+\rangle\langle+| \otimes \psi_{AB}\right), \\ \Sigma' :=& \mathcal{Q}^{\psi_{AB}}\left(|+\rangle\langle+| \otimes \psi_{AB}\right).\end{aligned} \tag{F1}$$

As $e^{i\psi_{AB}\pi} = \mathbb{I} - 2\psi_{AB}$, Σ can be written in a block matrix form as

$$\begin{aligned}\Sigma =& \left[\text{C-}(\mathbb{I} - 2\psi_{AB})\right]\left(|+\rangle\langle+| \otimes \psi_{AB}\right) \\ =& \frac{1}{2}\begin{bmatrix}\psi_{AB} & -\psi_{AB} \\ -\psi_{AB} & \psi_{AB}\end{bmatrix}.\end{aligned} \tag{F2}$$

Upon measuring the ancilla qubit of Σ in the Pauli-X basis, we have

$$\operatorname{Tr}\left[(|-\rangle\langle-| \otimes \mathbb{I}_d)\Sigma\right] = \operatorname{Tr}(\psi_{AB}) = 1, \tag{F3}$$

This indicates that we will certainly obtain the measurement result $|1\rangle$.

By Lemma 2, we have

$$\begin{aligned}
&\left\|\left[\text{C-}e^{-i(\psi_{AB})^{R_A}\pi}\right]-\left[\text{C-}e^{i\psi_{AB}\pi}\right]\right\|_\diamond\\
=&\left\|\left[e^{-i\left(|1\rangle\langle 1|\otimes(\psi_{AB})^{R_A}\right)\pi}\right]-\left[e^{i(|1\rangle\langle 1|\otimes\psi_{AB})\pi}\right]\right\|_\diamond\\
\leq&2\pi\left\|(\psi_{AB})^{R_A}+\psi_{AB}\right\|_\infty\\
&\exp\left(\pi\max\left\{\left\|(\psi_{AB})^{R_A}\right\|_\infty,\|\psi_{AB}\|_\infty\right\}\right)\\
\leq&2\pi e^{\pi}\|\mathbb{I}_A\otimes\mathrm{Tr}_A(\psi_{AB})-\psi_{AB}+\psi_{AB}\|_\infty\\
=&2\pi e^{\pi}\|\mathrm{Tr}_A(\psi_{AB})\|_\infty,
\end{aligned}\tag{F4}$$

where the second inequality follows from

$$\begin{aligned}
&\left\|(\psi_{AB})^{R_A}\right\|_\infty=\|\mathbb{I}_A\otimes\mathrm{Tr}_A(\psi_{AB})-\psi_{AB}\|_\infty\\
\leq&\max\left\{\|\mathbb{I}_A\otimes\mathrm{Tr}_A(\psi_{AB})\|_\infty,\|\psi_{AB}\|_\infty\right\}=1.
\end{aligned}\tag{F5}$$

The average purity of the reduced density matrix $\mathrm{Tr}_A(\psi_{AB})$ is known to satisfy [12]

$$\mathop{\mathbb{E}}_{\psi_{AB}\sim\text{Haar}}\mathrm{Tr}\left[\mathrm{Tr}_A(\psi_{AB})^2\right]=\frac{d_A+d_B}{d_Ad_B+1}=\frac{2\sqrt{d}}{d+1},\tag{F6}$$

where the last equality holds because we take $d_A=d_B=\sqrt{d}$ for simplicity. Note that

$$\begin{aligned}
\mathop{\mathbb{E}}_{\psi_{AB}\sim\text{Haar}}\|\mathrm{Tr}_A(\psi_{AB})\|_\infty\leq&\mathop{\mathbb{E}}_{\psi_{AB}\sim\text{Haar}}\sqrt{\mathrm{Tr}\left[\mathrm{Tr}_A(\psi_{AB})^2\right]}\\
\leq&\sqrt{\mathop{\mathbb{E}}_{\psi_{AB}\sim\text{Haar}}\mathrm{Tr}\left[\mathrm{Tr}_A(\psi_{AB})^2\right]},
\end{aligned}\tag{F7}$$

where the second inequality follows from Jensen's inequality. Thus, we have

$$\mathop{\mathbb{E}}_{\psi_{AB}\sim\text{Haar}}\|\mathrm{Tr}_A(\psi_{AB})\|_\infty\to 0,\ \text{as } d\to\infty.\tag{F8}$$

When d is sufficiently large, by Markov's inequality, we can guarantee $\|\mathrm{Tr}_A(\psi_{AB})\|_\infty\leq(2\pi e^{\pi})^{-1}/12$ with a probability of at least $4/5$. Thus, with a probability of at least $4/5$, we have

$$\left\|\left[\text{C-}e^{-i(\psi_{AB})^{R_A}\pi}\right]-\left[\text{C-}e^{i\psi_{AB}\pi}\right]\right\|_\diamond\leq 1/12.\tag{F9}$$

Combined with

$$\left\|\mathcal{Q}^{\psi_{AB}}-\left[\text{C-}e^{-i(\psi_{AB})^{R_A}\pi}\right]\right\|_\diamond\leq 1/12,\tag{F10}$$

we obtain

$$\left\|\mathcal{Q}^{\psi_{AB}}-\left[\text{C-}e^{i\psi_{AB}\pi}\right]\right\|_\diamond\leq 1/6.\tag{F11}$$

According to the definition of the diamond distance, we have $\|\Sigma'-\Sigma\|_1\leq 1/6$. By Lemma 3, we know

$$\left|\mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'\right]-\mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma\right]\right|\leq 1/6.\tag{F12}$$

Combined with Eq. (F3), we have $\mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'\right]\geq 5/6$. Now we can conclude that

$$\begin{aligned}
&\Pr\left[\text{obtain }|1\rangle\,\middle|\rho\in\mathscr{P}_1\right]\\
\geq&\Pr\left[\text{Eq. (F11) holds}\,\middle|\rho\in\mathscr{P}_1\right]\\
&\Pr\left[\text{obtain }|1\rangle\,\middle|\text{Eq. (F11) holds},\rho\in\mathscr{P}_1\right]\\
\geq&\frac{4}{5}\cdot\frac{5}{6}=\frac{2}{3}.
\end{aligned}\tag{F13}$$

On the other hand, when $\rho = \psi_A \otimes \psi_B$, we denote

$$\begin{aligned} \Gamma &:= \left[\text{C-}e^{-i(\psi_A\otimes\psi_B)^{R_A}\pi}\right](|+\rangle\!\langle+|\otimes\psi_A\otimes\psi_B), \\ \Gamma' &:= \mathcal{Q}^{\psi_A\otimes\psi_B}(|+\rangle\!\langle+|\otimes\psi_A\otimes\psi_B). \end{aligned} \tag{F14}$$

It can be observed that

$$e^{-i(\psi_A\otimes\psi_B)^{R_A}\pi} = e^{-i((\mathbb{I}_A-\psi_A)\otimes\psi_B)\pi} = \mathbb{I}_d - 2\Pi, \tag{F15}$$

where $\Pi = (\mathbb{I}_A - \psi_A)\otimes\psi_B$ is a projector that satisfies $\Pi(\psi_A\otimes\psi_B) = (\psi_A\otimes\psi_B)\Pi = 0$. This leads to the expression of Γ as follows:

$$\Gamma = \frac{1}{2}\begin{bmatrix} \psi_A\otimes\psi_B & \psi_A\otimes\psi_B \\ \psi_A\otimes\psi_B & \psi_A\otimes\psi_B \end{bmatrix}, \tag{F16}$$

which satisfies the property

$$\operatorname{Tr}\left[(|+\rangle\!\langle+|\otimes\mathbb{I}_d)\Gamma\right] = \operatorname{Tr}(\psi_A\otimes\psi_B) = 1. \tag{F17}$$

Since we have

$$\left\|\mathcal{Q}^{\psi_A\otimes\psi_B} - \left[\text{C-}e^{-i(\psi_A\otimes\psi_B)^{R_A}\pi}\right]\right\|_\diamond \le 1/12, \tag{F18}$$

it follows that

$$\left|\operatorname{Tr}\left[(|+\rangle\!\langle+|\otimes\mathbb{I}_d)\Gamma'\right] - \operatorname{Tr}\left[(|+\rangle\!\langle+|\otimes\mathbb{I}_d)\Gamma\right]\right| \le 1/12. \tag{F19}$$

Combined with Eq. (F17), we can deduce that

$$\Pr\left[\text{obtain } |0\rangle \,\middle|\, \rho\in\mathscr{P}_0\right] \ge \frac{11}{12}. \tag{F20}$$

Combining the above derivations, the probability of successfully distinguishing the sets satisfies

$$\begin{aligned} \Pr\left[\text{success}\right] =& \Pr\left[\rho\in\mathscr{P}_0\right]\Pr\left[\text{obtain } |0\rangle \,\middle|\, \rho\in\mathscr{P}_0\right] \\ &+ \Pr\left[\rho\in\mathscr{P}_1\right]\Pr\left[\text{obtain } |1\rangle \,\middle|\, \rho\in\mathscr{P}_1\right] \\ \ge& \frac{1}{2}\cdot\frac{11}{12} + \frac{1}{2}\cdot\frac{2}{3} > \frac{2}{3}. \end{aligned} \tag{F21}$$

$\square$

Appendix G: Entanglement Quantification

1. HME-based negativity estimation algorithm

Algorithm 1 Negativity Estimation

Input: The measurement accuracy ϵ, the failure probability δ, the system dimension $d = d_A d_B$.
Output: An estimation $\hat{N}(\rho)$ of $N(\rho)$.

```
for i = 1 to M do
    Randomly sample an integer l according to the probability distribution {p(l) = 8/(π²(2l−1)²)}_{l=1}^∞.
    if l > L = ⌈3d/(2πϵ) + 1/2⌉ then
        Set X̂_i = 0.
    else
        Run the quantum circuit shown in Fig. 4(c) in the main context to realize the C-e^{−iρ^{T_A} t} operation for t = 2l − 1 up
    to an error of π(2l−1)/(3(2+log(2L−1))) · ϵ/d using K(l) copies of ρ.
        if The measurement result of ancilla qubit is |0⟩ then
            Set X̂_i = −(π/2)d.
        else
            Set X̂_i = +(π/2)d.
        end if
    end if
end for
Set X̂ = MedianOfMeans(X̂_1, ··· , X̂_M).
Output N̂(ρ) = (1/2)((π/2)d + X̂ − 1).
```

2. Sample complexity of HME-based negativity estimation protocol

Theorem 5. *One needs an expected number of* $\widetilde{\mathcal{O}}\left(\log(\delta^{-1})\epsilon^{-3}d^2 d_A \left\|\rho^{T_A}\right\|_1\right)$ *copies of* ρ *to ensure* $\left|\hat{N}(\rho) - N(\rho)\right| \leq \epsilon$ *with a probability of at least* $1-\delta$*. Here, the* $\widetilde{\mathcal{O}}$ *notation suppresses logarithmic expressions for* d *and* ϵ*. Besides, the sampling times scales as* $M = \mathcal{O}\left(\log(\delta^{-1})\epsilon^{-2}d\left\|\rho^{T_A}\right\|_1\right)$*, and the copies of* ρ *needed in a single run of circuit in Fig. 4(c) of the main context when setting* $t = (2l-1)$ *scales as* $K(l) = \widetilde{\mathcal{O}}(\epsilon^{-1} d d_A l)$*.*

Proof. The Fourier series of $|x|$ on $[-\pi, \pi]$ is given by

$$|x| = \frac{\pi}{2} + \sum_{n=1}^{\infty} a_{2n-1}\cos((2n-1)x), \tag{G1}$$

where $a_{2n-1} = -\frac{4}{\pi(2n-1)^2}$. Then the trace norm of ρ^{T_A} is given by

$$\begin{aligned}
\left\|\rho^{T_A}\right\|_1 &= \mathrm{Tr}\left(\left|\rho^{T_A}\right|\right) \\
&= \mathrm{Tr}\left[\frac{\pi}{2}\mathbb{I}_d + \sum_{n=1}^{\infty} a_{2n-1}\cos\left((2n-1)\rho^{T_A}\right)\right] \\
&= \frac{\pi}{2}d + \sum_{n=1}^{\infty} a_{2n-1}\,\mathrm{Tr}\left[\cos\left((2n-1)\rho^{T_A}\right)\right].
\end{aligned} \tag{G2}$$

Let $\tau^{(2n-1)}$ be the state of the ancilla qubit before the Pauli-X basis measurement when the evolution time t is set to $t = 2n-1$ in the right circuit shown in Fig. 4(c) in the main context. Correspondingly, let $\widetilde{\tau}^{(2n-1)}$ be the state of the ancilla qubit before the Pauli-X basis measurement in the left circuit shown in Fig. 4(c) in the main context. Note that

$$\mathrm{Tr}\left(\tau^{(2n-1)}X\right) = \frac{1}{d}\,\mathrm{Tr}\left[\cos\left((2n-1)\rho^{T_A}\right)\right]. \tag{G3}$$

Now, we truncate the series in Eq. (G2) and consider only the first L terms to avoid divergence. We estimate the absolute value of the sum of the discarded terms as follows:

$$\begin{aligned}
&\left|\sum_{n=L+1}^{\infty} a_{2n-1}\operatorname{Tr}\left[\cos\bigl((2n-1)\rho^{T_A}\bigr)\right]\right| \\
\leq &\sum_{n=L+1}^{\infty} |a_{2n-1}|\left|\operatorname{Tr}\left[\cos\bigl((2n-1)\rho^{T_A}\bigr)\right]\right| \\
\leq &\frac{4d}{\pi}\sum_{n=L+1}^{\infty}\frac{1}{(2n-1)^2} \\
< &\frac{4d}{\pi}\int_L^{\infty}\frac{1}{(2x-1)^2}\,dx \\
= &\frac{2d}{(2L-1)\pi}.
\end{aligned} \tag{G4}$$

To ensure that $\frac{2d}{(2L-1)\pi} \leq \epsilon_1$, we choose $L = \lceil \frac{1}{\pi} d\epsilon_1^{-1} + \frac{1}{2} \rceil$. Here, ϵ_1 represents the truncation error threshold, which we will specify later.

Our goal is to estimate the sum

$$\sum_{n=1}^{L} a_{2n-1}\operatorname{Tr}\left[\cos\bigl((2n-1)\rho^{T_A}\bigr)\right] \tag{G5}$$

up to an error of ϵ_2. To achieve this, we introduce a new random variable $\mathbb{X}$, with the following probabilities for its values:

- $-\frac{\pi}{2}d$ with probability

$$\sum_{n=1}^{L}\frac{8}{\pi^2(2n-1)^2}\langle +|\,\tau^{(2n-1)}\,|+\rangle, \tag{G6}$$

- $\frac{\pi}{2}d$ with probability

$$\sum_{n=1}^{L}\frac{8}{\pi^2(2n-1)^2}\langle -|\,\tau^{(2n-1)}\,|-\rangle, \tag{G7}$$

- 0 with probability

$$1-\sum_{n=L+1}^{\infty}\frac{8}{\pi^2(2n-1)^2}. \tag{G8}$$

The random variable $\mathbb{X}$ serves as a good estimator for Eq. (G5) because its expected value is

$$\begin{aligned}
\mathbb{E}[\mathbb{X}] = &-\frac{\pi}{2}d\left(\sum_{n=1}^{L}\frac{8}{\pi^2(2n-1)^2}\langle +|\,\tau^{(2n-1)}\,|+\rangle\right. \\
&\left.-\sum_{n=1}^{L}\frac{8}{\pi^2(2n-1)^2}\langle -|\,\tau^{(2n-1)}\,|-\rangle\right) \\
= &-\frac{\pi}{2}d\sum_{n=1}^{L}\frac{8}{\pi^2(2n-1)^2}\operatorname{Tr}\left(\tau^{(2n-1)}X\right) \\
= &\sum_{n=1}^{L} a_{2n-1}\operatorname{Tr}\left[\cos\bigl((2n-1)\rho^{T_A}\bigr)\right].
\end{aligned} \tag{G9}$$

Correspondingly, we introduce a random variable $\widetilde{\mathbb{X}}$, with the following probabilities for its values:

- $-\frac{\pi}{2}d$ with probability

$$\sum_{n=1}^{L} \frac{8}{\pi^2(2n-1)^2} \langle +| \widetilde{\tau}^{(2n-1)} |+\rangle , \tag{G10}$$

- $\frac{\pi}{2}d$ with probability

$$\sum_{n=1}^{L} \frac{8}{\pi^2(2n-1)^2} \langle -| \widetilde{\tau}^{(2n-1)} |-\rangle , \tag{G11}$$

- 0 with probability

$$1 - \sum_{n=L+1}^{\infty} \frac{8}{\pi^2(2n-1)^2}. \tag{G12}$$

The bias between the random variables $\mathbb{X}$ and $\widetilde{\mathbb{X}}$ is given by

$$\begin{aligned} &\left|\mathbb{E}[\widetilde{\mathbb{X}}] - \mathbb{E}[\mathbb{X}]\right| \\ \leq &\sum_{n=1}^{L} \frac{4d}{\pi(2n-1)^2} \left|\mathrm{Tr}\left(\widetilde{\tau}^{(2n-1)}X\right) - \mathrm{Tr}\left(\tau^{(2n-1)}X\right)\right|. \end{aligned} \tag{G13}$$

To sample the random variable $\widetilde{\mathbb{X}}$, we first sample a random integer l according to the probability distribution

$$\left\{p(l) = \frac{8}{\pi^2(2l-1)^2}\right\}_{l=1}^{\infty}. \tag{G14}$$

If $l \geq L+1$, we set $\tilde{\mathbb{X}} = 0$. Otherwise, we use HME to generate the state $\widetilde{\tau}^{(2l-1)}$, and then perform a measurement in the Pauli-X basis. Set $\widetilde{\mathbb{X}} = -\frac{\pi}{2}d$ if the measurement outcome is $|+\rangle$, set $\widetilde{\mathbb{X}} = +\frac{\pi}{2}d$ if the measurement outcome is $|-\rangle$.

For $n = 1, \cdots, L$, we set

$$e_{2n-1} := \frac{\pi(2n-1)}{2\left(2+\log(2L-1)\right)} \frac{\epsilon_2}{d}. \tag{G15}$$

We use HME to approximate the controlled-$e^{-i(2n-1)\rho^{T_A}}$ operation with a precision of e_{2n-1}. This precision ensures that $\left\|\widetilde{\tau}^{(2n-1)} - \tau^{(2n-1)}\right\|_1 \leq e_{2n-1}$. By applying Lemma 3, we bound the absolute difference between $\mathrm{Tr}\left(\widetilde{\tau}^{(2n-1)}X\right)$ and $\mathrm{Tr}\left(\tau^{(2n-1)}X\right)$ as:

$$\left|\mathrm{Tr}\left(\widetilde{\tau}^{(2n-1)}X\right) - \mathrm{Tr}\left(\tau^{(2n-1)}X\right)\right| \leq e_{2n-1}. \tag{G16}$$

Therefore, the bias between $\widetilde{\mathbb{X}}$ and $\mathbb{X}$ is bounded by

$$\begin{aligned} \left|\mathbb{E}[\widetilde{\mathbb{X}}] - \mathbb{E}[\mathbb{X}]\right| \leq &\sum_{n=1}^{L} \frac{4d}{\pi(2n-1)^2} e_{2n-1} \\ = &\frac{2\epsilon_2}{2+\log(2L-1)} \sum_{n=1}^{L} \frac{1}{2n-1} \\ < &\frac{2\epsilon_2}{2+\log(2L-1)} \left(1 + \int_1^L \frac{1}{2x-1} dx\right) = \epsilon_2. \end{aligned} \tag{G17}$$

We sample $\widetilde{\mathbb{X}}$ for M times and obtain independent and identically distributed random variables $\widetilde{\mathbb{X}}_1, \cdots, \widetilde{\mathbb{X}}_M$. Now, let's estimate the required value of M to ensure the desired accuracy and success probability. To do so, we need to estimate the variance of $\widetilde{\mathbb{X}}$, denoted as $\mathrm{Var}[\widetilde{\mathbb{X}}]$. Note that

$$\begin{aligned}
&\left|\mathbb{E}[\widetilde{\mathbb{X}}]-\left(-\frac{\pi}{2}d\right)\right| \\
\leq&\left|\mathbb{E}[\widetilde{\mathbb{X}}]-\mathbb{E}[\mathbb{X}]\right|+\left|\mathbb{E}[\mathbb{X}]-\left(-\frac{\pi}{2}d\right)\right| \\
\leq&\epsilon_2+\left|\left\|\rho^{T_A}\right\|_1-\sum_{n=L+1}^{\infty}a_{2n-1}\operatorname{Tr}\left[\cos\left((2n-1)\rho^{T_A}\right)\right]\right| \\
<&\epsilon_2+\left\|\rho^{T_A}\right\|_1+\epsilon_1.
\end{aligned} \tag{G18}$$

Thus, we have

$$\begin{aligned}
&\mathrm{Var}[\widetilde{\mathbb{X}}]=\mathbb{E}[\widetilde{\mathbb{X}}^2]-\mathbb{E}[\widetilde{\mathbb{X}}]^2 \\
\leq&\left|\mathbb{E}[\widetilde{\mathbb{X}}^2]-\frac{\pi^2d^2}{4}\right|+\left|\mathbb{E}[\widetilde{\mathbb{X}}]^2-\frac{\pi^2d^2}{4}\right| \\
=&\frac{\pi^2d^2}{4}\sum_{n=L+1}^{\infty}\frac{8}{\pi^2(2n-1)^2}+\left|\mathbb{E}[\widetilde{\mathbb{X}}]-\frac{\pi d}{2}\right|\left|\mathbb{E}[\widetilde{\mathbb{X}}]+\frac{\pi d}{2}\right| \\
<&\frac{\pi}{2}\epsilon_1 d+\pi d\left(\left\|\rho^{T_A}\right\|_1+\epsilon_1+\epsilon_2\right)=\mathcal{O}\left(d\left\|\rho^{T_A}\right\|_1\right),
\end{aligned} \tag{G19}$$

where the last inequality is given by

$$\sum_{n=L+1}^{\infty}\frac{1}{(2n-1)^2}<\frac{\pi\epsilon_1}{4d}, \tag{G20}$$

which is guaranteed by our previous choice of L. Since $\mathrm{Var}[\widetilde{\mathbb{X}}]$ is upper bounded by $\mathcal{O}\left(d\left\|\rho^{T_A}\right\|_1\right)$, we can use the following lemma to bound the number of samples required.

Lemma 11. *Suppose we aim to estimate the expectation value $\mathbb{E}[R]$ of a random variable R with finite variance, using independently drawn samples $R_1,\cdots,R_M$ from its distribution. Then by employing the median of means estimation method and taking $M=\mathcal{O}\left(\log\left(\delta^{-1}\right)\mathrm{Var}[R]\epsilon^{-2}\right)$, we can guarantee*

$$\left|\mathbf{MedianOfMeans}(R_1,\cdots,R_M)-\mathbb{E}[R]\right|\leq\epsilon \tag{G21}$$

with a probability of at least $1-\delta$.

According to Lemma 11, to estimate $\mathbb{E}[\widetilde{\mathbb{X}}]$ with accuracy ϵ_3 and a success probability of at least $1-\delta$, it suffices to sample $\mathbb{X}$ for $M=\mathcal{O}\left(\log(\delta^{-1})d\left\|\rho^{T_A}\right\|_1\epsilon_3^{-2}\right)$ times.

By Proposition 1, the expected number of ρ required for a single sampling of $\widetilde{\mathbb{X}}$ is

$$\begin{aligned}
&\mathop{\mathbb{E}}_{l\sim p(\cdot)}\left[\mathcal{O}\left(e_{2l-1}^{-1}d_A(2l-1)^2\right)\right] \\
=&\mathcal{O}\left(\sum_{l=1}^{L}\frac{8}{\pi^2(2l-1)^2}\left(\frac{2\left(2+\log(2L-1)\right)}{\pi(2l-1)}\frac{d}{\epsilon_2}\right)d_A(2l-1)^2\right) \\
=&\mathcal{O}\left(\frac{16}{\pi^3}\frac{dd_A}{\epsilon_2}\sum_{l=1}^{L}\frac{(2+\log(2L-1))}{2l-1}\right) \\
=&\mathcal{O}\left(\frac{8}{\pi^3}\frac{dd_A}{\epsilon_2}\left(2+\log(2L-1)\right)^2\right) \\
=&\mathcal{O}\left(\epsilon_2^{-1}dd_A\log(L)^2\right).
\end{aligned} \tag{G22}$$

We take $\epsilon_1 = \epsilon_2 = \epsilon_3 = \frac{2\epsilon}{3}$. Now, with a success probability of at least $1-\delta$, we have

$$\begin{aligned}&\left|\left(\frac{\pi}{2}d + \textbf{MedianOfMeans}(\hat{\mathbb{X}}_1, \cdots, \hat{\mathbb{X}}_M)\right) - \left\|\rho^{T_A}\right\|_1\right| \\ \leq& \left|\textbf{MedianOfMeans}(\hat{\mathbb{X}}_1, \cdots, \hat{\mathbb{X}}_M) - \mathbb{E}[\widetilde{\mathbb{X}}]\right| \\ &+ \left|\mathbb{E}[\widetilde{\mathbb{X}}] - \mathbb{E}[\mathbb{X}]\right| + \left|\left(\frac{\pi}{2}d + \mathbb{E}[\mathbb{X}]\right) - \left\|\rho^{T_A}\right\|_1\right| \\ <& \epsilon_3 + \epsilon_2 + \epsilon_1 = 2\epsilon.\end{aligned} \tag{G23}$$

Thus, we can estimate $\left\|\rho^{T_A}\right\|_1$ to an accuracy of 2ϵ, which allows us to determine $N(\rho) = \frac{\left\|\rho^{T_A}\right\|_1 - 1}{2}$ to an accuracy of ϵ.

Putting everything together, the expectation value of the total number of copies of ρ we need is

$$\begin{aligned}&M \mathop{\mathbb{E}}_{l\sim p(\cdot)}\left[\mathcal{O}\left(e_{2l-1}^{-1} d_A (2l-1)^2\right)\right] \\ =& \mathcal{O}\left(\log(\delta^{-1}) d \left\|\rho^{T_A}\right\|_1 \epsilon_3^{-2}\right) \mathcal{O}\left(\epsilon_2^{-1} d d_A \log(L)^2\right) \\ =& \mathcal{O}\left(\log(\delta^{-1}) \epsilon^{-3} d^2 d_A \left\|\rho^{T_A}\right\|_1 \log(\epsilon^{-1} d)^2\right) \\ =& \widetilde{\mathcal{O}}\left(\log(\delta^{-1}) \epsilon^{-3} d^2 d_A \left\|\rho^{T_A}\right\|_1\right),\end{aligned} \tag{G24}$$

here the notation $\widetilde{\mathcal{O}}$, suppresses logarithmic factors of d and ϵ. □

3. Comparison with tomography-based negativity estimation protocols

To make a comparison with other negativity estimation protocols, we now discuss several properties of entanglement negativity. We first show that $\left\|\rho^{T_A}\right\|_1 \leq d_A$. Let $|\psi\rangle$ be a $d = d_A \times d_B$ dimensional pure state with Schmidt coefficients $s_1 \geq \cdots \geq s_{d_A} \geq 0$. According to Lemma 1 in [5], we have

$$\left\||\psi\rangle\langle\psi|^{T_A}\right\|_1 = (s_1 + \cdots + s_{d_A})^2. \tag{G25}$$

Since $s_1^2 + \cdots + s_{d_A}^2 = 1$, it follows that $\left\||\psi\rangle\langle\psi|^{T_A}\right\|_1 \leq d_A$. Let $\rho = \sum_{i=1}^d \lambda_i |\psi_i\rangle\langle\psi_i|$ be the spectral decomposition of ρ, then we have

$$\begin{aligned}\left\|\rho^{T_A}\right\|_1 =& \left\|\sum_{i=1}^d \lambda_i |\psi_i\rangle\langle\psi_i|^{T_A}\right\|_1 \\ \leq& \sum_{i=1}^d \lambda_i \left\||\psi_i\rangle\langle\psi_i|^{T_A}\right\|_1 \leq d_A.\end{aligned} \tag{G26}$$

We then show that for all d-diamonsional states ρ_1 and ρ_2, the following inequality holds:

$$|N(\rho_1) - N(\rho_2)| \leq \frac{\sqrt{d}}{2}\|\rho_1 - \rho_2\|_1. \tag{G27}$$

Let $\rho_1 - \rho_2 = \sum_{i=1}^d \mu_i |\phi_i\rangle\langle\phi_i|$ be the spectral decomposition of $\rho_1 - \rho_2$, then

$$\begin{aligned}&|N(\rho_1) - N(\rho_2)| = \frac{1}{2}\left|\left\|\rho_1^{T_A}\right\|_1 - \left\|\rho_2^{T_A}\right\|_1\right| \\ \leq& \frac{1}{2}\left\|\rho_1^{T_A} - \rho_2^{T_A}\right\|_1 = \frac{1}{2}\left\|(\rho_1 - \rho_2)^{T_A}\right\|_1 \\ \leq& \frac{1}{2}\sum_{i=1}^d |\mu_i| \left\||\phi_i\rangle\langle\phi_i|^{T_A}\right\|_1 \leq \frac{d_A}{2}\|\rho_1 - \rho_2\|_1 \\ \leq& \frac{\sqrt{d}}{2}\|\rho_1 - \rho_2\|_1.\end{aligned} \tag{G28}$$

As mentioned before, a straightforward negativity estimation protocol is to first generate an estimation of ρ using quantum state tomography and then compute the negativity based on it. Quantum state tomography uses multiple copies of ρ to learn a classical description $\hat{\rho}$ such that $\|\hat{\rho} - \rho\|_1 \leq \epsilon$. The sample complexity of quantum state tomography using single-copy measurements is proved to be $\Theta(\epsilon^{-2}d^3)$ [13, 14]. It seems that this single-copy strategy has a lower sample complexity than the HME-based protocol.

However, accurately learning a quantum state in trace distance is not equivalent to accurately estimating negativity. This is due to the relation $|N(\rho_1) - N(\rho_2)| \leq \frac{\sqrt{d}}{2}\|\rho_1 - \rho_2\|_1$ we derived before, where the equal sign can be asymptotically satisfied. Therefore, if one wants to estimate the negativity of any state to ϵ accuracy, an accuracy of $\Theta(\epsilon/\sqrt{d})$ in trace distance for state tomography is required. Substituting this into the complexity of quantum state tomography, the complexity of this tomography-based negativity estimation protocol scales as $\Theta(\epsilon^{-2}d^4)$, which is higher than that of the HME-based protocol.

Furthermore, calculating the negativity from the classical description of ρ requires conducting spectral decomposition of an exponentially large matrix, leading to requirements for exponentially large computational time and classical memory. While in the HME-based protocol, the number of classical calculations, $\mathcal{O}(M)$, is much lower than that of the spectral decomposition. Besides, the classical memory requirement of Algorithm 1 mainly comes from the subroutine **MedianOfMeans**, which scales logarithmically with d.

4. Sample complexity lower bound for estimating negativity

The exponential sample complexity of estimating negativity is unavoidable, even with highly joint operations. Based on the Holevo-Helstrom theorem, we prove that:

Proposition 4. *If there exists a protocol that can estimate $N(\rho)$ to ϵ accuracy with a probability of at least $2/3$ for any state ρ, the worst case sample complexity is lower bounded by $\Omega(\epsilon^{-2}d)$, where d is the dimension of the Hilbert space for ρ.*

Proof. Consider the case when $d_A = d_B = \sqrt{d}$. We define $\rho(x)$ as follows:

$$\rho(x) := x\frac{1}{\sqrt{d}}\Phi^+ + (1-x)\frac{1}{d}\mathbb{I}_d, \tag{G29}$$

where Φ^+ is the unnormalized maximally entangled state. This state provides an example that the difference in negativity will be exponentially larger than the difference in trace distance. For $x > \frac{1}{1+\sqrt{d}}$, we have

$$N[\rho(x)] = x\frac{\sqrt{d}}{2} - \frac{1}{2} + \frac{1-x}{2\sqrt{d}}. \tag{G30}$$

Thus, when $x_1 > x_2 > \frac{1}{1+\sqrt{d}}$, the difference in negativity will be

$$N[\rho(x_1)] - N[\rho(x_2)] = \left(\frac{\sqrt{d}}{2} - \frac{1}{2\sqrt{d}}\right)(x_1 - x_2). \tag{G31}$$

While at the same time,

$$\|\rho(x_1) - \rho(x_2)\|_1 = 2\left(1 - \frac{1}{d}\right)(x_1 - x_2). \tag{G32}$$

Now we use $\rho(x)$ to prove Proposition 4. For $0 \leq x_1, x_2 \leq 1$, the fidelity between $\rho(x_1)$ and $\rho(x_2)$ can be lower bounded as

$$\begin{aligned}
&F\big(\rho(x_1), \rho(x_2)\big) \\
=&\left[\sqrt{\left(x_1 + (1-x_1)\frac{1}{d}\right)\left(x_2 + (1-x_2)\frac{1}{d}\right)}\right. \\
&\left. + (d-1)\sqrt{(1-x_1)(1-x_2)\frac{1}{d^2}}\right]^2 \\
\geq&\left[\sqrt{x_1x_2} + \sqrt{(1-x_1)(1-x_2)}\right]^2,
\end{aligned} \tag{G33}$$

where the inequality holds for all $0 \leq x_1, x_2 \leq 1$ and all dimension d. Taking $\lambda = 6\epsilon/\sqrt{d}$,

$$N[\rho(1/2+\lambda)] - N[\rho(1/2)] = \lambda\left(\frac{\sqrt{d}}{2} - \frac{1}{2\sqrt{d}}\right) > 2.5\epsilon \tag{G34}$$

holds when $d > 6$. Thus, if we can estimate $N(\rho)$ to an accuracy of ϵ with a success probability of at least $2/3$ using K copies of the unknown state ρ, then we can distinguish between $\rho(1/2+\lambda)$ and $\rho(1/2)$ with a probability of at least $2/3$ using K copies of ρ. By Eq. (G33), we have

$$\begin{aligned} F\big(\rho(1/2+\lambda), \rho(1/2)\big) &\geq \left[\sqrt{\frac{1}{4}+\frac{\lambda}{2}} + \sqrt{\frac{1}{4}-\frac{\lambda}{2}}\right]^2 \\ &\geq 1 - 2\lambda^2, \end{aligned} \tag{G35}$$

where the second inequality holds for all $\lambda \in (0, 1/2]$.

By Lemma 7, in order to distinguish between $\rho(1/2+\lambda)$ and $\rho(1/2)$ with a probability of at least $2/3$ using K copies of the unknown state, K must be large enough to make sure

$$\left\|\rho(1/2+\lambda)^{\otimes K} - \rho(1/2)^{\otimes K}\right\|_1 \geq \frac{2}{3}. \tag{G36}$$

Note that

$$\begin{aligned} &\left\|\rho(1/2+\lambda)^{\otimes K} - \rho(1/2)^{\otimes K}\right\|_1 \\ \leq& 2\sqrt{1 - F\big(\rho(1/2+\lambda)^{\otimes K}, \rho(1/2)^{\otimes K}\big)} \\ =& 2\sqrt{1 - F\big(\rho(1/2+\lambda), \rho(1/2)\big)^K} \\ \leq& 2\sqrt{1 - (1-2\lambda^2)^K}. \end{aligned} \tag{G37}$$

Combine Eq. (G36) and (G37), K must satisfies

$$\begin{aligned} K &\geq \log\left(\frac{9}{8}\right)\log\left(\frac{1}{1-2\lambda^2}\right)^{-1} \\ &\geq \log\left(\frac{9}{8}\right)\frac{1}{4\lambda^2} \\ &\geq \Omega\left(\lambda^{-2}\right) = \Omega\left(\epsilon^{-2}d\right), \end{aligned} \tag{G38}$$

where the second inequality holds when $\lambda \in (0, 1/2]$. □

Appendix H: Quantum Noiseless State Recovery

1. Sample complexity

Theorem 6. *Let $\mathcal{E}$ be an invertible noise map. Denote $H_{\mathcal{E}^{-1}} := \Lambda_{\mathcal{E}^{-1}}^{T_1}$ as the Hamiltonian corresponding to the inverse noise channel, and let $F := \langle\psi|\,\sigma\,|\psi\rangle$ be the fidelity between $|\psi\rangle\!\langle\psi|$ and σ. Then we need at most $\mathcal{O}\left(\log(\delta^{-1})\epsilon^{-1}F^{-2}\|H_{\mathcal{E}^{-1}}\|_\infty^2\right)$ copies of $\mathcal{E}(|\psi\rangle\!\langle\psi|)$ to approximately produce $|\psi\rangle\!\langle\psi|$ with trace distance smaller than ϵ and a success probability at least $1-\delta$ using the circuit shown in Fig. 5 of the main context.*

Proof. We use $\mathcal{Q}$ to represent the channel constructed by HME, which approximates the channel $\left[\text{C-}e^{-i\psi\pi}\right]$. We select the number of sequential operations $K = \mathcal{O}(\tilde{\epsilon}^{-1}\|H_{\mathcal{E}^{-1}}\|_\infty^2)$ to guarantee $\left\|\mathcal{Q} - \left[\text{C-}e^{-i\psi\pi}\right]\right\|_\diamond \leq \tilde{\epsilon}$, where $\tilde{\epsilon}$ represents an error threshold that will be further defined in our analysis. Let

$$\begin{aligned} \Sigma &:= \left[\text{C-}e^{-i\psi\pi}\right]\left(|+\rangle\!\langle+| \otimes \sigma\right), \\ \Sigma' &:= \mathcal{Q}(|+\rangle\!\langle+| \otimes \sigma), \end{aligned} \tag{H1}$$

by definition of diamond distance, we have $\|\Sigma' - \Sigma\|_1 < \tilde{\epsilon}$. After measuring the ancilla qubit of Σ' in the Pauli-X basis and obtaining the outcome $|-\rangle$, the state evolves into

$$\mathrm{Tr}_c\left(\frac{(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'(|-\rangle\langle-|\otimes\mathbb{I}_d)}{\mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'\right]}\right), \tag{H2}$$

which closely approximates ψ. Note that

$$F' := \mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'\right] \tag{H3}$$

is the probability of obtaining the measurement result $|-\rangle$. By Lemma 3, we can observe that

$$\begin{aligned}&\left|F' - F\right|\\ =&\left|\mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'\right] - \mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma\right]\right|\\ \leq&\|\Sigma' - \Sigma\|_1 < \tilde{\epsilon},\end{aligned} \tag{H4}$$

where

$$F := \mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma\right] = \langle\psi|\,\sigma\,|\psi\rangle \tag{H5}$$

is the fidelity between the initial state σ and the target noiseless state ψ. We have

$$\begin{aligned}&\left\|\mathrm{Tr}_c\left(\frac{(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'(|-\rangle\langle-|\otimes\mathbb{I}_d)}{\mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'\right]}\right) - |\psi\rangle\langle\psi|\right\|_1\\ =&\left\|\mathrm{Tr}_c\left(\frac{(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'(|-\rangle\langle-|\otimes\mathbb{I}_d)}{\mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'\right]}\right) - \mathrm{Tr}_c\left(\frac{(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma(|-\rangle\langle-|\otimes\mathbb{I}_d)}{\mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma\right]}\right)\right\|_1\\ \leq&\left\|\frac{1}{F'}\mathrm{Tr}_c\left((|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'(|-\rangle\langle-|\otimes\mathbb{I}_d)\right) - \frac{1}{F'}\mathrm{Tr}_c\left((|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma(|-\rangle\langle-|\otimes\mathbb{I}_d)\right)\right\|_1\\ &+\left\|\frac{1}{F'}\mathrm{Tr}_c\left((|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma(|-\rangle\langle-|\otimes\mathbb{I}_d)\right) - \frac{1}{F}\mathrm{Tr}_c\left((|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma(|-\rangle\langle-|\otimes\mathbb{I}_d)\right)\right\|_1\\ \leq&\frac{1}{F'}\left\|(|-\rangle\langle-|\otimes\mathbb{I}_d)(\Sigma'-\Sigma)(|-\rangle\langle-|\otimes\mathbb{I}_d)\right\|_1 + \left|\frac{1}{F'} - \frac{1}{F}\right|\left\|\mathrm{Tr}_c\left((|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma(|-\rangle\langle-|\otimes\mathbb{I}_d)\right)\right\|_1\\ \leq&\frac{\tilde{\epsilon}}{F'} + \frac{\tilde{\epsilon}}{F'F}F = \frac{2\tilde{\epsilon}}{F'} \leq \frac{2\tilde{\epsilon}}{F-\tilde{\epsilon}}.\end{aligned} \tag{H6}$$

In order to guarantee

$$\left\|\mathrm{Tr}_c\left(\frac{(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'(|-\rangle\langle-|\otimes\mathbb{I}_d)}{\mathrm{Tr}\left[(|-\rangle\langle-|\otimes\mathbb{I}_d)\Sigma'\right]}\right) - |\psi\rangle\langle\psi|\right\|_1 < \epsilon \tag{H7}$$

for some given error threshold ϵ, we can take $\tilde{\epsilon} = F\epsilon/3$. Thus, $\mathcal{O}(\epsilon^{-1}F^{-1}\|H_{\mathcal{E}^{-1}}\|_\infty^2)$ copies of ρ are sufficient for us to get ψ up to an error of ϵ when the post-selection succeeds.

We repeatedly run the circuit shown in Fig. 5 in the main context for m times to increase the success probability to $1-(1-F')^m$. To ensure $1-(1-F')^m \geq 1-\delta$, it suffices to take

$$\begin{aligned}m =& \left\lceil \log(\delta^{-1})\log\left(\frac{1}{1-F'}\right)^{-1}\right\rceil\\ =&\mathcal{O}\left(\log(\delta^{-1})F'^{-1}\right) = \mathcal{O}\left(\log(\delta^{-1})F^{-1}\right).\end{aligned} \tag{H8}$$

Thus, the total number of copies of ρ we need to achieve a success probability of at least $1-\delta$ is

$$\mathcal{O}\left(\log(\delta^{-1})\epsilon^{-1}F^{-2}\|H_{\mathcal{E}^{-1}}\|_\infty^2\right). \tag{H9}$$

$\square$

2. Comparison with quantum error correction and mitigation

We should acknowledge that, in comparison to the well-developed quantum error correction and mitigation protocols, quantum noiseless state recovery still faces significant challenges. The exponentiation of quantum noiseless state recovery relies on rigorous assumptions, such as having precise knowledge of the noise channel to determine the Hamiltonian and an efficient method for exponentiating controlled Hamiltonian evolution. To apply quantum noiseless state recovery in practical scenarios, many efforts should be made to enhance its practicality. However, quantum noiseless state recovery exhibits fundamental differences from quantum error correction and mitigation, which may provide inspiration for alternative error management protocols. Therefore, a deeper understanding of the characteristics of quantum noiseless state recovery is meaningful.

Quantum noiseless state recovery combines the starting point of quantum error mitigation and the target of quantum error correction. When considering the state of processing, quantum noiseless state recovery is closer to quantum error mitigation, as it deals with states that have already been affected by noises. In contrast, quantum error correction begins with noiseless states and protects them from the influence of noises. Regarding the target, quantum noiseless state recovery aligns more closely with quantum error correction since both aim to recover the noiseless state rather than solely the noiseless expectation values. For a clearer comparison, we present a summary of the distinctive features of these methods in Table I.

	QEC	QEM	QNSR
Main Technique	Encoding and Decoding	Statistical Methods	HME and Quantum Phase Estimation
State of Processing	Noiseless	Noisy	Noisy
Target	State Recovery	Expectation Value Recovery	State Recovery
Qubit Overhead	High	Low	Moderate
Sample Complexity	Polynomial	Exponential	Exponential
General Noise	No	Yes	Yes

TABLE I. The comparison of QEC (quantum error correction), QEM (quantum error mitigation), and QNSR (quantum noiseless state recovery).

In terms of qubit overhead, quantum noiseless state recovery also falls between quantum error correction and mitigation. For the circuit of quantum noiseless state recovery, as shown in Fig. 5 in the main context, a system with $(2n+1)$ qubits can be utilized to recover a n-qubit noiseless state. In comparison, typical error correction codes such as the surface code or concatenated code [15, 16] may require a much larger number of ancilla qubits to protect a single logical qubit. Whereas, quantum error mitigation only requires a small number of ancilla qubits, but is thus limited to recovering the noiseless expectation values. It is worth noting that the ability to recover quantum states, rather than solely retrieving expectation values, offers several advantages [17]. For instance, in applications such as quantum storage, the primary goal is to obtain noise-free quantum states. Additionally, in practical scenarios such as the Shor algorithm [18] and quantum key distribution [19, 20], noiseless states are required to obtain single-shot measurement results instead of just expectation values.

Both quantum error mitigation and quantum noiseless state recovery suffer from higher sample complexities compared to quantum error correction. The presence of single-qubit independent noise poses a significant challenge for quantum devices by greatly limiting their ability to demonstrate quantum advantages and solve practical problems [21]. Studies have revealed that for independent single-qubit noises, the sample complexity of quantum error mitigation has an exponential lower bound [22, 23]. Unfortunately, quantum noiseless state recovery faces the same difficulty in this regard. Based on Theorem 2, we prove that:

Corollary 1. *Let $\mathcal{E}^{\otimes n}$ be an n-qubit noise channel, which is the n-fold tensor product of single-qubit invertible and non-unitary noise channel $\mathcal{E}$. Then, the sample complexity of realizing controlled-$e^{-i(\mathcal{E}^{-1})^{\otimes n}(\rho)t}$ evolution to an accuracy of ϵ in diamond distance has an exponential lower bound, $K \geq \epsilon^{-1}2^{\Omega(n)}t^2$.*

Proof. In this section, let $M_{\mathbf{r}} := \frac{1}{2}(\mathbb{I}_2+\mathbf{r}\cdot\boldsymbol{\sigma})$ be the Bloch representation of a two-dimensional Hermitian matrix, where $\mathbf{r} = (x, y, z) \in \mathbb{R}^3$ is the Bloch vector of $M_{\mathbf{r}}$, and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3) = (X, Y, Z)$. Additionally, $\|\mathbf{r}\| = \sqrt{x^2 + y^2 + z^2}$ represents the norm of $\mathbf{r}$, $D^3 = \{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$ represents the closed unit ball in $\mathbb{R}^3$, and $S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$ represents the unit sphere.

The action of a trace-preserving map in the Bloch representation can always be expressed as an affine map [8], which maps one ellipsoid to another ellipsoid. Note that affine maps preserve antipodal points of ellipsoids. We use the notation $\mathcal{E}$ to refer to both the CPTP noise channel and the corresponding affine map on $\mathbb{R}^3$. If $\mathcal{E}$ is not a unitary channel, it maps some pure state on S^2 to the interior of D^3. Consequently, $\mathcal{E}(D^3) \subsetneq D^3$, which implies that $D^3 \subsetneq \mathcal{E}^{-1}(D^3)$.

In the following, we will discuss two cases separately: when $\mathcal{E}$ is a unital map, and when it is not a unital map.

If $\mathcal{E}$ is a unital map, then $(0,0,0)$ is the fixed point of $\mathcal{E}$. Since $D^3 \subsetneq \mathcal{E}^{-1}(D^3)$, we can choose $M_{\pm\mathbf{r}} \in \mathcal{E}^{-1}(S^2)$ such that $\|\mathbf{r}\| > 1$. Besides, the two pure states, $|\psi_\pm\rangle\langle\psi_\pm| := \mathcal{E}(M_{\pm\mathbf{r}})$, are orthogonal since their Bloch vectors are antipodal points on S^2. Let the spectral decomposition of $M_{\pm\mathbf{r}}$ be

$$M_{\pm\mathbf{r}} = \frac{1 \pm \|\mathbf{r}\|}{2} |w_1\rangle\langle w_1| + \frac{1 \mp \|\mathbf{r}\|}{2} |w_2\rangle\langle w_2| , \tag{H10}$$

where $|w_1\rangle$ and $|w_2\rangle$ are two orthonormal pure states that are determined by $\mathbf{r}$. One can show that

$$A_n := \frac{1}{2} |\psi_+\rangle\langle\psi_+|^{\otimes n} - \frac{1}{2} |\psi_-\rangle\langle\psi_-|^{\otimes n} \tag{H11}$$

is an element of the feasible region $\mathscr{F}$, as $[M_{+\mathbf{r}}, M_{-\mathbf{r}}] = 0$. Acting $\mathcal{N}_n := (\mathcal{E}^{-1})^{\otimes n}$ on A_n, we have

$$\begin{aligned} \mathcal{N}_n(A_n) &= \frac{1}{2}(M_{+\mathbf{r}})^{\otimes n} - \frac{1}{2}(M_{-\mathbf{r}})^{\otimes n} \\ =& \frac{1}{2}\left[\left(\frac{1+\|\mathbf{r}\|}{2}\right)^n - \left(\frac{1-\|\mathbf{r}\|}{2}\right)^n\right] |w_1\rangle\langle w_1|^{\otimes n} \\ &+ \frac{1}{2}\left[\left(\frac{1-\|\mathbf{r}\|}{2}\right)^n - \left(\frac{1+\|\mathbf{r}\|}{2}\right)^n\right] |w_2\rangle\langle w_2|^{\otimes n} \\ &+ \cdots . \end{aligned} \tag{H12}$$

Thus we have

$$\begin{aligned} &R_* \geq R\left[\mathcal{N}_n\left(A_n\right)\right] \\ \geq& \frac{1}{2}\left[\left(\frac{1+\|\mathbf{r}\|}{2}\right)^n - \left(\frac{1-\|\mathbf{r}\|}{2}\right)^n\right] \\ &- \frac{1}{2}\left[\left(\frac{1-\|\mathbf{r}\|}{2}\right)^n - \left(\frac{1+\|\mathbf{r}\|}{2}\right)^n\right] \\ \geq& \Omega\left(\left(\frac{1+\|\mathbf{r}\|}{2}\right)^n\right) . \end{aligned} \tag{H13}$$

Since $\frac{1+\|\mathbf{r}\|}{2} > 1$, Theorem 2 implies that the sample complexity K for realizing $e^{-i\mathcal{N}_n(\rho)t}$ with accuracy ϵ has a lower bound of $K \geq \epsilon^{-1} 2^{\Omega(n)} t^2$.

When $\mathcal{E}$ is not a unital map, then $\mathcal{E}^{-1}$ is also not a unital map. Denote $\mathcal{E}^{-1}(\frac{\mathbb{I}_2}{2}) = M_{\mathbf{c}}$, then the center of the ellipsoid $\mathcal{E}^{-1}(D^3)$ is given by $\mathbf{c}$. The line passing through $(0,0,0)$ and $\mathbf{c}$ intersects the ellipsoid plane $\mathcal{E}^{-1}(S^2)$ at two points $\mathbf{s}$ and $\mathbf{t}$, which are antipodal points on $\mathcal{E}^{-1}(S^2)$ and satisfy $\mathbf{t} = -\frac{\|\mathbf{t}\|}{\|\mathbf{s}\|}\mathbf{s}$. Without loss of generality, we assume that $\|\mathbf{s}\| > \|\mathbf{t}\|$. Denote $|\phi_+\rangle\langle\phi_+| := \mathcal{E}(M_{\mathbf{s}})$ and $|\phi_-\rangle\langle\phi_-| := \mathcal{E}(M_{\mathbf{t}})$, they are antipodal pure states on S^2, thus $\langle\phi_+|\phi_-\rangle = 0$. Let the spectral decomposition of $M_{\mathbf{s}}$ and $M_{\mathbf{t}}$ be

$$\begin{aligned} M_{\mathbf{s}} &= \frac{1+\|\mathbf{s}\|}{2} |v_1\rangle\langle v_1| + \frac{1-\|\mathbf{s}\|}{2} |v_2\rangle\langle v_2| , \\ M_{\mathbf{t}} &= \frac{1+\|\mathbf{t}\|}{2} |v_2\rangle\langle v_2| + \frac{1-\|\mathbf{t}\|}{2} |v_1\rangle\langle v_1| , \end{aligned} \tag{H14}$$

where $|v_1\rangle$ and $|v_2\rangle$ are orthonormal pure states that are determined by $\mathbf{s}$. One can show that

$$B_n := \frac{1}{2} |\phi_+\rangle\langle\phi_+|^{\otimes n} - \frac{1}{2} |\phi_-\rangle\langle\phi_-|^{\otimes n} \tag{H15}$$

is an element of the feasible region $\mathscr{F}$, as $[M_{\mathbf{s}}, M_{\mathbf{t}}] = 0$. Acting $\mathcal{N}_n$ on B_n, we have

$$\begin{aligned} \mathcal{N}_n(B_n) &= \frac{1}{2}(M_{\mathbf{s}})^{\otimes n} - \frac{1}{2}(M_{\mathbf{t}})^{\otimes n} \\ =& \frac{1}{2}\left[\left(\frac{1+\|\mathbf{s}\|}{2}\right)^n - \left(\frac{1-\|\mathbf{t}\|}{2}\right)^n\right] |v_1\rangle\langle v_1|^{\otimes n} \\ &+ \frac{1}{2}\left[\left(\frac{1-\|\mathbf{s}\|}{2}\right)^n - \left(\frac{1+\|\mathbf{t}\|}{2}\right)^n\right] |v_2\rangle\langle v_2|^{\otimes n} \\ &+ \cdots . \end{aligned} \tag{H16}$$

Thus we have

$$R_* \geq R\left[\mathcal{N}_n\left(B_n\right)\right] \geq \Omega\left(\left(\frac{1+\|\mathbf{s}\|}{2}\right)^n\right). \tag{H17}$$

Since $\frac{1+\|\mathbf{s}\|}{2} > 1$, Theorem 2 implies that the sample complexity K for realizing $e^{-i\mathcal{N}_n(\rho)t}$ with accuracy ϵ has a lower bound of $K \geq \epsilon^{-1} 2^{\Omega(n)} t^2$.

Finally, note that

$$\text{C-}e^{-i\mathcal{N}_n(\rho)t} = e^{-i|1\rangle\langle 1|_c \otimes \mathcal{N}_n(\rho)t}, \tag{H18}$$

and $|1\rangle\langle 1|_c \otimes \mathcal{N}_n(\rho)$ has the same spectrum as $\mathcal{N}_n(\rho)$. Based on this observation, we can easily demonstrate that realizing controlled-$e^{-i\mathcal{N}_n(\rho)t}$ with accuracy ϵ also has a lower bound of $K \geq \epsilon^{-1} 2^{\Omega(n)} t^2$. □

It is important to emphasize that the higher sample complexity of quantum noiseless state recovery compared to quantum error correction does not solely stem from the fact that quantum noiseless state recovery uses fewer ancilla qubits. While both methods aim to obtain the noiseless state, quantum noiseless state recovery faces a more challenging task: recovering the noiseless quantum state from noisy states instead of protecting noiseless states from noises. As local noises will exponentially reduce the distinguishability of quantum states, recovery from noisy states will encounter exponential sample complexities, as suggested by the spirit of the Holevo-Helstrom theorem.

In practice, quantum noiseless state recovery excels in handling certain types of noises that pose challenges for quantum error correction. For instance, leakage error is a significant challenge for some quantum platforms [24], which occurs when the target system possesses a hidden space, and the quantum state may leak out of the computational space. Detecting and operating on the hidden space is typically challenging, making it difficult for quantum error correction to handle leakage errors effectively. In contrast, quantum noiseless state recovery can naturally handle such non-trace-preserving errors, as its inverse map is also Hermitian-preserving.

Appendix I: Other Applications

Although not as evident as in entanglement detection and noiseless state recovery, Hermitian-preserving maps actually emerge in a wide range of quantum information tasks. In this section, we will explore several other examples, such as expectation value measurements, quantum imaginary time evolution, and linear combinations of unitaries, where HME can be effectively applied.

1. Expectation value measurements

Except for protocols such as randomized measurements [25–27] and shadow tomography [28], which either require exponential repetition times or joint operations, we typically employ two methods to efficiently estimate the expectation value of a single observable, $\text{Tr}(O\rho)$. The first method is based on the fundamental principles of quantum mechanics. Suppose O can be diagonalized as $O = V_O \Lambda_O V_O^\dagger$, where V_O and Λ_O are unitary and diagonal matrices, respectively. In this case, we can evolve the state ρ using $V_O^\dagger$, measure it in the computational basis multiple times, and use these measurement results to estimate $\text{Tr}(O\rho) = \text{Tr}\left(\Lambda_O V_O^\dagger \rho V_O\right)$. The circuit for this method is shown in Fig. 5(a). The second method is based on embedding the observable into a unitary U_O. A common embedding technique is based on the textbook result $O = \frac{\|O\|_\infty}{2}(U_O + U_O^\dagger)$. With this embedding, we can perform the Hadamard test circuit shown in Fig. 5(b) using a controlled-U_O operation to acquire the value of $\text{Tr}\left[(U_O + U_O^\dagger)\rho\right]$. This method is particularly efficient when O is not only Hermitian but also unitary in certain scenarios.

In certain practical scenarios, O may correspond to a sparse or intrinsic Hamiltonian of physical systems. In such cases, implementing V_O and U_O may still require deep quantum circuits [29] as they do not exploit the structure information of O. While, the implementation of e^{-iOt} and controlled-e^{-iOt} evolutions can be quite efficient, with methods such as Hamiltonian simulation protocols [9, 10, 30] or Hilbert space enlargement techniques [31, 32].

An intriguing observation is that $\text{Tr}(O\rho)$ is also a Hermitian-preserving map that acts on ρ since $\text{Tr}(O\rho)$ is a real number. This suggests the possibility of conducting observable measurements through HME. To facilitate this application, we can define a new map $\mathcal{N}_O(\rho) = \text{Tr}(O\rho)\,|1\rangle\langle 1|$, where the output space is a single-qubit Hilbert space and $|1\rangle$ is one of the basis states in this space. According to Theorem 1 in the main context, the Hamiltonian for exponentiating $\mathcal{N}_O$ is $H_O = O \otimes |1\rangle\langle 1|$, and the evolution of $e^{-iH_O t}$ is actually the controlled-e^{-iOt} operation. With

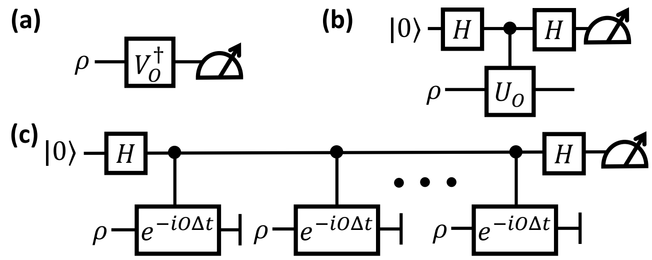

FIG. 5. The circuits of three direct expectation value measurement methods, where all the measurements are conducted in the Z basis. (a) V_O is the unitary that diagonalizes O. (b) U_O is obtained by decomposing O as $O = \frac{\|O\|_\infty}{2}(U_O + U_O^\dagger)$, and H represents the Hadamard gate. (c) The circuit utilizes HME to perform the observable measurement, utilizing controlled-$e^{-iO\Delta t}$ operations.

this Hamiltonian, HME enables the encoding of expectation values into relative phases of reference states. By setting the evolved state as $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, as shown in Fig. 5(c), the evolved state will approximately become

$$e^{-i\,\mathrm{Tr}(O\rho)t|1\rangle\langle 1|}\,|+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{-i\,\mathrm{Tr}(O\rho)t}\,|1\rangle\right). \tag{I1}$$

Subsequently, one can measure the evolved state in the Pauli basis to extract the value of $\mathrm{tr}(O\rho)$. By defining a new map as $\mathcal{N}(\rho) = \sum_i \mathrm{Tr}(O_i\rho)\,|i\rangle\langle i|$ and setting the evolved state to be a higher-dimensional coherent state, HME can encode multiple expectation values simultaneously. This HME-based phase encoding operation may also exhibit some advantages for other applications such as gradient estimation [33, 34].

2. Completely positive maps without trace preservation

Until now, most of the non-physical maps we have discussed, including positive maps, inverse maps of noise channels, and phase encoding maps, violate the complete positivity requirements for quantum channels. However, there also exist some completely positive but not trace-preserving maps that play important roles in quantum information processing tasks. One important example is the map $\mathcal{N}(\rho) = P\rho P^\dagger$, where P can be an arbitrary rectangular matrix. In the task of linear combination of unitaries [30], the matrix P is chosen as the sum of several unitaries, allowing for producing an arbitrary pure state or realizing a desired Hamiltonian evolution. Additionally, in quantum imaginary time evolution [35–37], P is set as $e^{-\beta H}$, where β represents the inverse temperature and H is a Hamiltonian. By increasing β, one can prepare a pure state that approximates the ground state of H to arbitrary precision.

It is worth noting that linear combination of unitaries and quantum imaginary time evolution share a common goal of preparing pure states, which is similar to quantum noiseless state recovery. To achieve this goal, one can adopt a similar circuit as shown in Fig. 5 in the main context. In particular, by carefully selecting a Hamiltonian, HME facilitates the realization of the controlled-$e^{-iP|0\rangle\langle 0|P^\dagger t}$ evolution, where $|0\rangle\langle 0|$ represents some initial pure state. If one has sufficient prior knowledge of P, the pure state corresponding to $P\,|0\rangle\langle 0|\,P^\dagger$ can also be generated using the quantum phase estimation circuit with a small number of ancilla qubits.

---

[1] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, 2018).

[2] M. C. Caro, H.-Y. Huang, M. Cerezo, K. Sharma, A. Sornborger, L. Cincio, and P. J. Coles, Generalization in quantum machine learning from few training data, Nature Communications **13**, 4919 (2022).

[3] S. Lloyd, M. Mohseni, and P. Rebentrost, Quantum principal component analysis, Nature Physics **10**, 631 (2014).

[4] S. Kimmel, C. Y.-Y. Lin, G. H. Low, M. Ozols, and T. J. Yoder, Hamiltonian simulation with optimal sample complexity, npj Quantum Information **3**, 13 (2017).

[5] N. Johnston and E. Patterson, The inverse eigenvalue problem for entanglement witnesses, Linear Algebra and its Applications **550**, 1 (2018).

[6] A. Holevo, Statistical decision theory for quantum systems, Journal of Multivariate Analysis **3**, 337 (1973).

[7] C. W. Helstrom, Quantum detection and estimation theory, Journal of Statistical Physics **1**, 231 (1969).

[8] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

[9] M. Suzuki, Fractal decomposition of exponential operators with applications to many-body theories and monte carlo simulations, Physics Letters A **146**, 319 (1990).

[10] M. Suzuki, General theory of fractal path integrals with applications to many-body theories and statistical physics, Journal of Mathematical Physics **32**, 400 (1991).

[11] A. Y. Kitaev, Quantum measurements and the abelian stabilizer problem, quant-ph/9511026 (1995).

[12] E. Lubkin and T. Lubkin, Average quantal behavior and thermodynamic isolation, International Journal of Theoretical Physics **32**, 933 (1993).

[13] R. Kueng, H. Rauhut, and U. Terstiege, Low rank matrix recovery from rank one measurements, Applied and Computational Harmonic Analysis **42**, 88 (2017).

[14] S. Chen, B. Huang, J. Li, A. Liu, and M. Sellke, Tight bounds for state tomography with incoherent measurements, arXiv:2206.05265 (2022).

[15] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, Phys. Rev. A **86**, 032324 (2012).

[16] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error, in *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97 (Association for Computing Machinery, New York, NY, USA, 1997) p. 176–188.

[17] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, Quantum error mitigation, arXiv:2210.00921 (2022).

[18] P. Shor, Algorithms for quantum computation: discrete logarithms and factoring, in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (1994) pp. 124–134.

[19] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (Bangalore, India, 1984) pp. 175–179.

[20] A. K. Ekert, Quantum cryptography based on bell's theorem, Phys. Rev. Lett. **67**, 661 (1991).

[21] D. Stilck França and R. García-Patrón, Limitations of optimization algorithms on noisy quantum devices, Nature Physics **17**, 1221 (2021).

[22] R. Takagi, S. Endo, S. Minagawa, and M. Gu, Fundamental limits of quantum error mitigation, npj Quantum Information **8**, 114 (2022).

[23] Y. Quek, D. S. França, S. Khatri, J. J. Meyer, and J. Eisert, Exponentially tighter bounds on limitations of quantum error mitigation, arXiv:2210.11505 (2022).

[24] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Simple pulses for elimination of leakage in weakly nonlinear qubits, Phys. Rev. Lett. **103**, 110501 (2009).

[25] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, The randomized measurement toolbox, Nature Reviews Physics **5**, 9 (2023).

[26] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Probing rényi entanglement entropy via randomized measurements, Science **364**, 260 (2019).

[27] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nature Physics **16**, 1050 (2020).

[28] S. Aaronson, Shadow tomography of quantum states, in *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018 (Association for Computing Machinery, New York, NY, USA, 2018) p. 325–338.

[29] M. Ueda, Quantum equilibration, thermalization and prethermalization in ultracold atoms, Nature Reviews Physics **2**, 669 (2020).

[30] A. Childs and N. Wiebe, Hamiltonian simulation using linear combinations of unitary operations, Quantum Information and Computation **12** (2012).

[31] X.-Q. Zhou, T. C. Ralph, P. Kalasuwan, M. Zhang, A. Peruzzo, B. P. Lanyon, and J. L. O'Brien, Adding control to arbitrary unknown quantum operations, Nature Communications **2**, 413 (2011).

[32] X.-Q. Zhou, P. Kalasuwan, T. C. Ralph, and J. L. O'Brien, Calculating unknown eigenvalues with a quantum algorithm, Nature Photonics **7**, 223 (2013).

[33] A. Gilyén, S. Arunachalam, and N. Wiebe, Optimizing quantum optimization algorithms via faster quantum gradient computation, in *Proceedings of the 2019 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* (Society for Industrial and Applied Mathematics, 2019) pp. 1425–1444.

[34] W. J. Huggins, K. Wan, J. McClean, T. E. O'Brien, N. Wiebe, and R. Babbush, Nearly optimal quantum algorithm for estimating multiple expectation values, Phys. Rev. Lett. **129**, 240501 (2022).

[35] S. Lu, M. C. Bañuls, and J. I. Cirac, Algorithms for quantum simulation at finite energies, PRX Quantum **2**, 020321 (2021).

[36] D.-B. Zhang, G.-Q. Zhang, Z.-Y. Xue, S.-L. Zhu, and Z. D. Wang, Continuous-variable assisted thermal quantum simulation, Phys. Rev. Lett. **127**, 020502 (2021).

[37] M. Motta, C. Sun, A. T. K. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, Nature Physics **16**, 205 (2020).