Noisy probabilistic error cancellation and generalized physical implementability

Check for updates

Tian-Ren Jin [1,2], Yu-Ran Zhang [3] ✉, Kai Xu[1,2,4,5,6,7] & Heng Fan [1,2,4,5,6,7] ✉

Decoherence severely limits the performance of quantum processors, posing challenges to reliable quantum computation. Probabilistic error cancellation, a quantum error mitigation method, counteracts noise by quasiprobabilistically simulating (non-physical) inverse noise operations. However, existing formulations of physical implementability, quantifying the minimal cost of simulating non-physical operations using physical channels, do not fully account for the experimental constraints, since noise also affects the cancellation process, and not all physical channels are experimentally accessible. Here, we generalize the physical implementability to encompass arbitrary convex sets of experimentally available quantum states and operations. Within this generalized framework, we demonstrate noiseless error cancellation with noisy Pauli operations and analyze the bias of noisy cancellation. Furthermore, we establish connections between generalized physical implementability and quantum information measures, e.g. diamond norm, logarithmic negativity, and purity. These findings enhance the practical applicability of probabilistic error cancellation and open new avenues for robust quantum information processing and quantum computing.

In quantum computation, an ideal quantum circuit is unitary[1]. However, the imperfection of quantum devices will lead to the noises in the performance of quantum circuits. The physical operations on a quantum system are thought to be completely positive and trace-preserving (CPTP), which is also called the quantum channel[2]. The Markov noise in quantum circuit can be depicted as a quantum noise channel $\mathcal{E}$[3].

In practice, the noise channel $\mathcal{E}$ can be evaluated by the quantum process tomography, even though the cost may be exponentially overwhelming[4]. With the knowledge of the noise channel $\mathcal{E}$, we would like to implement its inverse $\mathcal{E}^{-1}$ to cancel the impact of the noise. However, only the unitary channels have quantum channels as its inverse $\mathcal{E}^{-1}$ is also CPTP[5]. It is impossible to physically implement a quantum channel to cancel the incoherence error.

Although the inverse operation $\mathcal{E}^{-1}$ may not be a quantum channel but rather a Hermitian-preserving and trace-preserving (HPTP) operation, it can be simulated with a series of quantum channels. If it can be decomposed as the affine combination of quantum channels, the inverse operation $\mathcal{E}^{-1}$ can be simulated by the quasiprobability mixture, of which the absolute value of coefficients in the affine decomposition are normalized into a probabilistic distribution. The noise inverse operation is the probabilistic mixture of quantum channels with signatures of coefficients in this distribution up to normalization. This quantum error mitigation method is the probabilistic error cancellation (PEC)[6–17].

Furthermore, the reduced dynamics with correlated initial conditions might not always be CPTP[18,19]. The quasiprobability mixture technique allows for simulating such HPTP operations with quantum channels, which helps to analyze the non-Markov noise. The logarithm of the minimal cost, or overhead C_{em} in the context of quantum error mitigation[6,10,20], to simulate an HPTP operation with CPTP channels is defined as the physical implementability of the HPTP operation[10,21]. It has also been shown that the physical implementability of the noise inverse operation $\mathcal{E}^{-1}$ characterizes the decoherence effects of noise channel[22].

In the PEC method, by employing the Pauli-twirling technique[6,17,23–27], the error channel and the noises in experiments can be randomized compiled to be Pauli diagonal. By employing the cycle benchmarking[28,29] and error reconstruction technique[30,31], the error model of the noisy circuit can be constructed[17,26,32]. In principle, with the well-performed randomized compiling and sufficient estimation of error model, the PEC method is believed to be free of bias[6]. However, in practice, the PEC method will not always be unbiased. The bias can be induced from the violation of the error model[33,34]. This may be caused by the inaccurately randomized compiling or the inaccurate error model estimation[34].

[1]Institute of Physics, Chinese Academy of Sciences, Beijing, China. [2]School of Physical Sciences, University of Chinese Academy of Sciences, Beijing, China. [3]School of Physics and Optoelectronics, South China University of Technology, Guangzhou, China. [4]Beijing Academy of Quantum Information Sciences, Beijing, China. [5]Hefei National Laboratory, Hefei, China. [6]Songshan Lake Materials Laboratory, Dongguan, China. [7]CAS Center for Excellence in Topological Quantum Computation, UCAS, Beijing, China. ✉e-mail: yuranzhang@scut.edu.cn; hfan@iphy.ac.cn

Moreover, to cancel the error in noisy circuit, additional Pauli gates need to be introduced in the implementation of the PEC, which would also be affected by the noises in the additional circuit. Intuitively, the noises on Pauli gates $\mathcal{P}_i$ may be single-qubit errors, which is not comparable to the error to be canceled. However, due to the cross-talk between qubits[35], the noises of multi-qubit Pauli gates may not be the tensor product of noises of single-qubit Pauli gates in experiments. Therefore, the noises introduced in Pauli gates are not simply the single-qubit errors, and it is possible to yield a considerable bias of cancellation. If the noises in the cancellation are less than the error to be canceled, the PEC has a positive effect on the performance of a quantum circuit. Otherwise, the cancellation makes the performance of the quantum circuit even worse. The similar imperfection in quantum error correction leads to the famous threshold theorem[36–39], which may be the main obstacle to realizing fault-tolerant quantum computation. The error mitigation technique is not believed to have an inevitable threshold phenomenon[6]. Nevertheless, the noisy implementation of the cancellation will also lead to additional bias.

In this paper, we investigate the noisy cancellation of the error channel. We generalized the physical implementability to an arbitrary given convex set $\mathcal{F}$ of quantum operations and states that are available in experiments, and its general properties are introduced in Methods. Using the generalized physical implementability, we demonstrate the optimal method to noiseless cancel a given error channel with noisy Pauli gates in Results. We also consider the bias of the noisy cancellation with the noisy Pauli gates and the bias of the error model violation. Moreover, we illustrate the connection between the generalized physical implementability and other quantum information measures, such as the diamond norm, logarithmic negativity, and purity in Discussion.

Results

Noiseless cancellation with noisy Pauli Basis

Let $\mathcal{N}$ be an HPTP operation on system A, and then it can be decomposed as the affine combination of CPTP channels $\mathcal{N}_i$

$$\mathcal{N} = \sum_i n_i \mathcal{N}_i, \qquad (1)$$

where $\sum_i n_i = 1$. If all $n_i \geq 0$, it can be implemented by a probabilistic mixture of CPTP channels $\mathcal{N}_i$. Otherwise, for any $n_i \leq 0$, we should rewrite it as

$$\mathcal{N} = Z \sum_i \mathrm{sgn}(n_i) q_i \mathcal{N}_i, \qquad (2)$$

where $Z = \sum_i |n_i|$ and $q_i = |n_i|/Z$. Then, $\mathcal{N}$ can be simulated by using the probabilistic mixture of CPTP channels $\mathcal{N}_i$ with sign $\mathrm{sgn}(n_i)$. The quantity Z is the cost of CPTP channels for the implementation of a single HPTP operation. The physical implementability[21] is defined as the logarithm of the minimal cost of CPTP operations for HPTP operations

$$\nu(\mathcal{N}) = \log \min \left\{ \sum_i |n_i| : \mathcal{N} = \sum_i n_i \mathcal{N}_i, \mathcal{N}_i \in \mathcal{Q} \right\}, \qquad (3)$$

where $\mathcal{Q}$ denote the set of CPTP channels.

For the PEC quantum error mitigation method, the error channel is mitigated by simulating its inverse operation. Let $\mathcal{E} \equiv \mathcal{U}_\lambda \circ \mathcal{U}^\dagger$ be the error channel (in the left action) of an ideal quantum circuit $\mathcal{U}$ whose noisy circuit in experiments is $\mathcal{U}_\lambda$, as shown in Fig. 1. The quasiprobability decomposition of the inverse operation $\mathcal{E}^{-1}$ of error channel is

$$\mathcal{E}^{-1} = \sum_i r_i \mathcal{P}_i, \qquad (4)$$

where $\mathcal{P}_i$ are called noisy basis. With the Pauli-twirling technique, we consider the error model as the Pauli diagonal error, and the noisy basis $\mathcal{P}_i$ are Pauli gates. Then, the ideal expectation of the operator $\widehat{O}$ relative to the initial state ρ is

$$\left\langle \widehat{O} \right\rangle_0 = Z \sum_i \mathrm{sgn}(r_i) \frac{|r_i|}{Z} \mathrm{Tr}\left[\widehat{O} \mathcal{P}_i \circ \mathcal{U}_\lambda(\rho) \right]. \qquad (5)$$

The PEC mitigated unitary circuit $\mathcal{U}_{\mathrm{PEC}}$ is shown in Fig. 1a.

Ideally, the Pauli gates $\mathcal{P}_i$ are physically realizable quantum channels, but the inevitable noises in experiments lead to the noisy Pauli gates $\mathcal{K}_i$. Here, we encounter two different kinds of noise, one is the error channel $\mathcal{E}$ in the quantum circuit $\mathcal{U}$ of our interest, and another is the noises in the noisy basis $\mathcal{P}_i$, which is used to cancel $\mathcal{E}$. To clearly distinguish them, the former one $\mathcal{E}$ is called error, while the latter is called noise in the following. The noisy realization of the inverse of the error channel thus is

$$\mathcal{E}_\lambda^{-1} = \sum_i r_i \mathcal{K}_i \neq \mathcal{E}^{-1}, \qquad (6)$$

which cannot cancel the error channel completely, as Fig. 1a

Instead, if we ideally apply a modified PEC operation of the error channel, shown in Fig. 1b,

$$\mathcal{E}_{\mathrm{m}}^{-1} = \sum_i q_i \mathcal{P}_i, \qquad (7)$$

its noisy realization is

$$\mathcal{E}_{\lambda\mathrm{m}}^{-1} = \sum_i q_i \mathcal{K}_i. \qquad (8)$$
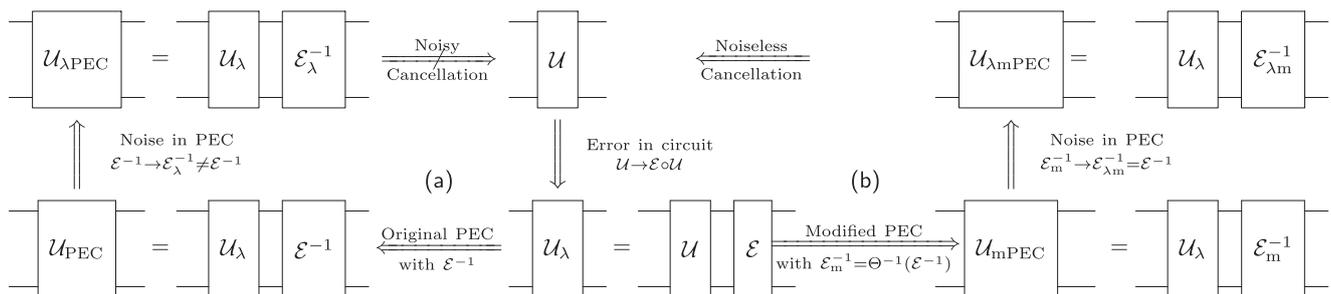

Fig. 1 | Diagram of the PEC method and its modification. a The noisy implementation $\mathcal{U}_\lambda$ of a quantum circuit $\mathcal{U}$ with error channel $\mathcal{E}$. The PEC method cancels this error by implementing its inverse operation $\mathcal{E}^{-1}$. **b** However, the noisy realization of the inverse operation $\mathcal{E}_\lambda^{-1}$ incompletely cancels the error $\mathcal{E}$. With the noises of cancellation, the PEC operation needs to be modified as $\mathcal{E}_{\mathrm{m}}^{-1} = \Theta^{-1}(\mathcal{E}^{-1})$, where Θ is the map from the ideal Pauli gates $\mathcal{P}_i$ to noisy gate $\mathcal{K}_i$, such that its noisy realization $\mathcal{E}_{\lambda\mathrm{m}}^{-1}$ cancels the error $\mathcal{E}$ complete.

We can select the parameters q_i to cancel $\mathcal{E}$ completely, i.e., $\mathcal{E}_{\lambda m}^{-1} = \mathcal{E}^{-1}$, based on the knowledge of the noises on Pauli gates.

In precise, with the Pauli twirling techniques, we still assume that the noisy Pauli gates are Pauli diagonal

$$\mathcal{K}_i = \Theta(\mathcal{P}_i) = \sum_j \Theta_{ij} \mathcal{P}_j. \tag{9}$$

The noise map Θ can be calibrated on the experimental devices. If the noisy Pauli gate $\mathcal{K}_i$ are not completely indistinguishable, i.e., $\mathcal{K}_i$ are linearly independent, the noise linear map Θ is invertible. Then, the modified PEC operation of the error channel is

$$\mathcal{E}_{m}^{-1} = \Theta^{-1}(\mathcal{E}^{-1}) = \sum_i q_i \mathcal{P}_i, \tag{10}$$

where $q_i = \sum_j r_j \Theta_{ji}^{-1}$.

There still exists a problem that which model of the noisy quasiprobability cancellation of error $\mathcal{E}$ is optimal. To depict the cost of noisy cancellation, We thus define the generalized physical implementability for any HPTP operation $\mathcal{N}$ with respect to noisy CPTP operations as

$$p_{\Theta(\mathcal{Q})}(\mathcal{N}) = \inf\left\{ \sum_i |x_i| : \mathcal{N} = \sum_i x_i \mathcal{N}_i, \mathcal{N}_i \in \Theta(\mathcal{Q}) \right\} \tag{11}$$

where the minimization is the quasiprobability decompostion $\mathcal{N} = \sum_i x_i \mathcal{N}_i$ with respect to noisy CPTP channels $\Theta(\mathcal{Q})$. Here, Θ denotes the noise map for the basis of the operation space. Moreover, for other quantum information processing tasks involving the quasiprobability decomposition, their optimal costs can also be quantified by several types of implementability functions similar to the generalized physical implementability as defined. For the general definition and properties of the implementability function, see Methods.

With the generalized physical implementability, the optimal cost of the inverse operation $\mathcal{E}^{-1}$ is $p_{\Theta(\mathcal{Q})}(\mathcal{E}^{-1})$. Assume that the linear map Θ is invertible, by the affine invariance of the implementability function (see Methods), we have

$$p_{\Theta(\mathcal{Q})}(\mathcal{E}^{-1}) = p_{\mathcal{Q}}(\Theta^{-1}(\mathcal{E}^{-1})) = p_{\mathcal{Q}}(\mathcal{E}_m^{-1}), \tag{12}$$

where $p_{\mathcal{Q}} = \exp \nu$ is the (exponential) physical implementability. Since the optimal cancellation of a mixed-unitary channel with ideal CPTP channels is the decomposition with the unitaries[21], the optimal cancellation of modified PEC cancellation operation $\mathcal{E}_m^{-1}$ is the quasiprobability decomposition with respect to the ideal Pauli channels $\mathcal{P}_i$ with quasiprobability

$$q_i = \sum_j r_j \Theta_{ji}^{-1}. \tag{13}$$

The optimal cancellation of $\mathcal{E}^{-1}$ with respect to the noisy CPTP channels $\Theta(\mathcal{Q})$ thus is the quasiprobability decomposition with respect to the noisy Pauli channels $\mathcal{K}_i$ with q_i in Eq. (13). In conclusion, with the matrix Θ_{ij} measured from the noisy Pauli basis $\mathcal{K}_i$, the inverse operation $\mathcal{E}^{-1}$ of the error channel can be optimally canceled under the noisy Pauli basis $\mathcal{K}_i$.

Here, we illustrate the result with simple examples. Assume the noise is a depolarizing error on one qubit, ideally, the error channel of evolution with error rate λ is

$$\mathcal{E} = \left(1 - \frac{3\lambda}{4}\right)\mathcal{I} + \frac{\lambda}{4}(\mathcal{X} + \mathcal{Y} + \mathcal{Z}). \tag{14}$$

It is not difficult to show that its inverse is

$$\mathcal{E}^{-1} = \frac{4-\lambda}{4(1-\lambda)}\mathcal{I} - \frac{\lambda}{4(1-\lambda)}(\mathcal{X} + \mathcal{Y} + \mathcal{Z}). \tag{15}$$

If the error rate $\lambda \ll 1$, we approximate $\Theta(\mathcal{P}_i) \approx \mathcal{P}_i$, for $\mathcal{P}_i = \mathcal{X}, \mathcal{Y}, \mathcal{Z}$, we have

$$\mathcal{U} = \frac{4-\lambda}{4(1-\lambda)}\mathcal{U}_\lambda - \frac{\lambda}{4(1-\lambda)}(\mathcal{X} + \mathcal{Y} + \mathcal{Z}) \circ \mathcal{U}_\lambda, \tag{16}$$

which is in coincidence the known results of the depolarizing error[7]. Then, we assume that $\Theta(\mathcal{I}) = \mathcal{I}$ and $\mathcal{K}_{\mathcal{P}_i} = \Theta(\mathcal{P}_i) = \mathcal{E}^\alpha \circ \mathcal{P}_i$ for $\mathcal{P}_i = \mathcal{X}, \mathcal{Y}, \mathcal{Z}$. It can be calculated as

$$\mathcal{E}^\alpha = \frac{1 + 3(1-\lambda)^\alpha}{4}\mathcal{I} + \frac{1 - (1-\lambda)^\alpha}{4}(\mathcal{X} + \mathcal{Y} + \mathcal{Z}). \tag{17}$$

Let $a = \frac{1+3(1-\lambda)^\alpha}{4}, b = \frac{1-(1-\lambda)^\alpha}{4}$, we have

$$(\Theta_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ b & a & b & b \\ b & b & a & b \\ b & b & b & a \end{pmatrix}. \tag{18}$$

Thus, the quasiprobability, Eq. (13), of optimal cancellation in terms of $(\mathcal{K}_I, \mathcal{K}_X, \mathcal{K}_Y, \mathcal{K}_Z)^T$ is obtained from Eqs. (15) and (18).

For an arbitrary Pauli diagonal error, since it is CPTP, in general, it can be written as the exponential of Lindblad operators[3].

$$\mathcal{E} = \exp \mathcal{L}, \tag{19}$$

where the Lindblad operator can be written as

$$\mathcal{L} = \sum_i \lambda_i(\mathcal{P}_i - \mathcal{I}). \tag{20}$$

Thus, the error channel is

$$\mathcal{E} = \bigcirc_i[\omega_i \mathcal{I} + (1 - \omega_i)\mathcal{P}_i], \tag{21}$$

where $\omega_i = (1 + e^{-2\lambda_i})/2$. This error model is called the Pauli-Lindblad noise model[17]. The inverse operation of the error channel is

$$\begin{aligned} \mathcal{E}^{-1} = \exp(-\mathcal{L}) &= \bigcirc_i(\mu_i \mathcal{I} + (1 - \mu_i)\mathcal{P}_i) \\ &= e^{2\sum_i \lambda_i}\bigcirc_i[\omega_i \mathcal{I} + (1 - \omega_i)(-\mathcal{P}_i)] \end{aligned} \tag{22}$$

where $\mu_i = (1 + e^{2\lambda_i})/2$. There are two ways to simulate the inverse operation of the error channel. One is to simulate $\mathcal{E}^{-1}$ as a whole channel, and the other is to simulate each layer $(\mu_i \mathcal{I} + (1 - \mu_i)\mathcal{P}_i)$, separately. No matter the ideal or noisy cancellations, the second way may have more cost than the first way[21], since the sub-multiplicity of the implementability function (see Methods). For the formalism simplicity, however, we only consider the noisy cancellation in the second way. By measuring the noise map Θ of Pauli gates in the experiment, the optimal noisy cancellation is given as

$$\mathcal{E}^{-1} = \bigcirc_i\left[\mu_i \mathcal{I} + (1 - \mu_i)\sum_j \Theta_{ij}^{-1}\mathcal{K}_j\right]. \tag{23}$$

Invertibility of noise map

The above discussion is based on the assumption that the linear map Θ is invertible. When the map Θ is not invertible, there is no perfect noisy cancellation. The invertibility of Θ is equivalent to $\det \Theta \neq 0$. In practice, assuming the true value Θ_0 of the noise map is not invertible, i.e., $\det \Theta_0 = 0$, the random fluctuation from the finite measurements will lead to the noise map Θ measured from the experiments to be invertible, $\det \Theta \neq 0$. Therefore, the condition $\det \Theta \neq 0$ calculated with the experimental data does not certainly imply the invertibility of the noise map Θ. In the following, we discuss the condition the noise map Θ is invertible under finite measurements in experiment.

For simplicity, denote the true value of the determinant of the linear map $\det \Theta_0$ as $\det_0$. With the Chebyshev's inequality, it can be shown that the probability for the case that the true value of the determinant of the linear map Θ is not invertible is expressed as

$$\mathbb{P}(\det_0 = 0) \leq \exp\left(-\frac{N}{2\|\Theta^{-1}\|_2^2}\right). \tag{24}$$

Therefore, for a given number N of measurements, the linear map Θ is invertible with the probability $(1 - \delta)$ if the map Θ satisfies that

$$\|\Theta\|_2 \geq \sqrt{\frac{2D\log\frac{1}{\delta}}{N}}, \tag{25}$$

where D is the dimension of the map Θ. For the details of calculation, see Supplementary Note V.

Bias of noisy cancellation

For the noisy cancellation, without mitigating the noises in simulation of inverse noise operation $\mathcal{E}^{-1}$, we would like to estimate the bias of the noisy cancellation $\mathcal{E}_\lambda^{-1} \circ \mathcal{E} = \Theta(\mathcal{E}^{-1}) \circ \mathcal{E}$. The bias of the expectation of Pauli operator $\widehat{O}$ is expressed as

$$\delta_\lambda = |\mathrm{Tr}[\widehat{O}\mathcal{U}(\rho)] - \mathrm{Tr}[\widehat{O}\mathcal{E}_\lambda^{-1} \circ \mathcal{E} \circ \mathcal{U}(\rho)]| \tag{26}$$
$$\leq p_\mathcal{Q}(\mathcal{I} - \mathcal{E}_\lambda^{-1} \circ \mathcal{E}),$$

where we denote $\mathcal{Q} = \mathcal{Q}(A \to A)$. With $\mathcal{E}_\lambda^{-1} = \Theta(\mathcal{E}^{-1})$, we have

$$\delta_\lambda \leq 2\Theta_\lambda p_\mathcal{Q}(\mathcal{E}_\lambda^{-1} \circ \mathcal{E}) \leq 2\Theta_\lambda p_{\Theta(\mathcal{Q})}(\mathcal{E}_\lambda^{-1}), \tag{27}$$

where $\Theta_\lambda = 1 - \min_i \Theta_{ii}$ is the maximal error probability of the noisy Pauli gates. This result allows for estimating the upper bound of the bias in experiments, with the cost of simulation $p_{\Theta(\mathcal{Q})}(\mathcal{E}_\lambda^{-1})$ and the calibration of Pauli gates Θ_λ. For the details of the calculation, see Supplementary Note V.

If the circuit consists of L layers of operations $\mathcal{U} = \prod_{i=1}^L \circ \mathcal{L}_i$, where $\mathcal{L}_i$ is the i-th layer, there are two strategies to realize the PEC method. One is to cancel the error of each layer $\mathcal{E}_i$ separately, and the other is to cancel the total error $\prod_i \mathcal{E}_i$ directly. The circuits are shown in Fig. 2. Let the noisy realization of circuit be $\mathcal{U}_\lambda = \overleftarrow{\bigcirc}_{i=1}^L \mathcal{L}_{i\lambda}$, where $\mathcal{L}_{i\lambda} = \mathcal{E}_i \circ \mathcal{L}_i$, then the error channel of the circuit is

$$\mathcal{E} = \overleftarrow{\bigcirc}_{i=1}^L \tilde{\mathcal{E}}_i, \tag{28}$$

where $\tilde{\mathcal{E}}_i = (\overleftarrow{\bigcirc}_{j > i} \mathcal{L}_j) \circ \mathcal{E}_i \circ (\overrightarrow{\bigcirc}_{j > i} \mathcal{L}_j^\dagger)$. Here, the arrow above the symbol $\bigcirc$ represents the acting direction of layers.

We consider the bias of noisy cancellation of the total error of the circuit for these two different strategies. For the separate cancellation method, the noisy realization of the error-canceled circuit is

$$\mathcal{U}_{\mathrm{PEC}} = \overleftarrow{\bigcirc}_{i=1}^L \mathcal{L}_{i\mathrm{PEC}}, \tag{29}$$

where $\mathcal{L}_{i\mathrm{PEC}} = \mathcal{E}_{i\lambda}^{-1} \circ \mathcal{E}_i \circ \mathcal{L}_i$. The error of the separate cancellation is

$$\mathcal{E}_\mathrm{S}^{-1} \circ \mathcal{E} = \overleftarrow{\bigcirc}_{i=1}^L \tilde{\mathcal{E}}_{i\mathrm{PEC}}, \tag{30}$$

where $\tilde{\mathcal{E}}_i = (\overleftarrow{\bigcirc}_{j > i} \mathcal{L}_j) \circ \mathcal{E}_{i\lambda}^{-1} \circ \mathcal{E}_i \circ (\overrightarrow{\bigcirc}_{j > i} \mathcal{L}_j^\dagger)$. The bias of the noisy separate cancellation is

$$\delta_{\lambda\mathrm{S}} \leq 2\Theta_\lambda \sum_{j=1}^L \prod_{i=1}^j p_{\Theta(\mathcal{Q})}(\mathcal{E}_{i\lambda}^{-1}). \tag{31}$$

For the detailed calculation, see Supplementary Note V. On the other hand, by applying Eq. (27), the bias of the direct cancellation method is

$$\delta_{\lambda\mathrm{D}} \leq 2\Theta_\lambda p_{\Theta(\mathcal{Q})}(\mathcal{E}_\lambda^{-1} \circ \mathcal{E})$$
$$\leq 2\Theta_\lambda \prod_{i=1}^L p_{\Theta(\mathcal{Q})}(\mathcal{E}_{i\lambda}^{-1}). \tag{32}$$

Since the upper bound of the bias of the direct cancellation method, Eq. (32), is less than the upper bound of the bias of the separate cancellation method, Eq. (31), the direct cancellation method appears to be more accurate than the separate cancellation method.

However, there are more results with the detailed analysis of the noisily canceled error $\mathcal{E}_\lambda^{-1} \circ \mathcal{E}$. We consider a particular case that the noisily canceled error $\mathcal{E}_\lambda^{-1} \circ \mathcal{E}$ is a CPTP quantum channel, $p_\mathcal{Q}(\mathcal{E}_\lambda^{-1} \circ \mathcal{E}) = 1$. For instance, if the noises are uniform on the Pauli gates, $\Theta(\mathcal{E}) = \mathcal{N} \circ \mathcal{E}$, with the commutativity of Pauli diagonal operation, it preserves $\mathcal{E}_\lambda^{-1} \circ \mathcal{E} = \mathcal{N} \in \mathcal{Q}(A \to A)$. The bias of the direct cancellation is bounded by

$$\delta_{\lambda\mathrm{D}} \leq 2\Theta_\lambda, \tag{33}$$

since $p_\mathcal{Q}(\mathcal{E}_\lambda^{-1} \circ \mathcal{E}) = 1$. For the separate cancellation method, assume the noisily canceled error for each layer $\mathcal{E}_{i\lambda}^{-1} \circ \mathcal{E}_i$ is a CPTP quantum channel; it

Fig. 2 | Noisy probabilistic error cancellations of multi-layer circuit. a Cancel the errors of the L-layer circuit separately in each layer. **b** Cancel the errors of the L-layer circuit directly as a whole error of the circuit.
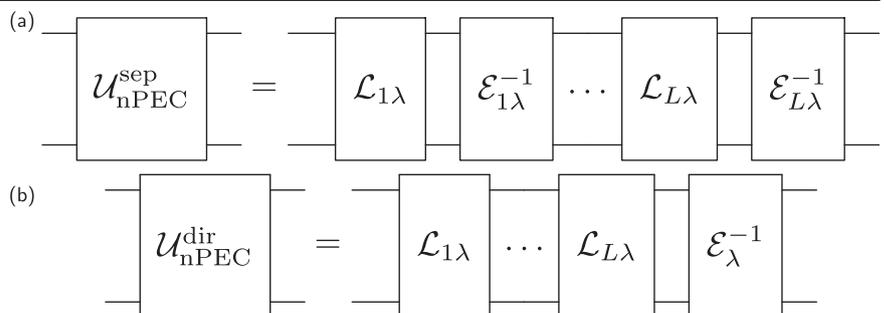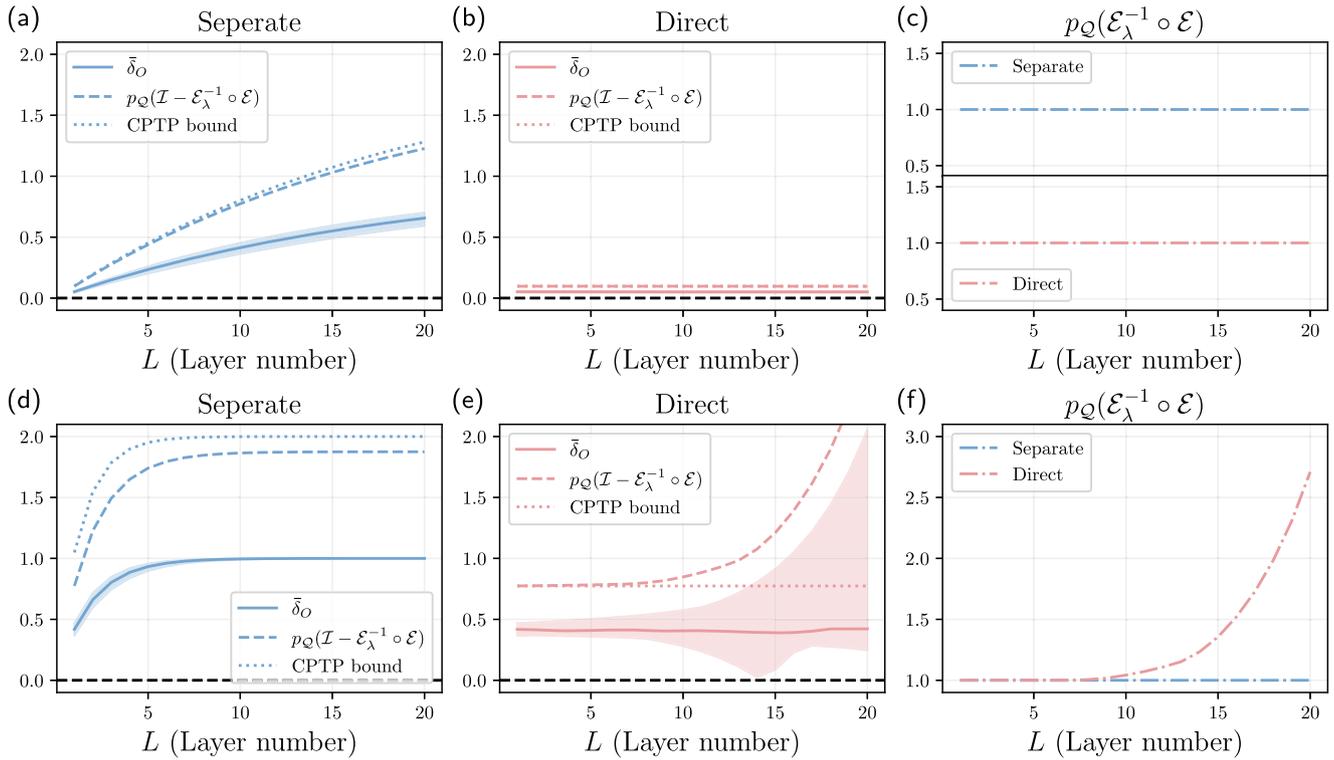
Fig. 3 | Numerical results for the biases of the noisy cancellations with separate and direct methods. a, b, d, and **e**, Colored filling regions denote the biases of expectations for different Pauli operators, and solid curves denote the median of the biases of the operator expectations. Dashed curves denote the distance between the imperfectly canceled error $\mathcal{E}_\lambda^{-1} \circ \mathcal{E}$ and $\mathcal{I}$ in the implementability function $p_\mathcal{Q}(\mathcal{I} - \mathcal{E}_\lambda^{-1} \circ \mathcal{E})$, which upper bounds the bias of the operator expectations by Eq. (26). Dotted curves denote the CPTP upper bound of $p_\mathcal{Q}(\mathcal{I} - \mathcal{E}_\lambda^{-1} \circ \mathcal{E})$, see Eqs. (33) and (34) for the direct and separate cancellation methods, respectively. The error $\mathcal{E}_j$ is assumed to be the same for each layer, $\mathcal{E}_j = \mathcal{E}_0$, and the error $\mathcal{E}_0$ as well as the noises $\mathcal{N}_i$ on Pauli gates $\mathcal{P}_i$ are randomly sampled from the Pauli-Lindblad error model (21), with a fixed single-layer error rate $\lambda = \sum_i \lambda_i$. **c** and **f**, Dotted dashed curves denote the implementability function $p_\mathcal{Q}(\mathcal{E}_\lambda^{-1} \circ \mathcal{E})$ of the imperfectly canceled error

for direct and separate cancellation methods. **a**--**c**, For a small error rate $\lambda = 0.05$, since the imperfectly canceled errors $\mathcal{E}_\lambda^{-1} \circ \mathcal{E}$ (**c**) are CPTP for circuits with a layer number less than 20, the biases of both the separate (**a**) and direct (**b**) cancellation methods are limited by the CPTP upper bound, Eqs. (33) and (34), and the bias of the direct cancellation is smaller than the separate cancellation method. **d**–**f** For a large error rate $\lambda = 0.5$, since the imperfectly canceled error $\mathcal{E}_\lambda^{-1} \circ \mathcal{E}$ for the direct cancellation method (**f**) will eventually not be CPTP with the cumulation of errors $\mathcal{E} = \mathcal{E}_0^L$. The bias increases exponentially and surpasses the CPTP bound (**e**), as the layer number L grows. However, the error $\mathcal{E}_\lambda^{-1} \circ \mathcal{E}$ for the separate cancellation method (**f**) is still CPTP, since each the error of individual layer $\mathcal{E}_{0\lambda}^{-1} \circ \mathcal{E}_0$ is CPTP. The bias of separate cancellation is still bounded by the CPTP upper bound (**f**). For details of simulation, see Supplementary Note VI.

can be shown that the bias is bounded by

$$\delta_{\lambda S} \le 2[1 - (1 - 2\Theta_\lambda)^{L/2}] \le 2. \tag{34}$$

For details of the calculation, see Supplementary Note V.

The numerical results for comparing the two cancellation methods with small and large error rates are shown in Fig. 3. When the error rate is sufficiently small, e.g., $\lambda = 0.05$ as used in Fig. 3a–c, the noisily canceled errors for both separate and direct cancellations are CPTP, Fig. 3c. Thus, the bias is limited by the CPTP upper bounds, as shown in Fig. 3a for the separate cancellation Eq. (34) and Fig. 3b for the direct cancellation Eq. (33). The bias of direct cancellation is much smaller than the separate cancellation method. When the error rate is moderately large, e.g., $\lambda = 0.5$ as used in Fig. 3d–f, the noisily canceled error $\mathcal{E}_{i\lambda}^{-1} \circ \mathcal{E}_i$ of each layer error is still CPTP, thus the noisily canceled error $\prod_i \mathcal{E}_{i\lambda}^{-1} \circ \mathcal{E}_i$ of separate cancellation is also CPTP, as shown in Fig. 3f, and the bias of separate cancellation is still limited by Eq. (34), as shown in Fig. 3d. For the direct cancellation method, the noisily canceled error $\mathcal{E}_\lambda^{-1} \circ \mathcal{E}$ is CPTP, when the layer is shallow. However, it will no longer be CPTP as the layer increases, as ilustrated in Fig. 3f. The bias of direct cancellation increases exponentially with the layer number accordig to Eq. (32) in and will surpass the CPTP upper bound of separate cancellation Eq. (34), see Fig. 3e where the CPTP upper bound is labeled by a

dotted curve. Thus, for a large error rate, the separate cancellation method is more effective than the direct cancellation method.

When the error rate is large enough, the noisily canceled error for the individual layer will not be CPTP. The bias of both the separate and direct cancellation methods will exponentially increase with the layer number, see Eqs. (32) and (31). However, for the direct cancellation method, the error of L layers of circuits may not be invertible under a finite measurement precision in the experiment, since its error rate is typically L times the error rate of the one for the individual layer. By considering the map $\Xi(\mathcal{N}) = \mathcal{E} \circ \mathcal{N}$, where $\mathcal{E}$ is the error to be canceled, the condition of invertibility of $\mathcal{E}$ is given by Eq. (25), as

$$\|\mathcal{E}\|_2 \ge \sqrt{\frac{2 \log \frac{1}{\delta}}{N}}. \tag{35}$$

In brief, for the case where the circuit is shallow and the error rate is sufficiently small, the direct cancellation method is more accurate. If the circuit is deep and the error rate is large, the separate cancellation method is more accurate.

Intuitively, one may expect a clear criterion for choosing between these two cancellation methods. The condition for the criterion requires the solution of a multi-variable polynomial equation $p_\mathcal{Q}(\Theta(\mathcal{E}^{-1}) \circ \mathcal{E}) = 1$ in

the components v_i of $\mathcal{E}$ and the element Θ_{ij} of Θ. However, the solution for the criterion is very complicated, where the error rate λ of $\mathcal{E}$ and the maximum error probability Θ_λ of the noise map Θ are not sufficient to determine the criterion. In experiments, to verify whether the noisily canceled error is CPTP requires the complete estimation of the noise map Θ, which provides sufficient information for the noiseless cancellation, thus it need not to consider the bias of noisy cancellation. Therefore, there is no feasible and reliable criterion to distinguish which method, direct or separate, has better performance. To quantitatively determine this criterion will be a topic in further investigations. Moreover, the error model used for PEC may be inaccurate, which also hinders the verification of the noisily canceled CPTP error.

Bias of Inaccurate error model

The case where the error model is inaccurate will also induce the bias of the PEC method. This is not the main target of this work and has been investigated in other works[34]. The bias of the PEC from the inaccurate error model is estimated by employing the diamond norm

$$
\begin{aligned}
\delta_O &= |\mathrm{Tr}[O\mathcal{U}(\rho)] - \mathrm{Tr}[O\widehat{\mathcal{E}}^{-1} \circ \mathcal{E} \circ \mathcal{U}(\rho)]| \\
&\leq \left\| \mathcal{I} - \widehat{\mathcal{E}}^{-1} \circ \mathcal{E} \right\|_\diamond \\
&\leq \sum_{j=1}^{L} \left\| \mathcal{I} - \widehat{\mathcal{E}}_j^{-1} \circ \mathcal{E}_j \right\|_\diamond \prod_{i=1}^{j-1} \left\| \widehat{\mathcal{E}}_i^{-1} \circ \mathcal{E}_i \right\|_\diamond,
\end{aligned}
\tag{36}
$$

where $\delta_O = |\langle O \rangle_{\mathrm{PEC}} - \langle O \rangle_{\mathrm{exact}}|$ is the bias of the PEC method with the inaccurate error model, $\mathcal{E}_i$ is the exact error channel of i-th layer, and $\widehat{\mathcal{E}}_i$ is the inaccurate error model learned from experiment. With the estimation of the diamond norm of the operation $\mathcal{I} - \widehat{\mathcal{E}}_j^{-1} \circ \mathcal{E}_j$, the bias is calculated as

$$
\delta_O \leq \sum_{j=1}^{L} \delta_j \prod_{i=1}^{j-1} \gamma_i,
\tag{37}
$$

where $\gamma_i = p_{\mathcal{Q}}(\widehat{\mathcal{E}}_i^{-1})$,

$$
\delta_j = |1 - v_0^{(j)}| + \gamma_j - v_0^{(j)}, \text{ or}
\tag{38}
$$

$$
\delta_j = |1 - v_0^{(j)}| + T(\{r_k^{(j)}\}).
\tag{39}
$$

$v_0^{(j)} = \frac{1}{4^n} \sum_k r_k^{(i)}$ is the component of $\widehat{\mathcal{E}}_j^{-1} \circ \mathcal{E}_j$ in $\mathcal{I}$, $r_k^{(i)} = f_{P_k}^{\mathrm{meas}}/f_{P_k}^{\mathrm{mod}}$ is the ratio of the measured fidelity to the model fidelity of the Pauli gate P_k, and

$$
T(\{r_k^{(j)}\}) = \frac{4^n - 1}{4^n} \sqrt{\sum_k r_k^{(j)} \left(r_k^{(j)} - \frac{1}{4^n - 1} \sum_{m \neq k} r_m^{(j)} \right)}.
\tag{40}
$$

Since the diamond norm is the implementability function $p_{\mathcal{Q}}$, Eq. (51), the bias from the inaccurate error model is compatible with the bias from the noisy cancellation operation estimated in this work. The total bias of the noisy PEC with the inaccurate error model is upper bounded as

$$
\begin{aligned}
\delta_{\lambda O} &= |\mathrm{Tr}[O\mathcal{U}(\rho) - \widehat{\mathcal{E}}_\lambda^{-1} \circ \mathcal{E} \circ \mathcal{U}(\rho)]| \\
&\leq p_{\mathcal{Q}}(\mathcal{I} - \widehat{\mathcal{E}}_\lambda^{-1} \circ \mathcal{E}) \leq \delta_O + \delta_{\lambda \mathrm{D,S}}.
\end{aligned}
\tag{41}
$$

We next consider the Pauli diagonal error model $\mathcal{E} = \exp \mathcal{L}(\lambda_i)$, where $\mathcal{L}(\lambda_i) = \sum_i \lambda_i (\mathcal{P}_i - \mathcal{I})$. Assume real error parameters $\{\lambda_i^{(j)}\}$ with deviations from the parameters $\{\widehat{\lambda}_i^{(j)}\}$, measured in experiments for the jth layer. The mitigated error channel for the jth layer is

$$
\widehat{\mathcal{E}}_j^{-1} \circ \mathcal{E}_j = \exp \mathcal{L}(\Delta \lambda_i^{(j)}),
\tag{42}
$$

where $\widehat{\mathcal{E}}_j$ is the target error channel for the mitigation, $\mathcal{E}_j$ is the real error channel, and $\Delta \lambda_i^{(j)} = \lambda_i^{(j)} - \widehat{\lambda}_i^{(j)}$. The total mitigated error channel is

$$
\widehat{\mathcal{E}}^{-1} \circ \mathcal{E} = \prod_j \widehat{\mathcal{E}}_j^{-1} \circ \mathcal{E}_j = \exp \mathcal{L}(\Delta \lambda),
\tag{43}
$$

where $\Delta \lambda_i = \sum_j \Delta \lambda_i^{(j)}$. The bias of the expectation of the Pauli operator $\widehat{O}$ is upper bounded by

$$
\begin{aligned}
\delta_O &\leq p_{\mathcal{Q}}(\mathcal{I} - \widehat{\mathcal{E}}^{-1} \circ \mathcal{E}) \\
&= p_{\mathcal{Q}}(\widehat{\mathcal{E}}^{-1} \circ \mathcal{E}) + |1 - v_0| - v_0.
\end{aligned}
\tag{44}
$$

It is crucial whether the mitigated error $\widehat{\mathcal{E}}^{-1} \circ \mathcal{E}$ is CPTP or not, and the error is called under-mitigated or over-mitigated if it is CPTP or not CPTP. The numerical results for the bias of under-mitigated and over-mitigated errors are shown in Fig. 4. For the under-mitigated error, $\widehat{\lambda}_i^{(j)} \leq \lambda_i^{(j)}$, we have
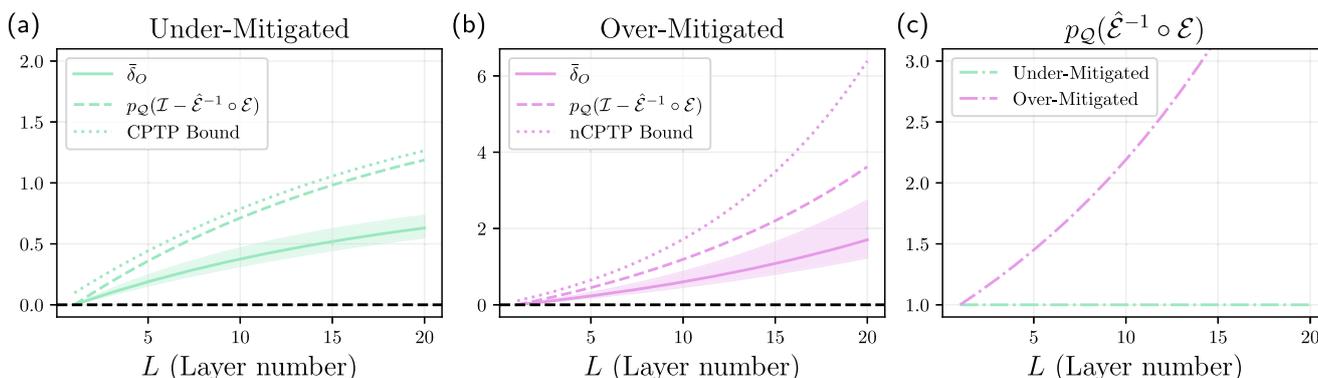

Fig. 4 | Numerical results for the biases of the inaccurate error model with under-mitigated and over-mitigated cases. a, b Colored filling regions denote the bias of expectations for different Pauli operators, and solid curves denote the median of the bias of the operator expectations. Dashed curves denote the distance between the mitigated error $\widehat{\mathcal{E}}^{-1} \circ \mathcal{E}$ and $\mathcal{I}$ characterized by the implementability function $p_{\mathcal{Q}}(\mathcal{I} - \widehat{\mathcal{E}}^{-1} \circ \mathcal{E})$, which upper bounds the bias of the operator expectations in Eq. (26). Dotted curves denote the CPTP upper bound of $p_{\mathcal{Q}}(\mathcal{I} - \widehat{\mathcal{E}}^{-1} \circ \mathcal{E})$, as shown in Eqs. (45) and (46) for the under-mitigated and over-mitigated errors, respectively. The mitigated error channel $\mathcal{E}_i^{-1} \circ \mathcal{E}_i$ for each layer is assumed to be the same, $\mathcal{E}_i = \mathcal{E}_0$, and randomly sampled from the Pauli-Lindblad error model in Eq. (21), with a fixed single-layer error rate $\Delta \lambda^{\mathrm{under}} = 0.05 \Delta \lambda_i^{\mathrm{under}} \geq 0$ for the under-mitigated error channel (**a**), and $\Delta \lambda_i^{\mathrm{over}} = -\Delta \lambda_i^{\mathrm{under}}$ for over-mitigated error channel (**b**). **c** Dotted dashed curves denote the implementability function $p_{\mathcal{Q}}(\widehat{\mathcal{E}}^{-1} \circ \mathcal{E})$ for the under-mitigated and over-mitigated errors. Since the under-mitigated error $\widehat{\mathcal{E}}^{-1} \circ \mathcal{E}$ is CPTP, the bias is lower than the CPTP upper bound as shown in (**a**). In contrast, the over-mitigated error channel $\widehat{\mathcal{E}}^{-1} \circ \mathcal{E}$ is not CPTP, the bias increases exponentially as the layer number L grows, which is bounded by the non-CPTP upper bound in Eq. (46) as shown in (**b**). For details of simulation, see Supplementary Note VI.

$p_{\mathcal{Q}}(\widehat{\mathcal{E}}^{-1} \circ \mathcal{E}) = 1$, as shown in Fig. 4c, and the upper bound is shown in Fig. 4a as

$$\delta_O \leq 2\left[1 - e^{-\Delta\lambda}\right], \qquad (45)$$

where $\Delta\lambda = \sum_i \Delta\lambda_i > 0$. For the over-mitigated error, $\widehat{\mathcal{E}}^{-1} \circ \mathcal{E}$ is not CPTP as illustrated in Fig. 4c, the upper bound increases as the layer number increases, as shown in Fig. 4b

$$\delta_O \leq e^{2\Delta\lambda_-} - 2e^{-\max\{\Delta\lambda, 0\}} + 1, \qquad (46)$$

where $\Delta\lambda_- \equiv \sum_i |\min\{\Delta\lambda_i, 0\}|$. Therefore, when considering the error model violation, the under-mitigated error channel outperforms the under-mitigated error channel.

For the error beyond the description with the Pauli diagonal error model, it can be randomly compiled into the Pauli diagonal error model by using the well-performed Pauli twirling technique. However, when the number N of shots for Pauli twirling is not sufficiently large, the inaccurate error model will yield a deviation from the estimation. This deviation will be suppressed with the increase of N as $\sim 1/\sqrt{N}$ according to the large number theorem, so it can be interpreted as a statistical error of the estimation. Given a tolerant precision δ of the estimation and a tolerant probability $\gamma \ll 1$ for the failure, the number of shots N should satisfies

$$N \gtrsim \frac{L^2}{\delta^2}(A - B\log\gamma), \qquad (47)$$

where $A, B \geq 0$ are constants for the specific error model. More detailed discussions can be found in Supplementary Note V.

Discussion

The implementability function is related to several quantities in quantum information theory and has further applications in quantum information processing. Here, we discuss the relationship between the implementability function and the diamond norm, logarithmic negativity, as well as purity. These results can reduce to the results about CPTP operations $\mathcal{Q}(A \to A)$ in other works. However, our technique does not stem from the semidefinite property of $\mathcal{Q}$ but uses the convexity and the free property of resource theories and can be applied to a general case.

We first consider the relationship between the implementability function and the diamond norm[38,40]. For a system A, given a convex free set of states $\mathcal{F}(A)$, the implementability function $p_{\mathcal{F}(A)}$ can be defined. Naturally, it induces a norm on the space of quantum operations $\mathcal{B}(A \to A)$

$$\|\mathcal{N}\|_{\diamond\mathcal{F}(A)} \equiv \max \frac{p_{\mathcal{F}(A)}(\mathcal{N}(\rho))}{p_{\mathcal{F}(A)}(\rho)}. \qquad (48)$$

This norm is a generalization of the diamond norm

$$\|\mathcal{N}\|_{\diamond} \equiv \max_{\rho \in \mathcal{Q}(A \otimes A)} \frac{\|id \otimes \mathcal{N}(\rho)\|_1}{\|\rho\|_1}, \qquad (49)$$

where the free set is $\mathcal{Q}(A \otimes A)$, since the implementability function $p_{\mathcal{Q}(A)} = \|\cdot\|_1$, with respect to the set of all quantum state $\mathcal{Q}(A)$ is the trace norm $\|\cdot\|_1$ (see Supplementary Note VII). The norm $\|\cdot\|_{\diamond\mathcal{F}(A)}$ is the Minkowski functional of the set $\mathcal{F}_{\max}(A \to A)$ of resource non-generating (RNG) operations[2], i.e., the maximal assignment of operations for the free set $\mathcal{F}(A)$ of state (see Supplementary Note VII for the proof)

$$\|\mathcal{N}\|_{\diamond\mathcal{F}(A)} = p_{\mathcal{F}_{\max}(A \to A)}(\mathcal{N}). \qquad (50)$$

Therefore, the diamond is the exponential of the physical implementability

$$\|\mathcal{N}\|_{\diamond} = p_{\mathcal{Q}(A \to A)}(\mathcal{N}). \qquad (51)$$

This result has also been proved with semidefinite programs[41], where $p_{\mathcal{Q}(A \to A)}$ is denoted as $\|\cdot\|_{\blacklozenge}$. We give a different proof (see Supplementary Note VII) without semidefinite programs, which is simpler.

In addition, we consider the relationship between the implementability function and the logarithmic negativity. The following results generalize the results of physical implementability[22]. The logarithmic negativity is defined as[42,43]

$$E_N(\rho) = \log\left\|\rho^{T_B}\right\|_1, \qquad (52)$$

which denotes the negativity of the partial transposed density matrix ρ^{T_B}, and is effective to detect entanglement[44,45]. Since $p_{\mathcal{Q}(AB)} = \|\cdot\|_1$, it is the logarithm of implementability function

$$E_N(\rho) = \log p_{T_B(\mathcal{Q})(AB)}(\rho) \qquad (53)$$

on the free set $T_B(\mathcal{Q})(AB)$ of partial transpose of quantum states, since the partial transpose $T_B(\rho) = \rho^{T_B}$ is a linear involution $T_B^2 = id_B$. Given a convex free set $\mathcal{F}(AB)$, for $\rho \in B(AB)$, we can define a generalized version $E_{N\mathcal{F}(AB)}$ of logarithmic negativity with respect to a convex set $\mathcal{F}(AB)$ as

$$E_{N\mathcal{F}}(\rho) = \log p_{T_B(\mathcal{F})(AB)}(\rho), \qquad (54)$$

where $T_B(\mathcal{F})(AB)$ is the partial transpose of the free set $\mathcal{F}(AB)$. With the so-defined generalized logarithmic negativity and the sub-multiplicity of implementability function, it follows that for $\rho = \mathcal{N}(\rho_0)$, where $\rho_0 \in \mathcal{B}(AB)$, $\mathcal{N} \in \mathcal{B}(AB \to AB)$

$$\begin{aligned} E_{N\mathcal{F}(AB)}(\rho) &- E_{N\mathcal{F}(AB)}(\rho_0) \\ &\leq \log p_{T_B(\mathcal{F})(AB \to AB)}(\mathcal{N}), \end{aligned} \qquad (55)$$

where the action of partial transpose on operations is induced as

$$T_B(\mathcal{N}) = T_B \circ \mathcal{N} \circ T_B. \qquad (56)$$

For the proof, see Supplementary Note VII. This inequality is tight when

$$p_{T_B(\mathcal{F})(AB \to AB)}(\mathcal{N}) = \|\mathcal{N}\|_{\diamond T_B(\mathcal{F})(AB)}, \qquad (57)$$

and by Eq. (50), i.e., the free set of operations is a resource nongenerating set

$$\mathcal{F}(AB \to AB) = \mathcal{F}_{\max}(AB \to AB). \qquad (58)$$

When the operation $\mathcal{N}$ is invariant under the partial transpose, namely

$$T_B \circ \mathcal{N} \circ T_B = \mathcal{N}, \qquad (59)$$

the inequality reduces to

$$E_{N\mathcal{F}(AB)}(\rho) - E_{N\mathcal{F}(AB)}(\rho_0) \leq \log p_{\mathcal{F}(AB \to AB)}(\mathcal{N}). \qquad (60)$$

In particular, it holds for the local operations $\mathcal{N} = \sum_i q_i \mathcal{N}_i^A \otimes \mathcal{N}_i^B$, which is invariant under the partial transpose.

Purity is the square of the Frobenius norm $\|A\|_2 = \sqrt{\mathrm{Tr}(A^\dagger A)}$. Let $\sigma \in \mathcal{B}(A)$, $\mathcal{N} \in \mathcal{B}(A \to A)$. For the Frobenius norm, we have

$$\frac{\|\mathcal{N}(\sigma)\|_2}{\|\sigma\|_2} \leq p_{\mathcal{F}(A)}(\mathcal{N}) \max_{l \in A^+ \cup A^-} \|\mathcal{F}_l\|_F, \qquad (61)$$

where $\|\cdot\|_F$ is an induced norm of the Frobenius norm

$$\|\mathcal{M}\|_F = \max_i \|\mathcal{M}_i\|_F = \max_i \max_{\sigma_i} \frac{\|\mathcal{M}_i(\sigma_i)\|_2}{\|\sigma_i\|_2}, \qquad (62)$$

where $\mathcal{M}_i$ is the Jordan decomposition of the operation $\mathcal{M}$ as the matrix acting on the vector space of state $\mathcal{B}(A)$. Here, $\mathcal{F}_l$ is the extreme points of the free set, where $\mathcal{N}$ has optimal decomposition on them. In particular, if the operation $\mathcal{N}$ is identity, i.e., $\mathcal{N}(I) = I$, denoting $\sigma = \rho - I/D$, we have

$$\log \frac{P(\mathcal{N}(\rho))D - 1}{P(\rho)D - 1} \leq 2 \log p_{\mathcal{F}(A)}(\mathcal{N}) \tag{63}$$
$$+ 2 \log \max_{l \in A^+ \cup A^-} \|\mathcal{F}_l\|_F.$$

Moreover, if $\mathcal{N} = \sum_l x_l \mathcal{U}_l$ is a unitary mixed operation, where the unitaries $\mathcal{U}_l \in \mathcal{F}(A \to A)$ are free, namely $\mathcal{F}_l = \mathcal{U}_l$, we have

$$\log \frac{P(\mathcal{N}(\rho))D - 1}{P(\rho)D - 1} \leq 2 \log p_{\mathcal{F}(A)}(\mathcal{N}). \tag{64}$$

For the proof, see Supplementary Note VII.

Conclusion

The implementability function $p_{\mathcal{F}}$ is the generalization of physical implementability to arbitrary convex free set $\mathcal{F}$, which evaluates the minimal cost to simulate a resource with the free resources by quasiprobability decomposition. We investigate the noisy realization of the PEC method based on the properties of the implementability function. We demonstrate the way to optimally simulate the inverse operation of the error channel with the noisy Pauli basis. This method requires that the noisy Pauli gates can be reversed into ideal Pauli gates. We give the condition under which the invertibility is guaranteed with a tolerant probability of failing under finite measurements in experiments. We also derive upper bounds of the bias of the noisy cancellation of error channel without canceling the noise on Pauli gates. It shows that to mitigate the error channels in circuits with many layers, canceling the total error directly has better performance than canceling the layer errors separately when the error rate and layer number are small. Otherwise, the separate cancellation has better performance.

Moreover, we discuss several quantum-information quantities that are relevant to the implementability function, including the diamond norm, logarithmic negativity, and purity. We propose a norm on quantum operations based on a given implementability function of state, which is a generalization of the diamond norm. This norm denotes the implementability function of the resource non-generating set of the operation. It directly implies that the diamond norm is the exponential of physical implementability, which has also been proved with semidefinite programs[41]. Moreover, we derive the relation between the implementability function and the logarithmic negativity and purity, which extends the existing results[22].

Our results are mainly based on the convexity of the free set. However, the composition of the operation may imply more properties for quantifying resources, which may be left to further research. We hope that our results will be useful for the probabilistic error cancellation method in practice. We also expect that the implementability function will have more applications in the field of quantum information processing.

Methods

Let $A, B, C, \dots$ denote the systems; the Hilbert spaces of these systems are $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C$. The set of bounded linear operators on the Hilbert space $\mathcal{H}_A$ is denoted as $\mathcal{B}(A)$, to which the density matrices of A system belong. The linear operation on the operators from $\mathcal{B}(A)$ to $\mathcal{B}(B)$ is denoted as $\mathcal{B}(A \to B)$.

The quantum resources form a convex set $\mathcal{Q} \subset \mathcal{B}/\mathbb{R}$ of space of normalized resources. Here, we focus on the quantum channels $\mathcal{Q}(A \to A)$, which are the completely positive and trace preserving (CPTP) quantum operations. For the noisy cancellation, the implementable channels can be expressed as $\Theta(\mathcal{Q})(A \to A)$, where Θ is the linear map from ideal channels to noisy channels. In this section, the symbol $(A \to A)$ for quantum operation will be omitted for simplicity.

The quasiprobability decomposition technique used in the PEC has also been used in other quantum information processing tasks, such as

entanglement forging[46], where the entangled states are simulated with separable states, and circuit knitting[47–49], where the nonlocal operations are simulated with local operations. Therefore, it is worth extending the definition of the physical implementability to any convex free set of quantum resources. (see Supplementary Note I for a brief introduction to the framework of quantum resource theories). In general, for a given convex set $\mathcal{F}$ of freely implementable resources, the implementability function can be defined as

$$p(N) = \inf_{\mathcal{F}} \sum_i |x_i|, \tag{65}$$

over all decomposition $N = \sum_i x_i E_i$ with $E_i \in \mathcal{F}$. Our definition does not take the logarithm, since the implementability function $p(N)$ so-defined is just the minimal number of free operations used to simulate a resourceful operation with quasiprobability decomposition technique. Meanwhile, this quantity is also related to the robustness measure R in the quantum resource theories[2,10,50,51] as

$$p = 2R + 1. \tag{66}$$

The robustness depicts the minimal amount of mixing with free resources to wash out all resources[51]. The implementability function operationally corresponds to the inverse problem of the robustness.

The infimum of the cost is attained if the convex set $\mathcal{F}$ is bounded closed

$$p(N) = \min_{\mathcal{F}} \sum_i |x_i| < \infty. \tag{67}$$

Hereafter, we assume that the free set $\mathcal{F}$ is a bounded closed convex set. Then, the implementability function has the following basic properties:
(i) Faithfulness

$$p(N) \geq 1, \text{ and } N \in \mathcal{F}, \text{ iff } p(N) = 1. \tag{68}$$

(ii) Sub-linearity

$$p(aN_1 + bN_2) \leq |a|p(N_1) + |b|p(N_2). \tag{69}$$

(iii) Composition sub-multiplicity

$$p(M \circ N) \leq p(M)p(N). \tag{70}$$

(iv) Tensor-product sub-multiplicity

$$p(M \otimes N) \leq p(M)p(N). \tag{71}$$

(v) If $\mathcal{F}_1 \subset \mathcal{F}_2$,

$$p_{\mathcal{F}_1} \geq p_{\mathcal{F}_2}. \tag{72}$$

In general, not all the resources can be implemented by the resources that can be freely implemented with quasiprobability decomposition. One example is that if the freely implementable resources are local operations and classical channels (LOCC), the CNOT gate is not implementable[22]. Therefore, we consider the affine space $\mathcal{A} = \text{aff}(\mathcal{F})$, and the vector space $V = \langle \mathcal{F} \rangle$ generated by the free set $\mathcal{F}$. Then, the implementability function $p_{\mathcal{F}}$ is the Minkowski gauge function[52] of the convex set $\mathcal{C} = \text{conv}(\mathcal{F} \cup - \mathcal{F})$, when constrained on the affine space $\mathcal{A} = \text{aff}(\mathcal{F})$, as

$$p_{\mathcal{F}} = p_{\mathcal{C}}|_{\mathcal{A}} \equiv \inf\{\alpha \geq 0 : N \in \alpha \mathcal{C}\}. \tag{73}$$

This implies that the implementability function is a ratio of the length of lines, which is an affine invariant preserved by affine transformations GA(V)[53]. In the vector space V, the implementability function is a norm, which is the so-called base norm for the free set $\mathcal{F}$[54,55]. Actually, for any norm $\|\cdot\|$, since its sub linearity, there is a balanced convex set $\mathcal{C}_{\|\cdot\|} = \{N : \|N\| \leq 1\}$, such that the norm is the Minkowski gauge function of $\mathcal{C}_{\|\cdot\|}$. Therefore, the implementability function may be potentially related to many norms widely used in quantum information theory. For the proofs of the properties above, see Supplementary Note II. For more mathematical properties and structures of the implementability function, see Supplementary Note III and Supplementary Note IV.

Data availability

Code availability

References
1. Nielsen, M. A. & Chuang, I. L. *Quantum computation and quantum information* (Cambridge University Press, 2010).
2. Chitambar, E. & Gour, G. Quantum resource theories. *Rev. Mod. Phys.* **91**, 025001 (2019).
3. Breuer, H.-P. & Petruccione, F. *The theory of open quantum systems* (Oxford University Press, USA, 2002).
4. Mohseni, M., Rezakhani, A. T. & Lidar, D. A. Quantum-process tomography: Resource analysis of different strategies. *Phys. Rev. A* **77**, 032322 (2008).
5. Wigner, E. *Group theory: and its application to the quantum mechanics of atomic spectra*, Vol. 5 (Elsevier, 2012).
6. Cai, Z. et al. Quantum error mitigation. *Rev. Mod. Phys.* **95**, 045005 (2023).
7. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
8. Endo, S., Benjamin, S. C. & Li, Y. Practical quantum error mitigation for near-future applications. *Phys. Rev. X* **8**, 031027 (2018).
9. Cai, Z. Multi-exponential error extrapolation and combining error mitigation techniques for NISQ applications. *npj Quantum Inf.* **7**, 80 (2021).
10. Takagi, R. Optimal resource cost for error mitigation. *Phys. Rev. Res.* **3**, 033178 (2021).
11. Takagi, R., Endo, S., Minagawa, S. & Gu, M. Fundamental limits of quantum error mitigation. *npj Quantum Inf.* **8**, 114 (2022).
12. Sun, J., Yuan, X., Tsunoda, T., Vedral, V., Benjamin, S. C. & Endo, S. Mitigating realistic noise in practical noisy intermediate-scale quantum devices. *Phys. Rev. Appl.* **15**, 034026 (2021).
13. Suzuki, Y., Endo, S., Fujii, K. & Tokunaga, Y. Quantum error mitigation as a universal error reduction technique: Applications from the NISQ to the fault-tolerant quantum computing eras. *PRX Quantum* **3**, 010345 (2022).
14. Strikis, A., Qin, D., Chen, Y., Benjamin, S. C. & Li, Y. Learning-based quantum error mitigation. *PRX Quantum* **2**, 040330 (2021).
15. Guo, Y. & Yang, S. Quantum error mitigation via matrix product operators. *PRX Quantum* **3**, 040313 (2022).
16. Piveteau, C., Sutter, D. & Woerner, S. Quasiprobability decompositions with reduced sampling overhead. *npj Quantum Inf.* **8**, 12 (2022).
17. Van Den Berg, E., Minev, Z. K., Kandala, A. & Temme, K. Probabilistic error cancellation with sparse Pauli–Lindblad models on noisy quantum processors. *Nat. Phys.* **19**, 1116–1121 (2023).
18. Pechukas, P. Reduced dynamics need not be completely positive. *Phys. Rev. Lett.* **73**, 1060 (1994).
19. Carteret, H. A., Terno, D. R. & Życzkowski, K. Dynamics beyond completely positive maps: Some properties and applications. *Phys. Rev. A* **77**, 042113 (2008).
20. Cai, Z. A practical framework for quantum error mitigation. arXiv: https://arxiv.org/abs/2110.05389 (2021).
21. Jiang, J., Wang, K. & Wang, X. Physical implementability of linear maps and its application in error mitigation. *Quantum* **5**, 600 (2021).
22. Guo, Y. & Yang, S. Noise effects on purity and quantum entanglement in terms of physical implementability. *npj Quantum Inf.* **9**, 11 (2023).
23. Wallman, J. J. & Emerson, J. Noise tailoring for scalable quantum computation via randomized compiling. *Phys. Rev. A* **94**, 052325 (2016).
24. Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X* **7**, 021050 (2017).
25. Cai, Z. & Benjamin, S. C. Constructing smaller Pauli twirling sets for arbitrary error channels. *Sci. Rep.* **9**, 11281 (2019).
26. Hashim, A. et al. Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor. *Phys. Rev. X* **11**, 041039 (2021).
27. Koczor, B. Exponential error suppression for near-term quantum devices. *Phys. Rev. X* **11**, 031057 (2021).
28. Erhard, A. et al. Characterizing large-scale quantum computers via cycle benchmarking. *Nat. Commun.* **10**, 5347 (2019).
29. Helsen, J., Roth, I., Onorati, E., Werner, A. & Eisert, J. General framework for randomized benchmarking. *PRX Quantum* **3**, 020357 (2022).
30. Flammia, S. T. & Wallman, J. J. Efficient estimation of Pauli channels. *ACM Trans. Quantum Comput* **1**, 32 (2020).
31. Harper, R., Flammia, S. T. & Wallman, J. J. Efficient learning of quantum noise. *Nat. Phys.* **16**, 1184 (2020).
32. Ferracin, S. et al. Efficiently improving the performance of noisy quantum computers. *Quantum* **8**, 1410 (2024).
33. Guimarães, J. D., Lim, J., Vasilevskiy, M. I., Huelga, S. F. & Plenio, M. B. Noise-assisted digital quantum simulation of open systems using partial probabilistic error cancellation. *PRX Quantum* **4**, 040329 (2023).
34. Govia, L. et al. Bounding the systematic error in quantum error mitigation due to model violation, arXiv:*2408.10985* (2024).
35. Zhao, P. et al. Quantum crosstalk analysis for simultaneous gate operations on superconducting qubits. *PRX Quantum* **3**, 020301 (2022).
36. Knill, E., Laflamme, R. & Zurek, W. Threshold accuracy for quantum computation. arXiv: https://doi.org/10.48550/arXiv.quant-ph/9610011 (1996).
37. Aharonov, D. & Ben-Or, M. Fault-tolerant quantum computation with constant error, in https://doi.org/10.1145/258533.258579 *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97 (Association for Computing Machinery, New York, NY, USA, 1997).
38. Kitaev, A. Y. Quantum computations: algorithms and error correction. *Russ. Math. Surv.* **52**, 1191 (1997).
39. Terhal, B. M. Quantum error correction for quantum memories. *Rev. Mod. Phys.* **87**, 307 (2015).
40. Watrous, J. *The theory of quantum information* (Cambridge University Press, 2018).
41. Regula, B., Takagi, R. & Gu, M. Operational applications of the diamond norm and related measures in quantifying the non-physicality of quantum maps. *Quantum* **5**, 522 (2021).
42. Vidal, G. & Werner, R. F. Computable measure of entanglement. *Phys. Rev. A* **65**, 032314 (2002).
43. Plenio, M. B. Logarithmic negativity: A full entanglement monotone that is not convex. *Phys. Rev. Lett.* **95**, 090503 (2005).

44. Peres, A. Separability criterion for density matrices. *Phys. Rev. Lett.* **77**, 1413 (1996).
45. Horodecki, M., Horodecki, P. & Horodecki, R. Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A* **223**, 1 (1996).
46. Eddins, A. et al. Doubling the size of quantum simulators by entanglement forging. *PRX Quantum* **3**, 010309 (2022).
47. Mitarai, K. & Fujii, K. Constructing a virtual two-qubit gate by sampling single-qubit operations. *N. J. Phys.* **23**, 023021 (2021).
48. Mitarai, K. & Fujii, K. Overhead for simulating a non-local channel with local channels by quasiprobability sampling. *Quantum* **5**, 388 (2021).
49. Piveteau, C. & Sutter, D. Circuit knitting with classical communication. *IEEE Trans. Inf. Theor.* **70**, 2734–2745 (2023).
50. Brandão, F. G. S. L. & Gour, G. Reversible framework for quantum resource theories. *Phys. Rev. Lett.* **115**, 070503 (2015).
51. Vidal, G. & Tarrach, R. Robustness of entanglement. *Phys. Rev. A* **59**, 141 (1999).
52. Osborne, M. S. & Osborne, M. S. *Locally convex spaces* (Springer, 2014).
53. Prasolov, V. V. & Tikhomirov, V. M. *Geometry*, Vol. 200 (American Mathematical Society, 2001).
54. Hartkämper, A. & Neumann, H. *Foundations of Quantum Mechanics and Ordered Linear Spaces: Advanced Study Institute Marburg 1973* (Springer, 1974).
55. Regula, B. Convex geometry of quantum resource quantification. *J. Phys. A* **51**, 045303 (2017).
56. Jin, T.-R., Zhang, Y.-R., Xu, K. & Fan, H. Numerical data and codes for "noisy probabilistic error cancellation and generalized physical implementability, *Figshare* https://doi.org/10.6084/m9.figshare.29290544.v1 (2025).

Author contributions
H.F. and Y.-R.Z. supervised the project; T.-R.J. proposed the idea; T.-R.J. performed the numerical simulations and discussed with Y.-R.Z. and K.X.; T.-R.J., Y.-R.Z. and H.F. co-wrote the manuscript, and all authors contributed to the discussions of the results and development of the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42005-025-02217-8.

Correspondence and requests for materials should be addressed to Yu-Ran Zhang or Heng Fan.

Peer review information *Communications Physics* thanks Luke Govia and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.